Pierre Pontarotti *Editor*

# Evolutionary Biology: Exobiology and Evolutionary Mechanisms

Springer

# Evolutionary Biology: Exobiology and Evolutionary Mechanisms

Pierre Pontarotti
Editor

# Evolutionary Biology: Exobiology and Evolutionary Mechanisms

Springer

*Editor*
Dr. Pierre Pontarotti
CNRS, Laboratoire Evolution Biologique et Modélisation
Université d'Aix-Marseille
Marseille
France

ISBN 978-3-642-38211-6 ISBN 978-3-642-38212-3 (eBook)
DOI 10.1007/978-3-642-38212-3
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013943333



Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

# Preface

For the 16th time, the Evolutionary Biology Meeting at Marseilles (EBM) took place. The goal of this annual meeting was to allow scientists of different disciplines, who share a deep interest in evolutionary biology concepts, knowledge and applications, to meet, exchange and start interdisciplinary collaborations.

The EBM at Marseilles is now recognised internationally as an important exchange platform and a booster for the use of evolutionary-based approaches in biology and also in other scientific areas.

This year more than 100 presentations were selected by the EBM scientific committee. These presentations really reflected the epistemological positioning of the meeting. We have selected 19 of the most representative ones for the book.

The book will give the reader an overview of the state of the art in the evolutionary biology field. The book is the sixth that we have published further to the meeting. I would like to underline that the six books are complementary to each another and should be considered as tomes.

The reader of evolutionary biology books as well as the meeting participants would maybe, like me, witness years after years during the different meetings and book editions a shift in the evolutionary biology concepts. Obviously, this shift is not only specific to this meeting but also reflects a general trend in the evolutionary biology community and the scientific community.

We shifted slowly in a few years from bifurcative phylogenetic trees thinking of evolutionary processes to a phylogenetic network thinking (Chaps. 10 and 11) indicating that the part of horizontal evolution should not be neglected. The HGT hybridization and the speciation process do not lead to clear-cut separation between species but to more or less mosaic evolution (incomplete lineage sorting).

In regard to the HGT concept, scientists interested in the origin of life questions do not think any more or much less of the Last Universal Common Ancestor hypothesis (concept linked to a tree view of evolution) but to a world with independent protocells, each of them bringing its own metabolic contribution to the protocells community (Chap. 6).

Another new discovery that will be important to shifts in evolutionary biology thinking is the revolutionary concept of transgenerational epigenetic showing that heredity is not only based directly on the genetic code (Chap. 13).

I am still puzzled that the adaptationist thinking stays very much alive and the fact that evolutionary models (mathematics and informatics) rely strongly on that.

In fact, the participation of the adaptive evolution on evolutionary shift is likely to be minor, but this is my hypothesis; let us see what will happen in the future and especially at the EBM.

As for the last book, we start the book with a chapter, then explain the history of great discoveries, and the life and scientific contribution of major contributors in the field.

The following articles are organised in the following categories:

Evolutionary Biology Concepts
Exobiology and Origin of Life
Evolutionary Mechanisms

I would like to thank all the authors, the meeting participants, the sponsors: Aix Marseille Université, Conseil Général 13, GDR BIM, ITMO, Region PACA, Ville de Marseille.

I also wish to thank Springer's edition staff and in particular Andrea Schlitzberger for her competence and help.

I wish also to thank members of the Association pour l'Etude de l'Evolution Biologique (AEEB) and the members of the Evolutionary Biology Modelling Laboratory.

Finally, I thank the AEEB coordinator: Marie Hélène Rome for the organisation of the 16th EBM and her help with the book.

Pierre Pontarotti
President of the AEEB

# Contents

# Part I
# History of Great Discoveries

# Chapter 1
# The First Experiments on Ascidian and Sea Urchin Eggs Fertilization

**Margherita Raineri and Erki Tammiksaar**

**Abstract** In 1845 Karl Ernst von Baer made the first experiments on artificial fertilization of ascidian and sea urchin eggs in Cornigliano near Genoa (Italy). These studies were continued in Venice and Trieste and the results were published in 1847 as a preliminary note. von Baer considered the eggs of sea urchins to be an excellent material for embryological investigations. He described the formation of the fertilization envelope, the nuclear migration to the center of the egg, and the nuclear divisions which determined the pattern of the cytoplasmic divisions. These observations led to the conclusion that all nuclei of larval and adult tissues are derived from the nucleus of the fertilized egg. The 1847 preliminary note did not include figures, but von Baer illustrated his findings with a pencil drawing that is published here for the first time.

## 1.1 Introduction

The popularity of ascidian and sea urchin eggs as suitable models for investigating fertilization and embryogenesis dates back to the second half of the nineteenth century. After the initial descriptive studies the focus of interest shifted to experimental embryology, addressing problems such as the block to polyspermy, the segregation of cytoplasmic materials and the metabolic changes at fertilization, the establishment of the body axes, the developmental potentialities of isolated

M. Raineri (✉)
Dipartimento di Scienze della Terra, dell'Ambiente e della Vita,
University of Genoa, 16132 Genoa, Italy
e-mail: raimrg@unige.it

E. Tammiksaar
Estonian University of Life Sciences, University of Tartu, Baer House,
Estonia, Europe
e-mail: erki.tammiksaar@emu.ee

P. Pontarotti (ed.), *Evolutionary Biology: Exobiology and Evolutionary Mechanisms*,
DOI: 10.1007/978-3-642-38212-3_1, 

blastomeres, and the dynamics of cell-to-cell interactions. As a result of these investigations, the eggs of sea urchins and ascidians came to be considered, respectively, as examples of regulative and mosaic development in every textbook on embryology. In addition, ascidians gained a central position in evolutionary thinking following the embryological studies of Kowalevsky (1866) and their interpretation by Haeckel (1868) and Darwin (1871) who considered them to be chordates, the closest invertebrate relatives of the vertebrates. At the present time, ascidians and sea urchins appear to be excellent material for studies of evolutionary developmental biology (EvoDevo) (Raff and Love 2004). In particular, embryonic and larval ascidians are investigated extensively to elucidate the genetic basis of chordate homologies and the developmental changes that might have led to the origin of vertebrates.

When considering the enormous amount of current and past publications, it seems surprising that the researcher who first performed successful artificial fertilization of ascidian and sea urchin eggs and described their initial development is ignored almost completely by the scientific literature. As far as sea urchins are concerned, recently Briggs and Wessel (2006) have drawn attention to three papers which anticipated for decades the report of Hertwig (1876) often considered to be the first truly significant account of these processes. These pioneering researches were carried out by von Baer, Dufossé, and Derbès, and all of them were published in 1847. In their review, however, Briggs and Wessel (2006) analyze almost exclusively the contribution of the latter author in comparison to that of Dufossé. Evidently, they could not retrieve the paper of von Baer, which is most probably the first published record of sea urchin fertilization and development. Its French abstract appeared on 26 May 1847 and was cited by Derbès, who published his own observations on August 1847, a few months after Dufossé. The same paper of von Baer (1847) also contains the first report on ascidian artificial fertilization. This priority was acknowledged by Krohn (1852) in a study of the embryogenesis of ascidians preceding that of Kowalevsky (1866). For all these reasons, we believe that this paper as well as the circumstances of its preparation should be rescued from oblivion. In his autobiography, however, von Baer (1865) gives an extremely brief account of the above-reported experiments; he only specifies that they were made in Trieste and Genoa in 1845 and 1846. The biographies of Stieda (1878) and Raikov (1968) provide a more detailed survey of these trips to Italy laying emphasis on their importance to the life of von Baer from a scientific and historical point of view. These studies, however, do not explain why he took up embryology again after he quit these studies in 1834 when he shifted from Königsberg to St. Petersburg.

## 1.2 Traveling from the North to the Mediterranean Sea

In the first half of the nineteenth century studies on marine biology were still in embryonic conditions, hence, visiting scientists coming to the Mediterranean Sea usually went to places where they already had connections with local students and/

or existing colonies of compatriots which could give information about the available research materials, provide facilities, etc. For all these reasons, their favored destinations in Italy were Nice on the Ligurian Sea, Trieste on the North of the Adriatic Sea, and Naples and Messina on the South of the Tyrrhenian Sea. In later years, indeed, Naples and Trieste were chosen to build well-equipped marine biological stations (Groeben 2008; Casellato 2008).

As reported by Stieda (1878), however, in 1845 von Baer went to another destination which never became a leading center of research on marine fauna. His return to embryology in this unusual location appears to be the contingent result of a rather complicated story.

After he settled permanently in St. Petersburg in 1834, von Baer realized that he would not be allowed to go abroad in the nearest future, since his return to Königsberg after his "academic interlude" in St. Petersburg in 1830 had raised some doubts that he would like to escape again for a longer period (Tammiksaar and Brauckmann 2004). Yet, at the beginning of 1845 his position as a member of the Imperial Academy of Sciences appeared to be so secure that he applied for a five-month permission to travel abroad (Raikov 1968). von Baer indicated Berlin, Breslau, Vienna, and London as the main destinations, and the study of histology, teratology, and craniological collections as the scientific program. If there were time enough, the trip could also include a visit to the western part of Germany and some educational institutions of Northern Italy.

Quite independently from this project, a summer journey to Italy was planned on April 1845 on request of the Grand Duchess Helene Pavlovna, wife of Michail Pavlovich, the youngest brother of Czar Nicholas I (Grünewaldt-Haackhof 1900, p. 218). This tour was organized for her two daughters Marie and Katharina for educational purposes, and von Baer was asked to take part in it as their home teacher since the end of the 1830s. Being a close friend of the Grand Duchess, von Baer neither could, nor did he want to refuse, but he sent his own trip plan to Helene Pavlovna in a letter of 12 April. He strongly recommended staying at Castellamare near Naples to take advantage of the hospitality of the King of Naples, and besides, as a personal benefit, of the possibility of attending the VII Conference of the Italian Scientists that was going to be held in Naples in September 1845.[1] However, to take part in a meeting would have not been sufficient reason for submitting to the Academy of Sciences such a radical change in his traveling plan. Not to discredit himself, von Baer had to mention some other serious scientific purpose. A developmental study of sea urchins in the Mediterranean Sea appeared to be appropriate, his eminence as a scientist being largely due to his previous research on animal embryology.

von Baer intended to leave St. Petersburg at the beginning of June, but he did not receive the Czar's permission. As he wrote to Helene Pavlovna, he was

[1] State Archives of the Russian Federation (Moscow), 647-1-3, sh. 1–2v. (von Baer to the Grand Duchess Helene Pavlovna, Sankt-Petersburg, 31. 03./12. 04. 1845).

convinced that the approval to his travel application was intentionally delayed.[2] The Grand Duchess, who was already in Europe, was so upset by this delay that she offered to cover the expenses of von Baer's trip to Italy. She held von Baer in high esteem as an educated and intelligent conversationalist and fellow traveler and their affinities became even closer as they had to face a similar tragedy, since on February 1845 Helene Pavlovna lost her daughter Elizabeth just as von Baer had lost his son Karl in 1843. On invitation of the Grand Duchess, von Baer arrived in Genoa on 8 September 1845, together with Otto von Madai, who had been formerly personal assistant to Elizabeth. Then, they went to the nearby village of Cornigliano, where Helene Pavlovna was taking a holiday of several weeks together with her daughters Marie and Katharina, her brother Prince August von Württemberg, Otto von Grünewaldt and Duke Adolf von Nassau, the husband of the departed Elizabeth (Stieda 1878). All these persons lived in Villa Serra, while von Baer, von Madai, and von Grünewaldt lived at Villa Brignole. They visited the Grand Duchess every day, in particular, von Madai and von Baer had to keep her company, respectively, for two and two hours and a half (Grünewaldt-Haackhof 1900, pp. 224, 226). These documents do not specify how much time von Baer could devote to embryological research. In any case, he performed successful artificial fertilization of ascidian and sea urchin eggs at the beginning of September[3] (the end of August following the Russian Calendar, von Baer 1847), as soon as he arrived at Cornigliano, and continued these studies in the following days (Grünewaldt-Haackhof 1900, p. 229). On 25 September he went to Florence together with von Madai, Prince von Württemberg, and Duke von Nassau for a short cultural visit. On 2 October they came back to Genoa to rejoin the Grand Duchess, who was living in a palace situated "in the highest point of the city" (Preller 1850). On 8 October Helene Pavlovna left Genoa together with her court, but von Baer stayed behind to devote himself completely to his embryological studies. These investigations fascinated him so much that he prolonged his stay as long as possible (letter of von Baer to Theodor L.W. Bischoff, Anonymous 1880). Just as in Cornigliano, while being in Genoa he had certainly some efficient assistance of local persons, since he regularly got seawater and sea urchin specimens from the port. The problem was that the commonest species, *Echinus lividus* Lam. (now *Paracentrotus lividus*), had immature eggs, while the bigger one, *Echinus brevispinosus* Risso (now *Sphaerechinus granularis*), had mature eggs, albeit less transparent, but it was much more rare and nearly impossible to collect when the sea was stormy (von Baer 1847). When also this species stopped to spawn in the Ligurian Sea, von Baer decided to shift to another seashore in order to continue his studies.

[2] State Archives of the Russian Federation, 647-1-697, sh. 1–2v. (von Baer to the Grand Duchess Helene Pavlovna, Sankt-Petersburg, 28. 05./09. 06. 1845).

[3] The Julian Calendar that was used in the Russian Empire followed the Gregorian Calendar with a delay of 13 days.

After his aristocratic interlude in Genoa, he embarked on a more adventurous exploration trying to find "that special kind of information which is essential to locate immediately the most suitable place" (von Baer 1847). Eventually, the Italian zoologist Filippo De Filippi, then at the University of Milan, told him that in Venice Dr. Giandomenico Nardo could provide all the facilities and materials for his research. von Baer left Genoa in the second half of October and on the very day of his arrival in Venice he got several specimens of *Echinus brevispinosus*, which was still spawning in the Adriatic Sea. He performed successful artificial fertilization in his hotel room, but the day after, the chambermaid thought that the bowls contained dirty water and threw all the embryos away. von Baer had not much time left to finish his studies. Since he could not obtain immediately other *Echinus* specimens, he went straight to Trieste hoping to find some German-speaking fishermen who could provide sea urchins. All his efforts came to nought, but the local newspaper *Horae Tergestinae* gave him some information about Heinrich Koch, a merchant who was interested in the Adriatic fauna. With the help of Koch, von Baer got every day many specimens of *E. brevispinosus* as well as *E. lividus*. The latter had more transparent eggs which allowed him to complete his previous observations on fertilization and embryonic development. The results were communicated to the Imperial Academy of Sciences of St. Petersburg on 13 November 1845, and a précis of this report from Trieste was published in 1847 without any figure, with the following preliminary remarks: "The author had expressed the wish this essay not to be published before he knew for sure that this year (1846) he could go once more to the Mediterranean Sea. Shortly before his departure he has added some information that will be completed subsequently in more detail." The research, then, was still in progress, but von Baer never published a full account of these studies. In the following paragraph we summarize the 1847 preliminary note illustrated by a pencil drawing of von Baer (Fig. 1.1) that was found in the archives and is published here for the first time.

My 100,000 sea urchins by myself procreated

These words of von Baer (1847) complaining about the inopportune chambermaid of Venice illustrate his characteristic wit as well as his complete commitment to his embryological studies.

At the beginning of his report from Trieste he tried to explain tactfully the reasons behind his decision to carry out a completely different trip to Italy as proposed originally to the Academy of Sciences. Of course, he did not say a word about the Grand Duchess and her role in the trip, nor did he give information about his living conditions and personal contacts in Cornigliano and Genoa. Instead, he laid emphasis on the fact that his Italian journey was aimed at addressing a scientific problem in the best experimental conditions.

To his surprise, he could not find in the literature any report on artificial fertilization of animals which spawn in seawater and develop outside the mother's body, although they appeared to be a very convenient material for such experiments. Thus, he planned to go to the seaside and do research in this field in order to observe hitherto unknown types of development. For this purpose, the

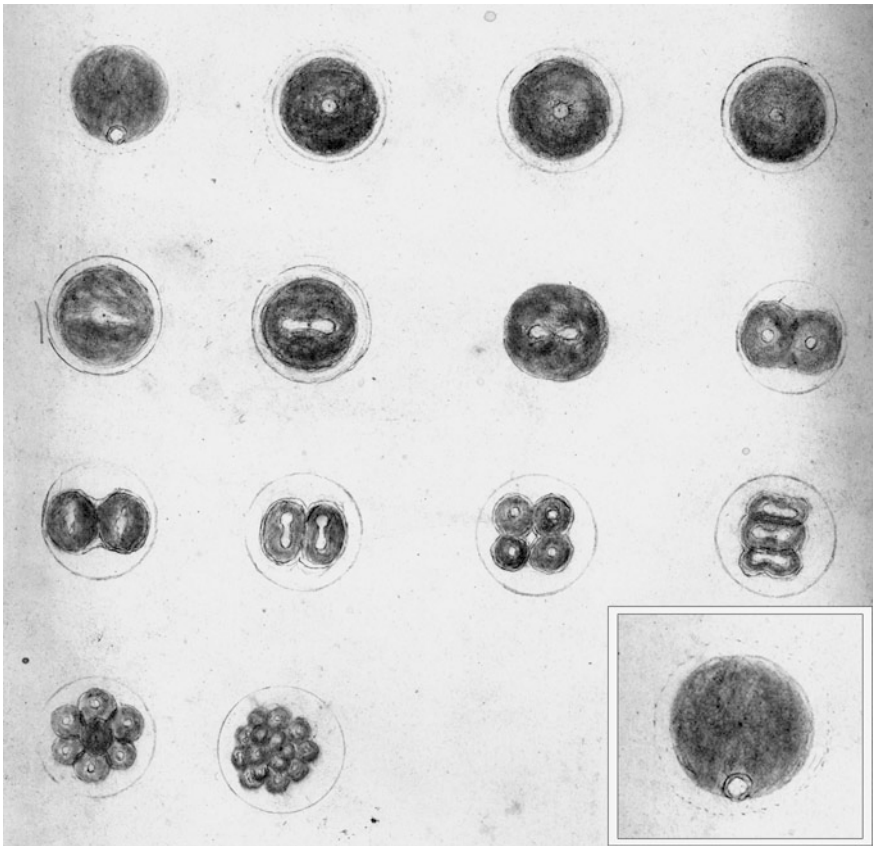


**Fig. 1.1** This pencil drawing made by von Baer was found without any legend. It shows the nuclear migration in the fertilized egg of sea urchin (*upper row*, from *left* to *right*) and the following stages of division. The supernumerary spermatozoa at fertilization are represented as dashes adherent to the fertilization envelope which appears as a line encircling the egg (see enlarged detail on the *right*, *bottom* of the figure). Courtesy of the Universitätsbibliothek Giessen, Nachlass von Baer, Band 22, Bl. 8

Mediterranean Sea appeared to be more promising than the Baltic Sea, the latter being more similar to a lagoon from the zoological point of view. Another reason for coming to Italy was the mild climate which in summer caused the room temperature not to differ much from that of the seawater. Taking into account the temperature sensitivity for development, the great difference between the outer frost and heated rooms was likely to be the reason for the failure of his attempts to investigate the embryology of fish in Königsberg as well as in St. Petersburg.

As reported above, he performed the first experiments on artificial fertilization of ascidian eggs at the beginning of September 1845. “Within a few hours these eggs showed the well-known yolk divisions and before the end of a day of 24 h the embryos left the egg envelope. In their external shape they were alike huge

cercarians with real or seeming eyespots and swam around very actively."He was eager to see the metamorphosis of these larvae, but all of them died a few hours later. The same day he performed artificial fertilization of sea urchin eggs. Hatching occurred after approximately 16 h and the ciliated embryos looked somewhat alike a larval medusa. Later, they changed in shape, "seemingly approaching at first the body plan of *Beroe* (a ctenophoran)," but on the fourth day they got an irregular shape and on the fifth day none of them was still alive.

This negative outcome was ascribed by von Baer to the lack of food, especially in the case of the ascidian larvae. Another problem was the small size of sea urchin embryos. Since von Baer had only a portable microscope without Mess apparatus, they were difficult to collect and got lost very easily in larger bowls, but they died earlier in smaller ones. He made several attempts to overcome these problems. For instance, to provide a regular supply of seawater containing food he put the eggs inside a bucket anchored on the seashore with a small opening that was closed by a net. In spite of these difficulties, while being in Genoa von Baer got a special interest in the development of sea urchins because "in the inner region of the egg just after fertilization, as well as on occasion of every yolk division, it was often possible to distinguish a clearer aspect that expanded lengthwise and remained visible for a while, but soon seemed to vanish again. A slight pressure applied to the egg allowed to detect an inner vesicle (or nucleus), or two of them facing each other…. When I marked on the stage the direction of the longitudinal axis of the inner clearer region, I could realize that the subsequent division of the yolk sphere always occurred in that direction, the separation between the two halves crossing the middle of that axis. No doubt the process of yolk division was anticipated and conditioned from the inside." Later, in Trieste von Baer observed these processes in more detail and also succeeded in keeping the larvae alive for a longer time, but never up to the adult form. "In the more transparent eggs (of *E. lividus*) I followed the process of yolk division until the formation of the swimming embryo, not only in its outer appearance, but also in its inner changes to a nearly complete extent, and moreover, in the same specimen, not only in adjacent eggs." He described fertilizations as follows. In the peripheral region of the egg there is a "clear circle approximately 1/8 the diameter on the entire egg …which sinks deeper inside as soon as the yolk sphere gets some motion inside a peripheral transparent envelope following the absorption of the sperm fluid" (Fig. 1.1). By manipulating the eggs in different ways, von Baer came to the conclusion that this circle, corresponding to the germinal spot or vesicle that was described in other eggs, was neither a spot, nor a vesicle or a cell. Von Baer called this body the nucleus of the egg. Contrary to the statement of Bagge (1841) and Kölliker (1843), it never disappeared completely, but a few minutes after fertilization it was no more visible at the surface since it moved towards the center of the egg. "Here the nucleus, until now spherical in shape, takes some period of rest, and then, it lengthens rather quickly in the meanwhile as it seems to sprout at both sides; both ends swell, but the middle part becomes thinner and soon divides completely, so that two comet-shaped nuclei lie with their tails opposite one another. Then, very quickly, the tail-shaped appendages pull themselves back into the spherical or vesicular masses,

and one has two nuclei. The appendages remain longer in sickly eggs. The original nucleus had already increased in volume before division; the same occurs to an even greater extent during division, so that each of the two new nuclei has approximately the size of the original one…. The two new nuclei remain for a while in close vicinity, then, they get apart one from the other, and eventually, a constriction of the yolk appears which divides it in two halves adherent one another, so that each nucleus now has his envelope of yolk substance…. Soon later, each of the two nuclei begins to sprout out in the same way and by dividing in the middle becomes transformed into two nuclei round which the yolk mass soon divides, and the whole egg becomes divided into four masses adhering to one another. Each quadrant changes into a round shape, so that immediately or during the next division the middle of the egg becomes empty. Quite similarly there follows the division of the quadrants, and more precisely in such a way that the direction of the new sprouts is perpendicular to the immediately preceding one. So it goes on with next divisions, a nucleus forming in advance for each portion of yolk by division of one that was produced earlier. These inner masses I have called nuclei seemed to me to be definitely fluid—a membranous envelope could form each time after a period of rest. At least, I noticed that a boundary line seemed to appear under the microscope that was sharper during the period of rest and could not be detected for sure during the period of sprouting. In the subsequent divisions the mass of yolk surrounding each nucleus becomes smaller, so that it is possible to distinguish more accurately that new cells never form inside a mother-cell by nuclear division. I could observe continuously these divisions up to 32 yolk bodies (up to this stage the process occurs quite regularly) (Fig. 1.1).” In more advanced stages the microscopic view was blurred by shadows, but in peripheral yolk bodies it was still possible to detect the same process of division. “Eventually, when the embryo has left the egg envelope and is swimming by means of cilia, each granule or histogenetic element (vulgarly called cell) has a very evident nucleus, and they all appear to be derived from the original nucleus of the egg. The following changes could not be observed directly owing to the movements of the larva. However, I have grounds for believing that also the permanent tissue constituents arise from the original ones by quite similar divisions. As a consequence, the yolk divisions were nothing but the beginning of the histogenetic differentiation which continued up to the eventual formation of the animal.

If this view is correct, no doubt how to answer the question about the pre-existence of the new individual before fertilization. The unfertilized egg is the embryo in latent life. Fertilization makes life active.”

## 1.3 The Embryos of Sea Urchins and the Cell Theory

Let us analyze the essay of von Baer (1847) starting from his account of fertilization. Unlike Derbés (1847), he did not mention either the jelly layer, or the spermatozoa, although he noticed that the embryos may be surrounded by

"secretions" and saw sperm clustering around the eggs (Fig. 1.1). However, without the benefits of higher magnifications and staining techniques it was impossible to distinguish in vivo the sperm entry inside the egg and the exocytosis of the cortical granules. Thus, von Baer limited his description to the observed facts and ascribed the earliest visible changes, i.e., the appearance of the fertilization envelope and the hyaline layer around the egg, to swelling due to the absorption of diluted seminal fluid.

In a following stage he paid considerable attention to the "metamorphosis" of the germinal vesicle. He called this transparent, spherical body the nucleus of the egg, although subsequently "one could be doubtful about this interpretation," because this body "came to occupy a considerable part of the egg." Later, however, he observed similar changes, as every "yolk division which could be seen from the outside was always preceded by the division of an inner nucleus which at first was transparent, but subsequently seemed to be completely fluid." Of course, the latter condition corresponded to the mitotic spindle, which could not be detected in living unstained eggs (Fig. 1.2). The great intuition of von Baer was that the nucleus or its substance never dissolves in the surrounding yolk, but splits into two equal parts. As a consequence, nuclei do not appear ex novo during development, but all of them derive ultimately from the nucleus of the fertilized egg. He stuck to this opinion in spite of the opposite view of Bagge (1841) and Kölliker (1843). In particular, the latter author asserted that in the eggs of nematodes "soon after fertilization germinal spot and germinal vesicle disappear completely, and the clear, transparent yolk contains nothing but sparse elementary granules (Elementarkörner)." It seems quite possible that Kölliker saw the chromosomes, as the eggs of *Ascaris* and other nematodes are fertilized before meiosis, hence, the germinal vesicle breaks down twice after fertilization. However, he mistook them completely, and owing to this error he failed to understand the true nature of the blastomeres. In his opinion, they were nothing but conglomerates of yolk containing Embryonalzellen, as a matter of fact, the nuclei, which he considered to be the progenitors of the definitive cells. To make things even worse, he thought that Embryonalzellen proliferated by endogeny, producing in their inside two small new cells which got free when the mother-cell dissolved. If we consider that Schleiden and Schwann, the putative fathers of the cell theory, also supported an endogenous or exogenous origin of the new cells (Baker 1953, 1955), we can fully appreciate the importance of the conclusion of von Baer (1847) that "new cells never form inside a mother-cell by nuclear division."

Conceivably, there was an element of luck in coming across the eggs of sea urchins which complete meiosis during maturation. As a consequence, it was much easier than in nematodes to interpret the centripetal migration of the germinal vesicle and the following nuclear division as a continuous process, the only missing (but crucial) detail being the contribution of the male pronucleus to the zygote. The favorable peculiarities of the sea urchin eggs allowed von Baer to give an account of nuclear multiplication that was "far fuller and more accurate than anything that had been published before"(Baker 1955) (Figs. 1.1 and 1.2). In relation to this, he made a fruitful generalization which gave him an insight into

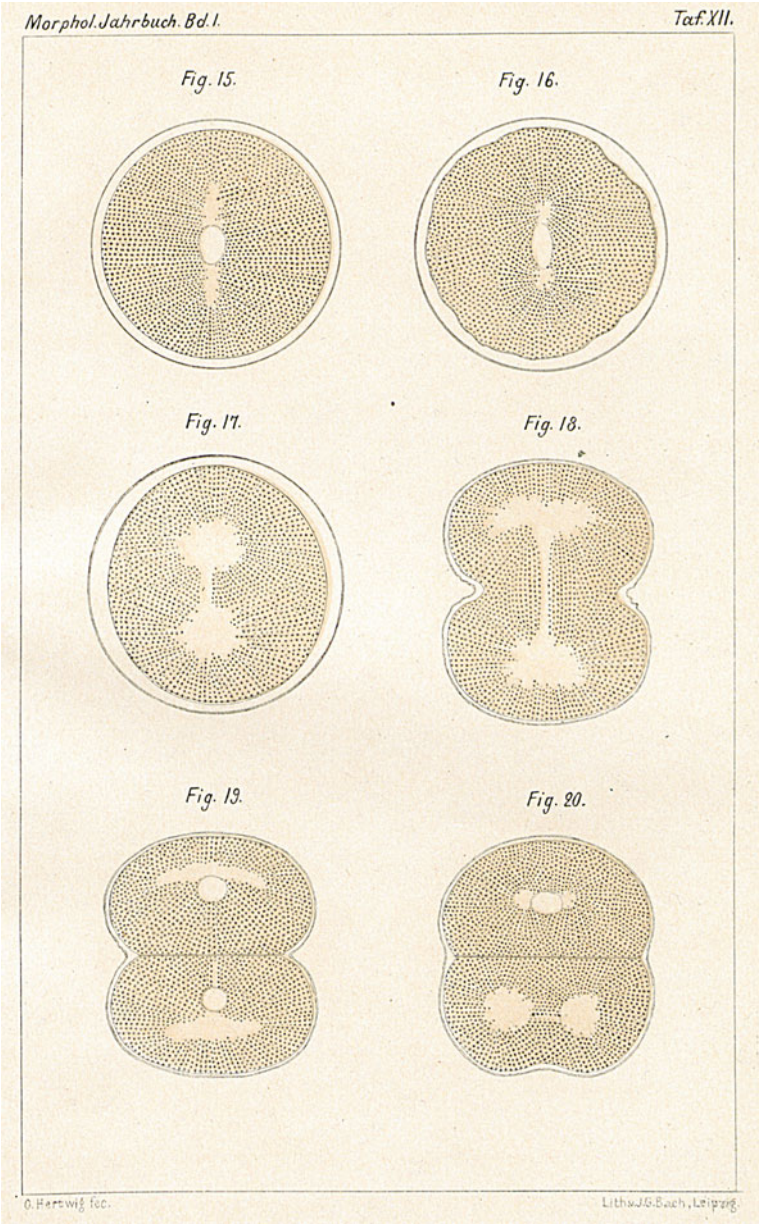



**Fig. 1.2** Hertwig's (1876) plate showing the first division of the sea urchin egg corresponds precisely to the description of von Baer (1847)

the nature of the blastomeres in every type of embryos. "I believe I have found the germinal vesicle also in the frog fertilized egg, but so condensed and perhaps devoid of an envelope, that it could not be described as a well-defined entity" (von

Baer 1847). This description corresponds precisely to the mature egg of the frog, which is arrested in meiotic metaphase II. As reported by von Baer (1865) in his autobiography, in the frog he could not detect the changes of the germinal vesicle and the following nuclear divisions because the eggs were too dark and large, and moreover, the nuclei dissolved in the acids he used to separate the cleavage products. This technique, however, demonstrated definitely that developing eggs do not undergo a "cleavage process," i.e., formation of furrows, but they are divided into smaller yolk bodies (von Baer 1834). Later, the study of the translucent eggs of sea urchins showed that each of these bodies contains its nucleus which derives from the nucleus of the fertilized egg by a continuous process of nuclear division preceding and conditioning yolk division.

The priority of the nucleus prevented von Baer from making the errors of Kölliker (1843) and other embryologists, including the above-mentioned Krohn (1852) who asserted that in dividing eggs of ascidians the nucleus disappears, and then, appears again in each newly formed blastomere. The same priority, however, might have contributed to outline a misleading parallelism between nuclear and cytoplasmic division as far as it concerns the behavior of the outer envelope. "I believe I have seen without any doubt that (in the eggs of sea urchins) on occasion of each division the products are devoid of outer membrane or covering of any kind, but very soon, after some period of rest, a very thin, somewhat discrete outer layer similar to a covering membrane can be distinguished, which vanishes again on occasion of each novel division" (von Baer 1865). At present, we know that the cell membrane, unlike the nuclear envelope, never disassembles, though, it is a fluid and dynamic layer. On the contrary, in the mid-1840s most cytologists were influenced by the work of Schwann (1839), who regarded as essentially identical plant and animal cells focussing attention on the cell wall and neglecting the nucleus. In a sense, von Baer (1865) came closer to the truth when he asserted that in dividing eggs of sea urchins the cell membrane could be compared to the film which forms at the drying surface of a protein solution, or the more complex membranes which appear when this solution makes contact with fats. Thus, although he never wrote the promised full-length paper, in his autobiography von Baer (1865) reported a few details which did not appear in the 1847 preliminary note. For instance, he gave a more precise description of the now-called metaphase-anaphase transition: when the nucleus expands in length, "for a very short time the mid-part is the widest one, but soon it constricts…. in the meanwhile as its substance retreats towards the opposite sides." More importantly, however, he explained why he considered that the discovery of the egg's division had brought him into "the innermost tabernacle of embryology."

For me it is embryology the beloved

As reported by Stieda (1878), in 1845 von Baer wrote an emotional letter to Nicholas Fuss, then secretary of the Imperial Academy of Sciences of St. Petersburg, where he compared his return to embryology during the Italian trip to the reunion of two lovers after a long separation. His enthusiasm stood in marked contrast to the rather cool attitude of the Academy of Sciences towards the above-

summarized results. Stieda himself thought that von Baer had overestimated his own discoveries due to nervous excitability following an excess of work. Later, however, von Baer (1865) made it clear that he considered his embryological studies in Genoa and Trieste to be an integral part of his life's task, which was to describe the main types of development and the main groups of organization of the animal kingdom. In his treatise Entwickelungsgeschichte der Thiere (1828) he had defined development as growth combined with gradual differentiation of tissues and organs from a few fundamental germ layers, and the type of organization as the pattern of relative positions of these organic elements. He also asserted that during development the special forms arise from the more general ones, starting from the simplest fundamental form which is common to every animal, the germ or vesicle, such as that of the chick egg. However, in 1828 nobody knew the process by which the constituents of the germ layers arise from the germ. According to von Baer (1865), this gap was filled by his embryological studies on frog and sea urchins, which demonstrated "the self-building of the substance by continuous division." This stage was common to every animal, in agreement with the above-summarized law of development which, as a corollary, implied a parallelism between lower and higher adult animals and early and late developmental stages of a given embryo. "Not only a similar process of division is observed in eggs of very different animals as a result of fertilization, namely, either division of the whole yolk mass, or of a small layer which I call germ, but this form of division, which appears to be so mechanical, is precisely the same by which the lowest organisms multiply, those in which growth and procreation are one and the same process, or those which cannot grow without multiplying and cannot multiply without having grown to this point by absorbing material from the external environment. There is always an inner part or nucleus, and a peripheral part. The nucleus is the dominant part" (von Baer 1865). He did not speculate about the role of the nucleus, which then was completely unknown, but his emphasis on the truly mechanical nature of division might suggest a parallelism between nuclear divisions and "the histogenetic differentiation (Sonderung = separation) which continued up to the eventual formation of the animal" (von Baer 1847). Conversely, the fact that all new elements arising from the egg division are nothing but "parts and expressions of the dominant unity" (von Baer 1865) might be related to the derivation of all nuclei from the nucleus of the fertilized egg.

These observations show that von Baer combined insightful generalizations with an empirical attitude. In his words, the basic rule was to never subscribe to theoretical views which could not be derived directly from concrete observations (von Baer 1865). When he carried out his studies on sea urchins, the echinoderms were still included in the Radiata together with cnidarians and ctenophorans, and therefore, when he compared the blastula and gastrula stages to a larval medusa and a ctenophoran he could have been influenced by his belief that the basic features of a body plan appear earlier in development. However, he was very cautious about this similarity and in a footnote he added that more advanced larvae (arguably the pluteus stage) have a quite different shape (von Baer 1847). To the contrary, Dufossé (1847) described larval sea urchins as radially symmetrical

throughout development, since he followed blindly the application to systematics of the above-mentioned law of development (Milne-Edwards 1844). Soon later, however, Derbès (1847) demonstrated definitely that these larvae are bilaterally symmetrical and never become attached to the substrate by means of a stalk, as reported by Dufossé. His conclusion was that sea urchins do not retain all along their existence the type of their conventional taxonomic group.

> The study of embryology is like climbing an Alpine peak (von Baer 1865)

In summer 1846 von Baer returned to the Adriatic Sea with the purpose of resuming his studies on marine invertebrate embryology. However, after a period of several months the Academy of Sciences of St. Petersburg did not give him permission for a longer stay abroad and he was obliged to go back to Russia before the sea urchin spawning season. Thereafter, von Baer never returned to embryology, but in the second half of the nineteenth century, Alexander O. Kowalevsky and Il'ya I. Metschnikoff completed and extended to other marine invertebrates his pioneering studies on ascidians and sea urchin embryology. These investigations were carried out under the influence of Darwin's evolutionary theory, but in spite of a few divergences, such as that concerning the interpretation of the ascidian larva (Kowalevsky 1866; von Baer 1873), they remained basically adherent to von Baer's research program (Raineri 2009, 2013). This experimental and conceptual work founded the Russian tradition of evolutionary comparative embryology which left important legacies to current evolutionary developmental biology (Gilbert 2003; Raff and Love 2004; Mikhailov 2012). The relationship between embryology and systematics that was suggested by von Baer did not lose its heuristic significance when the tree of life came to be ascribed to descent with modifications, as maintained by Darwin's theory. Although he was sympathetic to evolution, von Baer was critical of this theory. In his opinion, adaptation did not come about as the contingent result of random mutation and natural selection, but evolution could be understood by analogy with embryogenesis as a goal-directed process driven by a general law of development. The frustrating experiences in Venice and Trieste in 1845–1846 had stressed the need for well-equipped laboratories and international cooperation to establish embryology as a major field of research. Thus, while approaching the end of his life von Baer endorsed the project of Anton Dohrn and gave him full support for the foundation of the Zoological Station of Naples (Groeben 1993).

**Acknowledgments** This research was supported by the Estonian Ministry of Education. Grant No Pl3005SPT and No IUT 02-16.

## References

Anonymous (1880) Ein Brief Karl Ernst von Baers an Dr. Th. L. W. Bischoff (damals) in Giessen. Augsburger allgemeine Zeitung, Beilage Nr. 325, 20. 11

Bagge H (1841) Dissertatio inauguralis de evolutione *Strongyli auricularis* et *Ascaridis acuminatae viviparorum*. Ex officina Barfusiana, Erlangae

Baker JR (1953) The cell-theory: a restatement, history, and critique. Part IV. The multiplication of cells. Q J Microscopical Sci 94:407–440
Baker JR (1955) The cell-theory: a restatement, history, and critique. Part V. The multiplication of nuclei. Q J Microscopical Sci 96:449–481
Briggs E, Wessel GM (2006) In the beginning… animal fertilization and sea urchin development. Dev Biol 300:15–26
Casellato S (2008) Marine biology in the Adriatic Sea: an historical review. Proc California Acad Sci 59(suppl 1):193–205
Darwin CR (1871) The descent of man and selection in relation to sex. Murray, London
Derbès AA (1847) Observations sur le méchanisme et les phenomènes qui accompagnent la formation de l'embryon chez l'oursin comestible. Ann Sci Nat 3rd ser Zool 8:80–98
Dufossé A (1847) Observations sur le developpement des oursins. Ann Sci Nat 3rd ser Zool 7:44–52
Gilbert SF (2003) The morphogenesis of evolutionary developmental biology. Int J Dev Biol 47:467–477
Groeben C (1993) Karl Ernst von Baer (1792–1876) Anton Dohrn (1840–1909) Correspondence. Trans Am Philos Soc 83(3):1–156
Groeben C (2008) Tourists in science: 19th century research trips in the Mediterranean. Proc California Acad Sci 59(suppl 1):139–154
Haeckel E (1868) Natürlische Schöpfungsgeschichte. Reimer, Berlin
Hertwig O (1876) Beiträge zur Kenntniss der Bildung, Befruchtung und Theilung des thierischen Eies. Morphol Jahr 1:347–452 pls X–XIII
Kölliker A (1843) Beiträge zur Entwicklungsgeschichte wirbelloser Thiere. 1. Ueber die ersten Vorgänge im befruchteten Ei. Müller's Arch Anat Physiol wiss Med (no vol. number): 68–141 pls VI–VII
Kowalevsky AO (1866) Entwicklungsgeschichte der einfachen Ascidien. Mém Acad Imp Sci St. Pétersbourg, Ser VII, 10:1–19 pls I–III
Krohn A (1852) Ueber die Entwickelung der Ascidien. Müller's Arch Anat Physiol wiss Med: 312–333 pl VIII
Mikhailov AT (2012) Russian comparative embryology takes form: a conceptual metamorphosis toward "evo-devo". Evol Dev 14:9–19
Milne Edwards H (1844) Considérations sur quelques principes relatifs à la classification naturelle des animaux. Ann Sci Nat 3rd ser Zool 1:65–99
Preller L (1850) Carl Otto v. Madai, zur Erinnerung an ihn für seine Freunde. Leipzig
Raff RA, Love AC (2004) Kowalevsky, comparative evolutionary embryology, and the intellectual lineage of EvoDevo. J Exp Zool (Mol Dev Evol) 302B:19–34
Raikov BE (1968) Karl Ernst von Baer 1792–1876. Sein Leben und sein Werk. Übersetzt und ergänzt von H. v. Knorre, Acta Historica Leopoldina 5. Johann Ambrosius Barth, Leipzig
Raineri M (2009) On some historical and theoretical foundations of the concept of chordates. Theory Biosci 128:53–73
Raineri M (2013) Revisiting the past in view of the future. Evol Dev 15:1–2
Schwann T (1839) Mikroskopische Untersuchungen über die Uebereinstimmung in der Struktur und dem Wachstum der Thiere und Pflanzen. Sander'schen Buchhandlung, Berlin
Stieda L (1878) Karl Ernst von Baer. Eine biographische Skizze. Druck und Verlag von Friedrich Vieweg und Sohn, Braunschweig
Tammiksaar E, Brauckmann S (2004) Karl Ernst von Baer's 'Über Entwickelungsgeschichte der Thiere II' and its Unpublished Drawings. Hist Phil Life Sci 26:291–308
von Baer KE (1828) Ueber die Entwickelungsgeschichte der Thiere. Beobachtung und Reflexion. Erster Theil. Bornträger, Königsberg
von Baer KE (1834) Die Metamorphose des Eies der Batrachier vor der Erscheinung des Embryo und Folgerungen aus ihr für die Theorie der Erzeugung. Müller's Arch Anat Physiol wiss Med (no vol number):481–509 pl XI
von Baer KE (1847) Auszug aus einem Berichte des Akademikers v. Baer, aus Triest vom 1. (13.) November 1845. Bull Classe Phys-Math Acad Imp Sci St. Pétersbourg 5:231–240

von Baer KE (1865) Nachrichten über Leben und Schriften des Herrn Geheimrathes D[r.] Karl Ernst v. Baer, mitgetheilt von ihm selbst. Buckdruckerei der Kaiserlichen Academie der Wissenschaften, St. Petersburg

von Baer KE (1873) Entwickelt sich die Larve der einfachen Ascidien in der ersten Zeit nach dem Typus der Wirbelthiere? Mém Acad Imp Sci St. Pétersbourg, Ser VII, 19:1–35 pl I

von Grünewaldt-Haackhof O (1900) Vier Söhne eines Hauses. Zeit- und Lebensbilder aus Estlands Bd. 2.Vergangenheit. Leipzig: Duncker & Humblot

# Part II
# Evolutionary Biology Concepts

# Chapter 2
# Ontic Openness as Key Factor in the Evolution of Biological Systems

**Søren Nors Nielsen and Claus Emmeche**

**Abstract** The heterogenic character of biological systems has as a consequence that calculations of their possible combinatorial constellations very soon run into numerical explosions. This means, that the resulting numbers—so-called immense numbers—exhibit orders of magnitude beyond any physical meaning. Such a high number of possibilities cause another property—named ontic openness by the physicist W. M. Elsasser—to emerge within such systems. All biological systems possess the feature of being ontic open and this is of great importance to evolution, as ontic openness not only guarantees a development of the system to take place, but also interferes with our chances to fully comprehend this evolutionary processes sensu lato. Thus ontic openness implies an extremely high level of uncertainty and unpredictability. On the one hand, we have a certainty that "something" is bound to happen within the system—on the other hand, we can be totally sure that we will never be able to forecast exactly whatever that "something" will be. At lower levels of biological hierarchy, e.g., the molecular level represented by molecules like DNA, RNA, and proteins, ontic openness seems pretty easy to comprehend. When it comes to more aggregate and even conglomerate systems, i.e., at higher levels of biological hierarchy, the emergence as well as the expression of this property becomes increasingly obscure. Although definitely present, the property at superior levels tends to be overlooked or neglected. Although the calculations may take different forms—and in spite of finding different causes—the property penetrates through all levels of biological hierarchy. To prevent systems from ending up in a situation where the evolutionary state described by calculations that are incomprehensible or even intractable constraints

S. N. Nielsen (✉)
Department of Learning and Philosophy, Aalborg University, Copenhagen,
A.C. Meyers Vænge 15, 2450 Copenhagen SV, Denmark
e-mail: soerennorsnielsen@gmail.com

C. Emmeche
Center for the Philosophy of Nature and Science Studies, University of Copenhagen,
Copenhagen, Denmark
e-mail: cemmeche@ind.ku.dk

P. Pontarotti (ed.), *Evolutionary Biology: Exobiology and Evolutionary Mechanisms*,
DOI: 10.1007/978-3-642-38212-3_2, 

of the systems are needed. From the different levels some systematic patterns seem to be recognizable. Whereas lower levels find causes inside–upwards to be dominating, at upper levels causes become dominated by outside–inwards interactions. Eventually, the ontic openness is likely to be limited not only by physical dimensions but is also constrained by downward acting factors. One reason for this is that space and time scales are well-known to be tightly coupled throughout the biological hierarchy—smaller scales have fast reaction rates as opposed to large scale with slower functions. Thus, space and time scales become important to the realization of ontic openness. At the same time, a shift occurs that stresses information exchange and treatment together with cognitive processes to be increasingly dominant in the biosemiotics of the ongoing processes. The whole leads to a shift from dominance of objective factors to more subjective ones in the process of evolution. Viewing evolutionary systems as ontic openness systems and pursuing the constraints influencing them may turn out to be a fruitful strategy to the investigation of all developmental processes.

## 2.1 Introduction

Most researchers in areas of science, such as physics, chemistry, and biology have most likely during their career given thoughts to the remarkable property we call "life" as opposed to the much more inert system we also deal with within the natural sciences for instance in physics. What exactly makes the two types of systems differ? Many words have through time been used to designate what are the essential properties that define life or life processes, among others, possibility to replicate, storage and treatment of information, self-similarity, self-organization, network properties, etc.

At the same time, probably not many have done any considerations about the real "uniqueness" of living systems (Ulanowicz 2009). The state of this uniqueness is among others a derivative of the works of and considerations done by the physicist Walter M. Elsasser (1904–1991, 1998). Although being trained in the area of nuclear and quantum physics, Elsasser's career forced him into activities of many other only remotely related disciplines in areas such as meteorology, and in particular earth sciences. He is probably best known for his contributions concerning Earth's magnetism, where it was his findings that eventually lead to the discovery of plate tectonics. For further information we refer to a recently published biography (Rubin 2005).

Less known is probably a quite large part of his production inspired among others by Niels Bohr in which he dealt with considerations on the actual difference between living and physical systems. In a series of papers (Elsasser 1963, 1964, 1969a, b, 1972, 1981a, b, 1982, 1983) and four books (Elsasser 1958, 1966, 1975, 1998) he developed a synoptic framework with a more or less specific nomenclature to describe his views and predominantly phenomenological observations

concerning the characterization of physical versus biological systems. For a brief, yet more detailed introduction see later. All in all, his findings have a great potential impact on the way we understand evolutionary biology.

In general, evolutionary (neo-) Darwinian theory includes four aspects that form the core idioms of the theory, (as for instance stated and presented in Christiansen and Fenchel 2009) namely (1) an excessive production of offspring, (2) that the individuals produced are different and that this difference is heritable, (3) that the most fitted species have the highest probability to survive and reproduce, i.e., will be *selected* for. The previous points will (4) lead to a change between generations and that the species will be adapted to meet environmental changes (modified from Christiansen and Fenchel 2009).

Out of all this, the *principle of selection* (points #3 and #4 above) has received major attention and the discussions in this area have led to many treatments abundant in the scientific literature. Beyond no doubts, this is for several reasons, such as the weaknesses in the definition of "who" is carrying out this selection, or exactly "what" causes it to come by? What is it exactly that is selected for? And if fitness, then how is this fitness defined? Debates in this area have not ended and are not likely to be settled for years, although some indications exist (see Jørgensen et al. 2007) that fitness is most likely to be best understood as some sort of optimal energetic relation with the surroundings or even optimum thermodynamic efficiency (Jørgensen 2002).

As opposed to selection the necessary *principle of variation* that according to #2 above lies beneath the whole has received much less attention at least when it comes to the way it is considered here. The principle of variation is included in the "differences" among the excessively produced number of individuals, but so far close to nothing is stated about how these differences emerge nor has an underlying principle such as ontic openness been described.

Obviously, the genetics of mutations and the discovery of the genomic structure were necessary to build a platform for understanding the variation properly. In brief, having gained this insight variation seems to be what comes by when chromosomes undergo point and other mutations, and when genes becomes shuffled around during the mitotic and meiotic processes at lower levels of hierarchy and when individuals exhibit sexual reproduction are involved (see later). All in all, a plausible and sufficient explanation to variation at this level of hierarchy, but it leaves a picture that is far from complete. Although for instance, Mendelian laws allow for prediction to be made about the frequencies of certain characteristics within a given species, the calculations are still deterministic in the sense that for instance frequencies of alleles predicted from these laws are fixed.

Attempts of modeling gene distributions and fitness landscapes using stochastic methods often leads to distributions dominated by "noise", which is disturbing to both conclusions and computing time. The problem about such models is probably a lack of constraints that lead to too many degrees of freedom and an unrealistically high variability. What one would expect to see is rather a decrease in variability and noise as an effect of the system being constrained. A conjecture can be proposed that this would be expressed in a corresponding decrease in an entropy

like expression (or any isomorphic forms of equations) when calculated and observed from the system development over time.

It is possible that the sufficiency of such an explanation have had the consequence to draw the attention of the evolutionary science away from a more obscure and profound "variation principle" that lies inherent in all such systems. Such a principle is nevertheless the direct outcome of the considerations done by Elsasser in his theoretical works. The difference in view is now that the variation we pursue and investigate here is about anything else but deterministic (see earlier comment above).

Another reason seem to be that the philosophical works of Elsasser have only recently begun to get some attention. In spite of the fact that the philosophical issues are not too difficult to grasp, we are left with the fact that the messages brought to us some times may seem overwhelming. In fact, many physicists are reluctant to discuss the issues emerging from Elsasser theories for the reason that his calculations reach numbers that according to them are beyond physically any meaning. Meanwhile, such numbers belong to the everyday life of not only biological researchers, and mathematics allows us to calculate them. Therefore, we should not ignore the message they bring.

We can at this point no longer avoid mentioning that the key property of biological system by Elsasser is termed *ontic openness*. All systems which are sufficiently *heterogenic*,—which may be roughly explained as being composed of a large variety of distinguishable elements with difference properties,—will possess this property. Such systems are said to be *ontic open*. Ontic openness in short expands the system by giving it some extraordinary high degrees of freedom, at a level much higher than we used to think of. The "extra" degrees of freedom provide such a system with an intrinsic property that on the one hand ensures evolutionary powers of the system as evolution in this context has no limits. On the other hand, it can be concluded that evolution therefore must be met with limits of some kind, i.e. needs to be constrained. Variation and ontic openness must be viewed together with selection and constraints.

While an earlier contribution (Nielsen and Ulanowicz 2011) has advocated ontic openness as an essential feature at lower levels of biological hierarchy and of ecosystems, this paper argues that the property penetrates to all intermediate levels and examines the various ways that ontic openness emerges. In this context, the concept of code duality (digital and analog codes) introduced in biosemiotics by Hoffmeyer and Emmeche (1991) may be used in order to understand the importance of processes behind ontic openness and eventually shaping its outcome by constraints. Moving up the hierarchy, the binary code dependency still exists but is of reduced importance, whereas the analog codes become increasingly dominant. Thus, a totally randomized outcome of the ontic openness may be objectively and physically constrained at lower levels. The higher the level in the biological hierarchy the more a subjective layer of ontic openness is added to the semiotic processes. The constraints at the same time become involved and dominate these activities, although the system still has to obey physical constraints (Nielsen 2009).

## 2.2 Elsasser and Ontic Openness

The basic distinction between physical and biological systems introduced by Elsasser was that the first type of ensembles (sets) could be characterized as *homogeneous*, whereas the latter type were dominated by a high degree of *heterogeneity*. It is a well-known fact that many physical systems may be described as assemblies of elements or particles that all share approximately the same properties or to use Elsasser's formulation they are homogeneous. It is this homogeneity which makes it possible for us to describe and predict their collective behavior from statistical (mechanical) considerations.

The possibility to treat higher level, conglomerate systems by such average considerations is impossible, simply because their behavior and properties of the elements come to deviate too much from one another. Adding up to this comes also the fact that being conglomerate indicates that they may appear in quite different constellations leading to variations also in quantitative and qualitative properties. Theory of sets may assist in developing an increased understanding in this area.

Heterogeneity is a prerequisite for the possibility to bring around variation in constellations, just as it is the 4–5 nucleic acids making up strings of RNA and DNA that may vary in time and space,—the simplest spatial view being the two-dimensional one presented here. The variability is not restricted to being one–two-dimensional as it is clearly demonstrated already by the proteins and enzymes.

The central issue here is now that heterogeneity introduces a vast number of possibilities that may never be reached neither in time nor space, i.e., evolutionary possibilities that will never be realized (Nielsen and Ulanowicz 2011). This is for two reasons which are fundamental, though few biologists have giving them any thoughts. First, evolution did not have that much time actually to realize all possibilities. In fact, often the argument has been heard, that if only evolution had time enough, this or that feature would be developed. The second argument is striking in all its simplicity. The universe does not have particles enough to allow evolution to construct all possibilities, not to say allow them to evaluated against each other at the same time. Thus, it is an imperative that a major part of the evolutionary space is void.

Elsasser's argument around his concept of ontic openness is in fact based on some quite simple considerations. Given the time of existence of universe ($14 \times 10^9$ y)—and its number of participating elements ($10^{83}$)—and the fastest rate with which events in the system (our universe) could happen (one event per $10^{-9}$ s)—a total of all possible number of events or constellations of the participating elements (elementary particles) that could have been realized can be calculated. This reaches a number in the order of $4.4 \times 10^{109}$.

Elsasser introduces a level where any number of more than $10^{100}$ (a number referred as Googol) is referred to as *immense*. Systems where combinatorial considerations on possible number of constellations lead to numbers higher than Googol are said to be *ontic open* or possess *ontic openness*. The property is an intrinsic feature existing in all heterogeneous systems and leaves us with an

overwhelming number of possible realizations of such systems. Each constellation realized has no identical twin and is fully unique.

The numbers are conceived of as being beyond any physical meaning, hence the resistance among physicists to discuss them, and yet they can be calculated as objective possibilities. In fact reaching such high numbers, described as inconceivable is quite easy in biology. In the world of biologists in general, e.g., genetics, evolutionary biology, physiology, and ecology, such numbers are often reached when attempting to describe possible constellations. Rather than refusing to discuss such numbers, it seems more obvious to face the problem and try to interpret what they mean to the problem we attempt to analyze.

As said, to reach immense numbers during calculations of possible constellations of a biological system is quite easy only the background on which it is calculated differ depending on the ontological consideration. We therefore need to define the level of ontology at which the property of ontic openness emerges. In fact, referring to everyday life we face many situations where ontic openness manifests itself outside a strictly biological context, in the books we read, in the music we hear, in the environment we perceive, and may conclusively only perceive part of. Meanwhile, perception really belongs to the discipline of studying communication among organisms in its widest sense and may therefore be seen as included in biology domain too (as Bateson, Uexküll, and other theoretical biologists have been aware of).

What really stands out when trying to comprehend this inconceivable fact is that the existing life forms seem to have emerged by choosing among the members and realizing sets from an "urn" covering an unfathomable number of possibilities, numbers that seem to be even beyond the infinite. Ontic openness prevails and really seems to be the rule rather than the exception.

When trying to condense some statements relevant to evolutionary biology out of the above characteristics describing the calculations on ontic open system, the following issues seem to be more important.

In an ontic open system, unexpected things are bound to happen and in fact repetition of events is unlikely, or, to take the full consequence, impossible. What has happened once will never occur again. Something similar maybe,—but never exactly the same. For an illustrative example of the (im-)possibility of catching a given number of persons in exactly the same place after a year please refer to Ulanowicz (2009).

This brings the message to us that not only are all states of ontic open systems unique, they are also totally unpredictable. At the same time as we have a "guarantee" that something will happen in them, we can be sure that we will not be able to predict exactly what will happen, and we can be sure that once something has happened precisely the same thing will **never** happen again.

The question immediately raises how to bring in such considerations into not to say understand an evolutionary biology context. What are the extras needed in order to bring us to understand the evolutionary mechanism better? How do we understand the relatively stable evolutionary picture we observe from a situation which forecast such a high level of indeterminacy and uncertainty?

## 2.3 Ontic Openness and Evolutionary Phase Space(s)

It is implicitly clear that such views carry some unpleasant, but also, important messages not only to evolutionary ecology but to all parts of biology that involves ontic open systems.

In fact, ontic openness has the consequence that not only species are unique but so are also the individuals within a population of each of the many species that make up the diversity of the world. As pointed out by Williams (1998, 1st ed. 1956), investigating the physiology of humans, we share no absolute similarity to any other humans, except from ourselves. Even identical twins are not totally identical in this context.

Yet we consider all 7.5 billions of persons on Earth as belonging to the species of Homo sapiens and we will immediately recognize a member even if we are definitely not identical even if we do share our looks and are at least superficially similar.

In order to perceive this, species are not any longer small twigs or branches of an evolutionary tree as it may appear from a macroscopic point of view. When zooming in, we will identify the evolutionary lines of species as bundles of individuals each with their own unique history.

Ontic open systems are provided with an intrinsic "mechanism" of development or evolution. Every time it is possible to take new avenues they are seen available and will be taken.

Furthermore, this type of development in itself does not take any direction but our interpretation, depending on a chosen frame of reference specifying a specific fitness landscape, will allow us to describe a given step as being in a progressive direction (i.e., leading to more adapted, better fitted organisms), or if we will describe it as regressive (leading to a less adapted, lower fitness organism), or if the inevitable variation will introduce any change at all.

At the same time, it is worth to notice that ontic openness of the genome (1) allows for and explains the existence of punctuated equilibria (sensu Eldredge and Gould 1972; Gould 1989), and (2) also implicitly determines that part of the evolutionary phase space must be void and only parts of the phase space close to the presently realized are valid solutions, that is only the "adjacent" is "possible" (Kauffman 1995, 2000).

## 2.4 Ontic Openness and Levels of Hierarchy

If it is really the distinction between systems being either homogeneous or heterogeneous that makes ontic openness possible and if taking the stance that ontic openness has importance in evolution, it seems fair to ask the question when exactly do the systems and components get complex enough, i.e., get sufficiently heterogeneous as seen through the eyes of evolutionary biologists, to talk about

genuine heterogeneity, and thus biological evolvability. Complexity is here used in its widest sense and may be a result of variety in compositional elements and their interrelations both as physical flows or any type of regulatory interactions.

Unfortunately, the answer is not an easy one although a sketch can be drawn. Clearly, to have a world composed of simple and relatively inert inorganic molecules is not enough. We may today find a number of this type of compounds in the order of magnitude of some hundreds (less than 700) atomic(/molecular?) subcomponents under natural conditions. A typical example of homogeneous systems are crystalline structures. Under pre-biotic conditions, the number must have been considerably lower as oxygen was not available. We know, from the experiments of Miller and Urey (Miller 1953; Miller and Urey 1959) that the presence of simple organic molecules of various kinds is possible when a (closed) system is exposed to inputs of energy. All in all, this would result in systems where building blocks were ready but still without ontic openness.

Several setups may have facilitated in the production of essential molecules, localized structures with simple cycles facilitating transports and thereby chemical conversions, such as Turing structures, Benard—Cells, Beluzov–Zhabotinsky reaction like (oscillating spatial) patterns and hypercycling (Eigen and Schuster 1979). At a later state the systems became thermodynamic open, encapsulated in micelles formed by bipolar compounds spontaneously produced and allowing free movement of energy and matter over the boundaries. In addition, the fact that carbon is the core element of life brings about heterogeneity to the living world. Not-surprisingly, the number of organic molecules exceeds 10 millions.

The giant leap must have occurred at a time when some of the molecules joined each other to form longer and longer chains making variable combinations possible. Ontic openness was possible from the time where proteins, RNA and DNA had been formed. Most of these molecules of these types demonstrate even in the simplest forms we know today the property of ontic openness. This is based on a one-dimensional perception (linear organization form) only and not considering other secondary, tertiary, or quaternary configurations.

### *2.4.1 Macromolecules*

At this level, the property of ontic openness stems from the various possibilities of putting together the hereditary material, its storages, messengers, and products, such as enzymes. We find this property among all the macromolecules inside the systems. This has several consequences for a system to possess an ever qualitatively changing material. First of all, because ontic openness has the consequence that replication can never be carried out precisely, that is, without the introduction of deviations from the original. These errors are the source of all evolution, development, and aging. Second, even if proofreading is carried out even this process will not be without mistakes either. The response of the system will be a result depending on the functional changes introduced to the system. The effect of

the changes may be improvements, degrading, or even potentially lethal (see also the above). At this time, it should be noted that redundancy of the genetic material (in addition to proofreading) could be seen as ways to compensate for the ontic openness of the genome. Also it is known that codons that may be replaced without inducing any qualitative changes at, for instance, enzymatic activity is placed on the outside, whereas essential codons are wrapped up and protected by the tertiary structure.

### *2.4.2 Cells*

At the cellular level, i.e., cells including organelles, all functions come by as a response to several types of molecules and the specific functions now connected to the various organelles, like in the nucleus, endoplasmic reticulum, Golgi-apparatus, mitochondria, etc. The organization retains its ontic openness if for nothing else due to its close connection to the macromolecular level.

Second, the compartmentalization can be seen as making the system even more open since the division of processes allows for even more variations to occur and for instance the possibility of even larger deviation from thermodynamic equilibrium (Nielsen 2000). A sufficiently large collection of (individual) cells (as opposed to the organs in the following) will likewise be ontic open.

### *2.4.3 Organs*

A major consequence of collecting cells with similar function in the larger constructs is apparently a reduction in ontic openness. Ontic openness becomes reduced already during differentiation and determination and the fact that the cells have similar functions points in the direction of homogeneity. Other processes like communication between the organ's cells seem to enhance the uniform action of the cells. Whereas ontic openness if for sure retained among the genetic mechanisms being expressed in the main processes of the organs it seems heavily controlled. The function of organs may well be the first level where ontic openness is really reduced by constraints from higher levels. This type of communication should be included in considerations on the next level.

Meanwhile, dedicating certain cells to specific functions indirectly has the consequence that more types of cells must emerge and not even particular organ types have only one type of cells. The ontic openness is shifted to the combinatorial possibilities that exist between all cell types.

### 2.4.4 Individuals

Individuals are systems composed of cells containing genomes and organelles, put together in organs, endo- or exo-skeletons and an encapsulating membrane, in the case of humans—our skin. Such a basic view holds throughout the plant and animal kingdoms. As the hierarchical levels embedded in the organisms are all ontic open it implies that systems when viewed at the level of individuals are also ontic open. Meanwhile, the findings of Williams (1998) indicating that we are only identical to ourselves may appear as overwhelming. In medicine and pharmaceutical sciences, new strategies have been set up to search for "individual medicine". If this is taken too literally, the fact that the physiology of the human organism is ontic open may well force this paradigm into a crisis and sequentially to be reformulated. The best we can hope for is treatment of sets and maybe this would lead to a better research strategy.

The level of individuals when covering the whole evolutionary tree does span over quite a big range of organizational forms and functionalities, from monocellular to multicellular organisms, from autotrophs to heterotrophs. Ontic openness is inherent in all the organisms if for nothing else due to the complexity of the genome. But the constraints seem to change. The difference seems to be best understood in terms of changes in the way that communication takes place in the system. For this purpose, the code-duality perspective on the semiotic processes in biology as introduced by Hoffmeyer and Emmeche (1991) is convenient.

In brief, all exchange of information and communication within or among biological systems may be understood within the framework of biosemiotics, a recent scientific field based in part of the works of Jakob von Uexküll (1926). According to Hoffmeyer and Emmeche, this informational process can be seen as either digital or analog in character where the digital, binary ways of communication can be related the molecular level (especially DNA as a digital memory), whereas most of the other exchanges of information are taking form of analog processes. In multicellular organisms with perception and action cycles, there is a continuous exchange of information between organism and environment, and this "sign action" helps the organism to adapt to the physical environment by representing and processing significant cues of the environment (the creation of an inner Umwelt) upon which actions are based.

When viewed in this manner a general trend can be observed. Focusing at the lower and inner levels of the biological hierarchy ontic openness is a consequence of fluctuations in the digital coding systems. When moving upwards in the hierarchy these "inner" fluctuations still exist but become constrained and leveled out by other communicative processes that are analog in character.

A separate ontic openness is expressed through the internal relations of the individuals. Any interaction between individuals is considered to belong to the population level.

### 2.4.5 Population

Individuals belonging to the same species that are in themselves ontic open come together in larger collections usually referred to as a population. The exchange between the genomes of individuals as it happens for instance in connection to sexual reproduction only serves to an increase in ontic openness. This is probably not the right way of formulating it since either a system is ontic open or it is not. Rather, we should in this situation describe the process of sexual reproduction as adding an extra layer that is also ontic open. This is a detail that only affects the way we would calculate the number of combinatorial possibilities and has no effect to the conclusions made for the systems.

The shift in code duality now becomes more apparent when we talk about populations as societies where members of a species exhibits some sort of cooperation such as it can be observed in bees and ants. Communication certainly becomes an important factor that serves to shape the outcome of the events going on between members within such groups of species.

It is also clear from population ecology that it takes populations to be over a certain "viable" size to persist. In other words, they are to be kept ontic open. What we have done when we as humans interfere with the distribution of various other species in our environment is in fact that we most often decrease (bio-) diversity, that is, reducing the ontic openness of the systems. When talking about endangered or extinct species, we are referring to a reduction in ontic openness to a level where the populations' size is not viable any longer.

### 2.4.6 Communities

The introduction of this concept introduces an ontic openness that is strongly correlated with landscapes and the spatial variational component in the distribution of species. As remarked by Allen and Hoekstra communities are composed of individuals belonging to a number of species are therefore to be considered complex (Allen and Hoekstra 1992). The composition alone makes the communities ontic open.

Communities of organisms are mostly seen as belonging to the same (trophic) group of individuals and that communities have both a vegetational and a zoological side to it, which is often neglected. On the one hand, adding such a perspective seems to add up to the ontic openness. Meanwhile, the increase might not be correlated to a simple combinatorial calculation but it is more likely that many parallel communities exist, that may look similar but where certain species of plants rely on particular animal and vice versa for their existence.

### *2.4.7 Ecosystems*

Ecosystems are of course ontic open too (Jørgensen et al. 2007). In general, as pointed out by Ulanowicz, it takes no more for a system than to be composed of around 75 distinguishable components (or tokens) for it to be ontic open (Ulanowicz 2006). It is quite a task to analyze and fully understand an ecosystem. In fact, it can be shown to be physically impossible (Jørgensen et al. 2007). Not surprisingly, this requirement of the necessary complexity is easily fulfilled by ecosystems. Considering that most ecosystems easily demonstrate a composition with a number of members in this order of magnitude, and to this we should consider to add also variety in flows and in chemical species, as well as include all regulatory (cybernetic) mechanisms, then the need for complexity in terms of state variables is probably even lower.

It seems not to be necessary for this purpose to argue that ontic openness also penetrates to levels of hierarchy like landscapes, regions, and even the biosphere. The greatest difficulty when looking upon this in a higher level context is the fact that the ontic openness of other processes, for instance, influences from the human society comes to act as noise to the system (Nielsen and Müller 2009). Issues concerning these levels are therefore omitted from this treatment.

## 2.5 Subconclusion

Ontic openness is a reality at all levels of biological hierarchy. Even if the way it can be calculated depends on the ontology imposed on the system this is not considered to have any effect on the conclusions about the evolutionary potentials of such a system. The types of hierarchy varies from being embedded (2.4.1–2.4.4 above) to be of a scalar type of hierarchy (2.4.4–2.4.7 above). By "embeddedness" here it is meant that the sub-systems are successively included within the series of physical boundaries. The ontic openness of the upper levels includes the ontic openness of lower layers but adds extra dimensions to it. Each organizational layer seems to possess an additional layer type of ontic openness of its own based on an analog type of semiotic information exchange. The upper level ontic openness acts to constrain the lower levels.

## 2.6 Discussion

As it is clear from the introduction, all possible constellations of an ontic open system cannot be realized as the universe (1) neither have components enough, (2) nor does it or will exist long enough. Two severe limitations imposed on biological systems of which the prior seems quite the strongest. In other words, parts of the evolutionary phase space has to be left void.

As we have seen ontic openness in itself has no values to it. What exactly comes of possessing this feature may be difficult to say. It is simply a basic feature of our universe, and although it highlights some important aspects of its evolution, we can only imagine very artificial and highly idealized universes lacking this feature (e.g., simple mathematical models of evolution). What we can say is that it seems valid for an evolutionary line to exist at some new points that may possibly be taken up by a further realization of ontic openness that, although they are totally random, still fulfill the rules of the game. "Totally random" here does not mean instantly created from scratch or by design, but based upon a long process of already evolved structures, and thus being contingent relative to the pre-history of the system in question, but not pre-determined regarding its further evolution (cf. Gould 1989).

The fact that the ontic openness of the upper levels includes the ontic openness of lower layers points to the fact that all upper layers have an inside–outside, inward–upward component to their ontic openness (a parallel view to the findings in Nielsen 2009).

Considering evolutionary time and space scales at large it seems that the random component of ontic openness must have been most active and important during the early evolutionary states. Very early in the emergence and invention of new functional principles, cf. photosynthesis the evolutionary phase space is struck by a lock-in situation, where certain demands are to be fulfilled and other solutions are not valid any longer. The importance to later states is probably reduced to ensure the variational principle but the degrees of freedom with which ontic openness can be realized is heavily constrained.

Another point is that time and space scales are correlated. Normally, systems with small space scales of the live and react on short time scales, whereas larger systems possess longer time scales.

It would be tempting to relate this observation with the changes in the dominant part of the code duality. The digital code forms seem to have the necessary time of reacting, whereas the analog coding seems more suitable at higher levels. Meanwhile, to reach such a conclusion seems not to be valid as many digital and analog coded processes react on approximately the same time scale. The statements around this are illustrated in Table 2.1.

Rather it seems that different types of semiosis could be involved and that in particular at higher levels, it is beneficial that other cognitive processes are integrated within the system. Organisms with longer time scales and mammals in particular have an advantage in memory and use of experience, whereas this is not an option to a May fly. Again, cognition may be seen as the ultimate constraint.

Ultimately, a rough deduction can be made. Ontic openness may be argued to exist at all levels of biological hierarchy, but can only be calculated. The exact value calculated is dependent on ontological views, but once a system is ontic open in one context or another, it is demonstrated that the phenomenological properties of such systems applies. These properties involve an intrinsic nondirectional evolution of the system. Meanwhile, direction and valuation comes into play when higher levels come into play.

**Table 2.1** A summary of ontic openness at various levels, how the property emerges, what is ultimately constraining it, together with some semiotic perspectives

| Level | Ontic openness reached by… | Constrains | Semiotic perspectives |
|---|---|---|---|
| Molecular | Hardly | Mostly simple molecules | – |
| Macromolecular | Combinatorials calculations of possible constellations of strings of RNA, DNA and proteins | Short time adaptational problem solving at long time scale viable solutions to be produced | Digital code system |
| Cells and organelles | The combined effect of compartmentalization and specialization of various activities at previous level | Performance regulated from inside while request appears from upper levels | Digital and minor part analog |
| Organs | Combined effect from lower level expressions and the different type of cells in the organs | Should act as a whole fulfilling demands of the organism | Digital and increasing part analog |
| Individuals | Organism includes all the previous levels | Fulfilling phenotype demands and adequate semiotic responses | Dominance of analog part for all senses |
| Populations | Interaction among individuals in (meta-) populations | Space and resources,—mostly intraspecific competition | Increased importance of analog signaling among members, e.g. pheromones |
| Communities | Intracommunity interactions | Available “niches” and interspecific competitions | Interaction, communication, and cybernetics |
| Ecosystems | Intra- plus inter population/community interactions | Availability of matter and energy as well as recycling | Higher levels of network interaction, communication, and cybernetics |

In order to fully comprehend evolution and development at all scales of biological hierarchy, it seems fruitful and therefore important to add the perspective of ontic openness together with studies of its constrains in future research in the area.

## References

Allen TFH, Hoekstra TW (1992) Toward a unified ecology. Columbia University Press, New York, p 384

Christiansen FB, Fenchel T (2009) Evolution – den forudsigelige vilkårlighed. Aarhus Universitetsforlag, p 226

Eigen M, Schuster P (1979) The hypercycle—a principle of natural self-organization. Springer, Heidelberg

Eldredge N, Gould SJ (1972) Punctuated equilibria: an alternative to phyletic gradualism. In: Schopf TJM (ed) Models in paleobiology. Freeman Cooper, San Francisco, pp 82–115

Elsasser WM (1958) The physical foundation of biology. An analytical study. Pergamon Press, London

Elsasser WM (1963) Note on evolution in organismic theory. J Theor Biol 4:166–174

Elsasser WM (1964) Synopsis of organismic theory. J Theor Biol 7:53–67

Elsasser WM (1966) Atom and organism. A new approach to theoretical biology. Princeton University Press, Princeton

Elsasser WM (1969a) A causal phenomena in physics and biology: a case for reconstruction synopsis of organismic theory. Am Sci 57(4):502–516

Elsasser WM (1969b) The mathematical expression of generalized complementarity. J Theor Biol 25(2):276–296

Elsasser WM (1972) A model of biological indeterminacy. J Theor Biol 7:627–633

Elsasser WM (1975) The chief abstractions of biology. North Holland, Amsterdam

Elsasser WM (1981a) A form of logic suited for biology. In: Rosen R (ed) Progress in theoretical biology, vol 6. Academic Press, New York, pp 23–63

Elsasser WM (1981b) Principles of a new biological theory: a summary. J Theor Biol 89(1):131–150

Elsasser WM (1982) The other side of molecular biology. J Theor Biol 96(1):67–76

Elsasser WM (1983) Biological application of the statistical concepts used in the second law. J Theor Biol 105(1):103–116

Elsasser WM (1998) Reflections on a theory of organisms. Holism in biology. John Hopkins University Press, Baltimore

Gould SJ (1989) Wonderful life: the burgess shale and the nature of history. Hutchinson, London

Hoffmeyer J, Emmeche C (1991) Code-duality and the semiotics of nature. In: Anderson M, Merrell F (eds) On semiotic modeling. Mouton de Gruyter, Berlin and New York, pp 117–166

Jørgensen SE (2002) Integration of ecosystem theories: a pattern. Springer, Berlin, p 420

Jørgensen SE, Fath B, Bastianoni S, Marques JC, Müller F, Nielsen SN, Patten BC, Tiezzi E, Ulanowicz RE (2007) Chapter 3, ecosystems have ontic openness, p 35–57. In: Jørgensen SE, Fath B, Bastianoni S, Marques JC, Müller F, Nielsen SN, Patten BC, Tiezzi E, Ulanowicz RE (2007) A new ecology: systems perspective. Elsevier, Amsterdam, p 275

Kauffman S (1995) At home in the universe: the search for the laws of self- organization and complexity. Oxford University Press, New York, p 325

Kauffman SA (2000) Investigations. Oxford University Press, Oxford, p 302

Miller SL (1953) Production of amino acids under possible primitive earth conditions. Science 117(3046):528–529

Miller SL, Urey HC (1959) Organic compound synthesis on the primitive Earth. Science 130(3370):245–251

Nielsen SN (2000) Information in ecosystems. Chapter II.2.2, pp 195–216. In: Jørgensen SE, Müller F (eds) Handbook of ecosystem theories and management. Lewis Publishers, Boca Raton, p 584

Nielsen SN (2009) Thermodynamic constraints of life as downward causation in ecosystems. Cybern Hum Knowing 16(1–2):27–49

Nielsen SN, Müller F (2009) Understanding the functional principles of nature—proposing another type of ecosystem services. Ecol Model 220(16):1913–1925

Nielsen SN, Ulanowicz RE (2011) Ontic openness: an absolute necessity for all developmental processes. Ecol Model 222(16):2908–2912

Rubin H (2005) Walter Elsasser, prophet of biological complexity, seeker of simplifying rules. Cell Mol Biol 51:599–606

Ulanowicz RE (2006) Process ecology: a transactional worldview. J Ecodyn 1:103–114

Ulanowicz RE (2009) A third window. Natural life beyond Newton and Darwin. Templeton Foundation Press, West Conshohocken, p 196

von Uexküll J (1926) Theoretical biology. Harcourt, Brace & Co., New York

Williams RJ (1998) Biochemical individuality. The basis for the genetotrophic concept. Keats Publishing, New Canaan, p 267

# Chapter 3
# Effects of Epistasis and Pleiotropy on Fitness Landscapes

**Bjørn Østman**

**Abstract** The factors that influence genetic architecture shape the structure of the fitness landscape, and therefore play a large role in the evolutionary dynamics. Here the NK model is used to investigate how epistasis and pleiotropy—key components of genetic architecture—affect the structure of the fitness landscape, and how they affect the ability of evolving populations to adapt despite the difficulty of crossing valleys present in rugged landscapes. Populations are seen to make use of epistatic interactions and pleiotropy to attain higher fitness, and are not inhibited by the fact that valleys have to be crossed to reach peaks of higher fitness.

## 3.1 Introduction

Rugged fitness landscapes have been suggested to put a halt to adaptation (Gavrilets 2004; Whitlock et al. 1995). When there are multiple peaks in the fitness landscape, populations must cross valleys to achieve higher fitness, but because crossing a valley implies that organisms will have lower fitness, they are likely to get stuck on local fitness peaks. Several solutions have been proposed to the problem of valley-crossing, including non-static landscapes (Mustonen and Lässig 2009; Whitlock 1997; Whitlock et al. 1995), subpopulations crossing by drift (Wright 1932), and circumventing valleys by using neutral ridges (Gavrilets 1997). Here we will see that populations can indeed cross fitness valleys as long as the mutation-supply rate (product of population size and mutation rate) is not unrealistically low. When the supply of mutations is large enough, some organisms can endure lower fitness and still manage to reproduce, thereby giving them a chance to ascend adjacent fitness peaks. An initially maladapted population will most often climb the nearest peak,

B. Østman (✉)
Microbiology and Molecular Genetics, Michigan State University, 567 Wilson Rd, East Lansing, MI 48824-4320, USA
e-mail: ostman@msu.edu

P. Pontarotti (ed.), *Evolutionary Biology: Exobiology and Evolutionary Mechanisms*,
DOI: 10.1007/978-3-642-38212-3_3, 

and if this peak happens not to be the global peak, it can then achieve higher fitness by relying on the stochastic nature of evolution. The more rugged a landscape is, the harder it becomes to cross valleys, but it turns out that not only can populations overcome relatively high levels of ruggedness, but high ruggedness also implies that the global peak is higher, leading to more efficient adaptation. Epistatic interactions between genes therefore not only constrain adaptation when the population gets stuck on a local peak, it actually boosts it. Consequently, deleterious mutations are seen not as a hindrance to adaptation, but as a necessary component without which adaptation would grind to a halt.

The structure of the fitness landscape and the ruggedness that it exhibits are shaped by the interactions between the genotypes and the mutations that take place when a descending organism moves between neighboring genotypes. Epistasis and pleiotropy have distinct but related effects whereby the fitness landscape acquire a structure that either inhibits or enables an evolving population to attain higher fitness. Evolutionary dynamics is thus largely determined by three key parameters: population size, mutation rate, and the fitness landscape. With adequate information about the fitness landscape, and the factors that underlie genetic architecture, the extent to which populations can successfully utilize deleterious mutations and locate the fittest genotype can be assessed. Results from simulation evolving populations in rugged fitness landscapes are here presented showing that adaptation is not slowed down in moderately rugged landscapes, but rather allows populations to attain higher fitness than in landscapes with no epistasis. Three hypotheses concerning epistasis and pleiotropy springing from the model employed in this work are presented for future investigation.

## 3.2 NK Model

To investigate the effect of epistatic interactions on the adaptive process, we employ the NK model, which is a simple system previously used to study interactions between loci with different alleles. The NK model consists of $N$ loci in circular, binary sequences (Kauffman 1993; Kauffman and Levin 1987). Each of the $N$ loci contributes to the fitness of the organism via an interaction with $K$ adjacent loci. For each locus $i$ a lookup-table consisting of uniform random numbers represents the fitness component $\omega_i$ of a binary sequence of length $K + 1$. For example, $K = 1$ (interaction with one other locus) is modeled by creating random numbers for the four possible binary pairs 00, 01, 10, 11 for each of the $N$ loci, that is, the fitness component at one locus is conditional on the allele at one other locus. The overall fitness of an organism is usually given by the average of the N fitness components, but here we use the geometric mean (motivated by the fact that one could then introduce lethal mutations by setting one or more elements in the lookup-tables to zero):

$$W = \left( \prod_{i}^{N} \omega_i \right)^{1/N}.$$

Because the objective is to study the adaptive phase of evolutionary dynamics [as opposed to mutation-selection balance (Desai and Fisher 2007)], the simulations are started with a population of low fitness, allowing the population to increase in fitness. Every computational update 10 % of 5,000 asexual organisms are removed at random, and the remaining organisms replace them by reproduction. The organisms that get to reproduce are chosen with a probability proportional to fitness. That is, if the fitness of an organism is twice that of another, it has twice the chance to reproduce. Since this process is stochastic, organisms of lower fitness are not doomed to extinction, but with luck can become the ancestors of later generations. Reproduction is simulated by making a copy of the chosen organism and allowing each of the $N$ loci to mutate from 0 to 1 or from 1 to 0 at a rate $\mu$. Three different mutation rates are used in order to study the effect of the mutation-supply rate on the ability of the populations to cross valleys in the fitness landscape. Every organism has a genotype that consists of a binary string of length $N = 20$ loci, and the effect of varying degrees of ruggedness on adaptation is investigated by running simulations with different values of $K$, which modulates the amount of epistatic interactions. Each simulation is run for 2,000 updates, which is enough in most instances to attain the highest fitness possible given the fitness landscape, population size, and mutation rate. Because there are no features of the dynamics that allow more than transient coexistence of different genotypes, the most recent common ancestor is never far in the past. Consequently, all organisms surviving at the end of a simulation run shares most of their history, and we can therefore reconstruct the line of descent (LOD) that is common to all surviving organisms and still cover most of the 2,000 computational updates. Epistasis is measured between pairs of consecutive mutations, $i$ and $j$, on this shared LOD in the following way:

$$\varepsilon_{\mathrm{ij}} = \log\left( \frac{W_0 W_{\mathrm{AB}}}{W_A W_B} \right),$$

where $W_0$ denotes fitness before either mutation, $W_{\mathrm{A}}$ is fitness with the first mutation, $W_B$ is fitness with the second mutation, and $W_{\mathrm{AB}}$ is fitness with both mutations. The genotype with only mutation $B$ does not occur on the LOD, so $W_B$ has to be reconstructed afterwards for epistasis to be calculated. Epistasis on the LOD is calculated as the average of each pair of mutations. Epistasis can readily be calculated both between non-consecutive mutations as well as between more than two mutations, but we are here interested in how epistatic interactions enable populations to cross fitness valleys, and the most recent mutations have the largest effect on the ability to do this.

## 3.3 Results

The NK model was used to investigate the extent to which adapting populations make use of epistatic interactions. Simulations were carried out for three different mutations rates of $\mu = 10^{-4}$, $10^{-3}$, and $10^{-2}$. With a constant population size of 5,000, this gives mutation-supply rates of 0.5, 5, and 50, which spans values on either side of the limit demarcating the strong-selection weak-mutation regime (SSWM) from the regime where multiple mutations go to fixation at the same time. This regime has been investigated analytically under the assumption that the mutations do not interact (Unckless and Orr 2009), but here this assumption is removed allowing for epistasis to affect the adaptive process. Adaptation is observed and the fitness attained at the end of the 2,000 updates is recorded as $\Omega$, and used to compare the efficiency with which different populations adapt (Fig. 3.1).

As $K$ is increased from 0 (smooth landscape with no epistasis) to 10 (highly rugged landscape), the number of substitutions, i.e., mutations on the LOD, drops significantly (Fig. 3.2). The fraction of substitutions that are beneficial stays approximately constant across this range of $K$, while the amount of epistasis on the LOD increases (Fig. 3.3). The attained fitness increases up to $K \approx 5$ (depending on the mutation rate), after which a decline is observed (Fig. 3.4). So despite there are fewer beneficial substitutions for higher $K$, the population is still

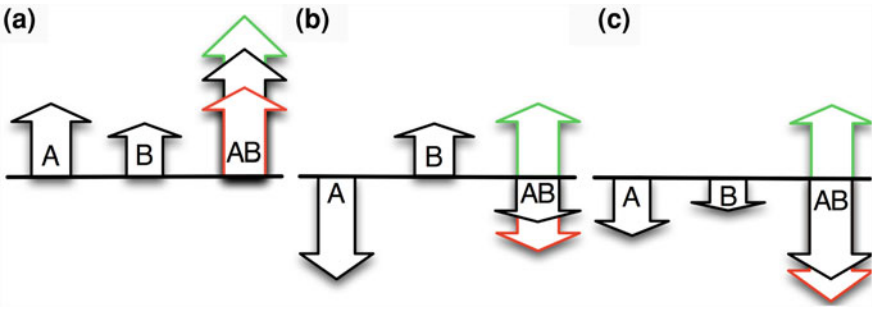



**Fig. 3.1** Illustration of epistasis. **a** Two beneficial mutations occurring together in one organism have an expected non-epistatic value equal to the product of individual fitness, $W_A \times W_B = W_{AB}$ (*black arrows*). Fitness values higher than this correspond to positive epistasis (*green arrow*), and fitness values lower correspond to negative epistasis (*red arrow*). **b** Sign epistasis denotes situations where the effect of a mutation chances depending on the genetic background on which it occurs. Here mutation $B$ is beneficial when occurring alone, and the expectation is that the effect of $A$ and $B$ together is dominated by the larger effect size of the deleterious mutation $A$. When the actual combined effect changes from the expected deleterious to beneficial (*green arrow*), we observe sign epistasis. **c** Two deleterious mutations have the non-epistatic expectation of being deleterious. Reciprocal sign epistasis occurs when the combined action of the two mutations reverses the effect on fitness. Reciprocal sign epistasis occurs when a valley in the fitness landscape is crossed, and it is a necessary condition for multiple peaks to exist. From Østman et al. (2012)

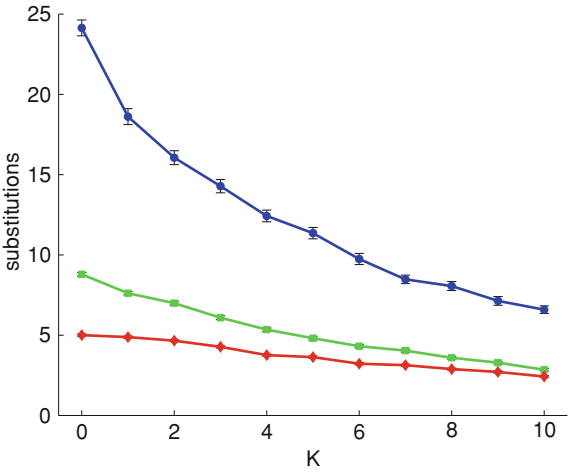



**Fig. 3.2** The number of substitutions declines as the ruggedness of the landscapes increases. *Red diamonds* $\mu = 10^{-4}$, *green squares* $\mu = 10^{-3}$, *blue circles* $\mu = 10^{-2}$. Data are averages over 200 simulation runs. Error bars are s.e.m. From Østman et al. (2012)

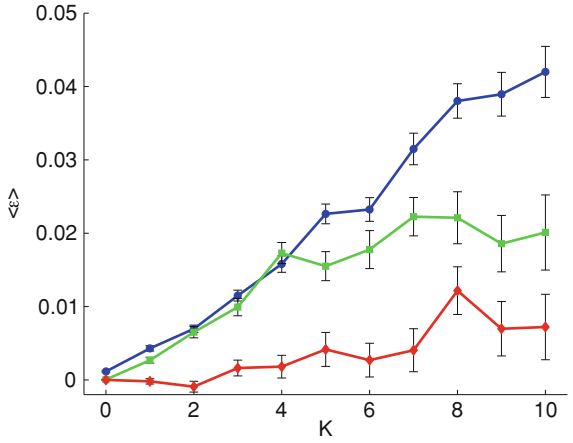



**Fig. 3.3** Mean epistasis of all pairs of substitutions increases with landscape ruggedness. Colors and data as in Fig. 3.2. From Østman et al. (2012)

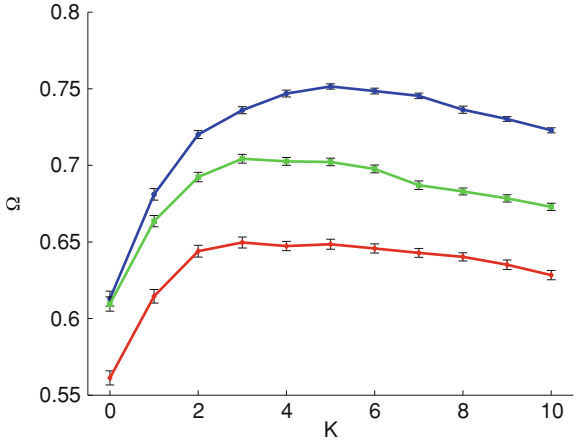



**Fig. 3.4** Attained fitness increases with landscape ruggedness to a point after which a decline is observed. Colors as in Fig. 3.2. From Østman et al. (2012)

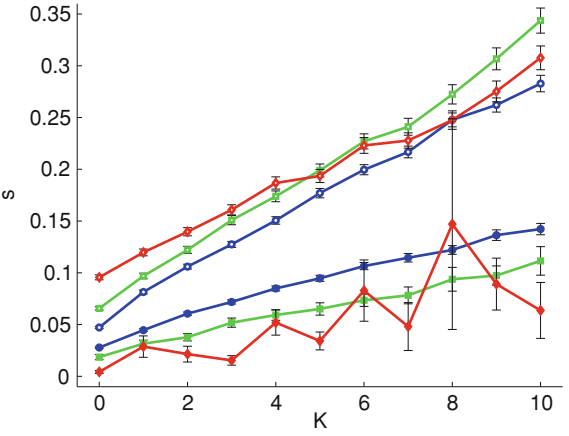



**Fig. 3.5** Selection coefficients increase with landscape ruggedness. Colors as in Fig. 3.2. *Open symbols* are beneficial substitutions, and *solid symbols* are deleterious substitutions. From Østman et al. (2012)

able to attain a higher fitness. This seems contradictory, for surely more beneficial substitutions would result in higher fitness. The answer is that the structure of the fitness landscape is the main determinant of the extent to which an evolving population can adapt. In the NK model epistasis causes changes in fitness that deviate from the non-epistatic expectation. The effect of each mutation is modulated, and the selection coefficients increase with $K$ (Fig. 3.5). This increase in selection coefficients is an effect of both epistasis and pleiotropy. Epistasis causes the fitness landscape to be more rugged and therefore to contain more peaks, which will then have steeper slopes. Pleiotropy increases the effect a mutation can have on fitness by affecting more than one trait at a time. Both of these cause the distribution of selection coefficients to be broader, opening up more opportunities for mutations of large effect.

## 3.4 Discussion

### *3.4.1 Epistasis Causes Fitness Landscape Ruggedness*

Fitness landscapes that contain only a single peak are often called smooth, whereas landscapes with multiple peaks are denoted as rugged (Fig. 3.6). Rugged landscapes can vary both in the number of peaks and in the range of fitness values between the least fit genotypes and the fitness of the global peak. This ruggedness of the fitness landscape is caused entirely by epistasis, with selection determining the fitness of individual genotypes. When there is no epistasis, the landscape is smooth. In this case, adaptation is straightforward, as an evolving population will eventually reach the top of the peak. When more than one peak is present, there is always epistasis present as well (Poelwijk et al. 2011). Take for example a fitness landscape in one dimension where fitness is a function of body-size. If small and

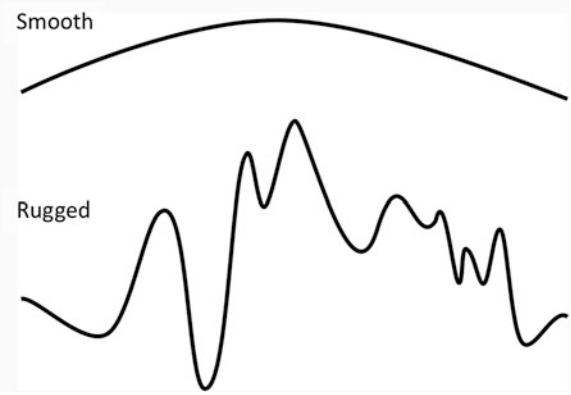



**Fig. 3.6** Examples of smooth and rugged fitness landscapes in one dimension. The smooth landscape has a single peak whereas the rugged landscape has multiple local peaks. The ruggedness is caused by epistatic interactions between genes or mutations, and because epistasis and pleiotropy are coupled in the NK model, increased epistasis also leads to higher levels of pleiotropy. Ruggedness therefore implies not only more peaks but also a greater range in fitness, here represented by the scale of the *vertical axes*

large bodies are favored by selection but intermediate sizes have lower fitness, then there are two peaks. Moving from one peak to the other necessarily results in epistasis between at least one pair of mutations: at the bottom of the valley, one mutation causes an increase in fitness only on the background of another, without which there would be a decrease in fitness.

While fitness landscape ruggedness is caused by epistasis by increasing the number of peaks, it is pleiotropy that is responsible for the increased range in fitness values. If we imagine the fitness landscape as a one-dimensional string, then the smooth landscape is simply a string with a single peak (Fig. 3.6). It has a range from one of the ends of the string to the height of the global peak. If we increase the number of epistatic interactions between genes and mutations, then the string becomes wrinkled, resulting in multiple local peaks. However, if we were to do this without affecting pleiotropy, then the fitness range would be unaffected. On the other hand, if we only increase the average level of pleiotropy while keeping the number of epistatic interactions constant, then the fitness range would expand without affecting ruggedness. This happens because pleiotropy results in mutations affecting more than one trait at a time, thereby broadening the distribution of the fitness effect of mutations, whether deleterious or beneficial. We can express this as epistasis modulates the frequency of the landscape, while pleiotropy modulates the amplitude. In the NK model epistasis and pleiotropy are coupled, and $K$ directly affects both at the same time. In natural systems this is generally not the case. Rather, genes group into modules that affect individual phenotypic traits, and through their joint action on this trait the genes (an mutations affecting them) interact epistatically. In contrast, while genes generally interact with many other genes, the level is pleiotropy is neither tied to epistasis nor is it as prevalent. The emerging picture from

natural systems (e.g. yeast) is that there are relatively few pleiotropic links between modules, resulting in well-defined modules of genes affecting single traits (Costanzo et al. 2010). Consequently, the effects of epistasis and pleiotropy are not coupled in natural systems, and decoupling them in the NK model could therefore increase the realism of the model. If the level of pleiotropy is lower in natural systems, then the range in fitness will also be affected, and the linear relationship found between selection coefficients and K will probably be less steep.

### *3.4.2 Future Directions*

It is generally assumed that the fitness effects of mutations in pleiotropic genes are uncorrelated (e.g., Orr 1998). Because most mutations are either neutral or deleterious, a mutation that is beneficial in one trait is considered most likely to be neutral or deleterious in the other traits that the pleiotropic gene affects. However, considering that the gene encodes a protein that is likely to have the same function in several or all the traits, it seems more likely that if the mutation is beneficial for one trait, then it would also likely be beneficial in the linked traits. For example, if the protein has the same function in the linked traits, then mutations that improve thermal stability or affinity of the active site are likely to be beneficial in all the traits. If the function of the protein is different in the linked traits, then there may be no correlation between the fitness effects, but for proteins whose biochemical function is similar in the linked traits, a correlation in fitness is likely. Such pleiotropically correlated fitness effects could have consequences for adaptation, causing beneficial mutations in pleiotropic genes to have larger effect, thereby increasing both the speed and probability of fixation.

In the weak mutation regime, early adaptation in NK is dominated by non-interacting beneficial mutations with some negative epistasis. Later in the adaptive process and nearer the peak epistasis shifts to become predominantly positive and to include some sign epistasis (Draghi and Plotkin 2012). The diminishing returns observed among beneficial mutations in adapting microbial populations (Chou et al. 2011; Khan et al. 2011) thus appears to be due to regression to the mean. Experimental populations are generally not seen crossing valleys in the fitness landscape, so the positive and reciprocal sign epistasis that would then be observed have rarely been reported in the literature [but see Dawid et al. (2010); Kvitek and Sherlock (2011)]. Valley can be crossed when the mutation-supply rate is large enough, and the waiting time to new mutations is short and deleterious mutations can be tolerated. Reciprocal sign epistasis is a necessary condition for multi-peaked fitness landscapes (Poelwijk et al. 2011), so empirical observations would therefore shed light on this important aspect of fitness landscape structure.

In the NK model, global peak height increases with landscape ruggedness (Østman et al. 2012). However, this may not be an artifact of the NK model alone. Rather, because the synergy between genes in modules is dependent on the number of genes, modules of genes encoding traits will have a larger effect on fitness the

more genes are available. In other words, larger modules confer higher fitness. A comparison between different species that share traits (e.g., vision) should reveal a correlation between the fitness conferred by the module and the number of genes in the module. This could be measured by counting the number of genes that contribute to, say, vision among similar organisms and scoring the trait on a scale of how well it functions. Most likely the increase in fitness as a function of the number of genes is less than linear, because the function cannot be improved indefinitely.

## References

Chou H, Chiu H, Delaney N, Segrè D, Marx CJ (2011) Diminishing returns epistasis among beneficial mutations decelerates adaptation. Science 332:1190–1192

Costanzo M, Baryshnikova A, Bellay J, Kim Y, Spear ED, Sevier CS, Ding H, Koh JLY, Toufighi K, Mostafavi S, Prinz J, St Onge RP, Vandersluis B, Makhnevych T, Vizeacoumar FJ, Alizadeh S, Bahr S, Brost RL, Chen Y, Cokol M, Deshpande R, Li Z, Lin ZY, Liang W, Marback M, Paw J, San Luis BJ, Shuteriqi E, Tong AHY, van Dyk N, Wallace IM, Whitney JA, WEirauch MT, Zhong G, Zhu H, Houry WA, Brudno M, Ragibizadeh S, Papp B, Pal C, Roth FP, Giaever G, Nislow C, Troyanskaya OG, Bussey H, Bader GD, Gingras AC, Morris QD, Kim PM, Kaiser CA, Myers CL, Andrews BJ, Boone C (2010) The genetic landscape of a cell. Science 327:425–431

Dawid A, Kiviet DJ, Kogenaru M, de Vos M, Tans SJ (2010) Multiple peaks and reciprocal sign epistasis in an empirically determined genotype-phenotype landscape. CHAOS 20(2):026105

Desai MM, Fisher DS (2007) Beneficial mutation selection balance and the effect of linkage on positive selection. Genetics 176(3):1759–1798

Draghi JA, Plotkin JB (2012) arXiv:1212.4114v1 [q-bio.PE]

Gavrilets S (1997) Evolution and speciation on holey adaptive landscapes. Trends Ecol Evol 12(8):307–312

Gavrilets S (2004) Fitness landscapes and the origin of species. Princeton University Press, Princeton and Oxford

Kauffman S (1993) The origins of order. Oxford University Press, New York

Kauffman S, Levin S (1987) Towards a general theory of adaptive walks on rugged landscapes. J Theor Biol 128(1):11–45

Khan AI, Dinh DM, Schneider D, Lenski RE, Cooper TF (2011) Negative epistasis between beneficial mutations in an evolving bacterial population. Science 332:1193–1196

Kvitek DJ, Sherlock G (2011) Reciprocal sign epistasis between frequently experimentally evolved adaptive mutations causes a rugged fitness landscape. PLoS Genet 7(4):e1002056

Mustonen V, Lässig M (2009) From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation. Trends Genet 25(3):111–119

Orr HA (1998) The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution 52(4):935–949

Østman B, Hintze A, Adami C (2012) Impact of epistasis and pleiotropy on evolutionary adaptation. Proc Roy Soc B: Biol Sci 279:247–256

Poelwijk FJ, Tănase-Nicola S, Kiviet DJ, Tans SJ (2011) Reciprocal sign epistasis is a necessary condition for multi-peaked fitness landscapes. J Theor Biol 272(1):141–144

Unckless RL, Orr HA (2009) The population genetics of adaptation: multiple substitutions on a smooth fitness landscape. Genetics 183:1079–1086

Whitlock MC (1997) Founder effects and peak shifts without genetic drift: adaptive peak shifts occur easily when environments fluctuate slightly. Evolution 51(4):1044–1048

Whitlock MC, Phillips PC, Moore FBG, Tonsor SJ (1995) Multiple fitness peaks and epistasis. Ann Rev Ecol Syst 26:601–629

Wright S (1932) The roles of mutation, inbreeding, crossbreeding, and selection in evolution. In: Proceedings of the sixth international congress of genetics, vol 1. pp 356–366

# Part III
# Exobiology and the Origin of Life

# Chapter 4
# The Origin of Life, Evolution, and Functional Organization

**Wim Hordijk, Mike Steel and Stuart Kauffman**

**Abstract** The idea that autocatalytic sets played an important role in the origin of life is not new. Neither is the idea that autocatalytic sets can tell us something about the evolution and functional organization of living systems. However, most of these ideas have, until recently, remained at a conceptual level, and very few concrete, mathematically sound, and practically applicable results had been achieved. In this chapter, we review and discuss recent results on a mathematical framework of autocatalytic sets that could take the idea out of the conceptual realm, and provide a formal and powerful way to investigate autocatalytic sets in the context of the origin of life, evolution, and functional organization.

## 4.1 Introduction

Origin of life research seems divided between two paradigms: *genetics-first* and *metabolism-first*. However, one element that both these paradigms have in common is that of *autocatalysis*: molecules catalyzing their own or each others replication, in a mutually beneficial way. The idea of such *autocatalytic sets* could form a bridge between the opposing paradigms.

Autocatalytic sets are generally defined as a set of molecules and catalyzed chemical reactions between them, where each molecule type is created by at least

W. Hordijk (✉)
SmartAnalytiX.com, Lausanne, Switzerland
e-mail: wim@WorldWideWanderings.com

M. Steel
Biomathematics Research Centre, University of Canterbury, Christchurch, New Zealand
e-mail: mike.steel@canterbury.ac.nz

S. Kauffman
University of Vermont, Burlington, VT, USA
e-mail: stukauffman@gmail.com

P. Pontarotti (ed.), *Evolutionary Biology: Exobiology and Evolutionary Mechanisms*,
DOI: 10.1007/978-3-642-38212-3_4, 

one reaction from this set, and each reaction is catalyzed by at least one molecule type from the set. Moreover, an autocatalytic set is usually required to be *self-sustaining* in the sense that every molecule in the set can be built up from some ambient 'food' source (molecules assumed to be available from the environment), by using reactions only from within the autocatalytic set itself.

This idea of collectively autocatalytic sets is not new. It was introduced more or less independently several times from the 1970s onward (Kauffman 1971, 1986; Eigen and Schuster 1977, 1979; Dyson 1982, 1985), and used in various origin of life models and scenarios (Wächterhäuser 1990, 2007; Gánti 1997; Rosen 1991; Letelier et al. 2006). Of course the idea has received criticism as well (Lifson 1997; Orgel 2008; Vasas et al. 2010), but recent experimental evidence shows the possibility and viability of such sets (Sievers and von Kiedrowski 1994; Lee et al. 1997; Ashkenasy et al. 2004; Hayden et al. 2008; Taran et al. 2010; Vaidya 2012).

In some of the original arguments to support the idea, it was claimed that in sufficiently complex chemical reaction systems (CRSs) (i.e., with a large enough diversity of molecule types), autocatalytic sets will arise almost inevitably (Kauffman 1971, 1986, 1993). This was subsequently disputed by pointing out that this argument actually requires an exponential growth rate in catalytic activity with increasing system size, something which is chemically highly unrealistic (Lifson 1997). So, the question of the probability of autocatalytic sets arising spontaneously in arbitrary CRSs remained open, with arguments both for and against mostly being based on informal qualitative arguments, and a thorough mathematical analysis was still lacking.

In more recent work on autocatalytic sets (Steel 2000; Hordijk and Steel 2004, 2012b, a; Mossel and Steel 2005; Hordijk et al. 2010, 2011, 2012), we have placed these ideas within a firm mathematical framework, and investigated the probability and structure of autocatalytic sets both mathematically and computationally. Not only has this led to a confirmation of the original claim (although with a more complicated mathematical proof), it has also provided many new and interesting insights into the conditions under which autocatalytic sets are likely to emerge, and their actual structure such as size and composition. Furthermore, our formal framework includes an efficient algorithm for actually detecting autocatalytic sets in arbitrary CRSs (an important tool that was not available before).

Given the recent experimental, theoretical, and computational progress, it seems more and more likely that autocatalytic sets played an important role, perhaps in more than one way, in the origin of life (Morowitz et al. 2000; Braakman and Smith 2012; Martin and Russel 2007; Penny 2005; Lane 2009; Hordijk et al. 2010; Kauffman 2011). Furthermore, we claim that autocatalytic sets might also play an important role in the evolution and functional organization of living systems in general. As such, we argue for a generalized theory of autocatalytic sets that could also be applied to living cells, entire ecologies, and perhaps even the economy.

In this chapter, we provide a non-technical overview of our formal framework for autocatalytic sets, some of the main results obtained with it so far, and our arguments for how it might be generalized and applied beyond the origin of life.

For full details, we refer to the original publications (Kauffman 1986, 1993; Steel 2000; Hordijk and Steel 2004, 2012a, b; Hordijk et al. 2010, 2011, 2012).

## 4.2 A Formal Framework of Autocatalytic Sets

### 4.2.1 Chemical Reaction Systems

In our framework, a *CRS* is defined as a set $X$ of molecule types, a set $\mathscr{R}$ of possible chemical reactions between these molecule types (transforming reactants into products), and a catalysis set $C$ indicating which molecule types can catalyze which reactions. A catalyst speeds up the rate at which a reaction happens (sometimes by several orders of magnitude), without being "used up" in the reaction itself. Mathematically, a CRS is written as a collection $Q = \{X, \mathscr{R}, C\}$. We also include the notion of a *food set* $F$, which is a (small) subset of molecule types that can be assumed freely available in the environment.

When using mathematical models to describe real-world systems, it is desirable to start with the simplest type of model that still captures the essence of the behavior of interest in the real system. Here, we use a simple model of CRSs known as the *binary polymer model* (Kauffman 1986, 1993). In this model, the molecule types are represented by bit strings up to a certain length $n$. The food set $F$ consists of a few small bit strings, for example all bit strings of length one and two (i.e., six food molecules in total). The possible reactions are simply *ligation* and *cleavage*: two bit strings are "glued" together into one longer one (ligation), or one bit string is "cut" into two smaller pieces (cleavage).

The catalysis set $C$ in the binary polymer model is generated at random as follows. Assume a probability $p$ that a given molecule $x \in X$ catalyzes a given reaction $r \in \mathscr{R}$. To generate an instance of the binary polymer model, each possible molecule-reaction pair $(x, r) \in (X, \mathscr{R})$ is included in the catalysis set $C$ independently with probability $p$. The main idea behind modeling catalysis at random is that, overall, little is known about which molecules can catalyze which reactions [predicting catalysis is known to be a hard problem (Kayala et al. 2011)].

Together, the set $X$ of bit strings up to length $n$, the set $\mathscr{R}$ of possible ligation and cleavage reactions, and a catalysis set $C$ generated at random with probability $p$, form a CRS $Q = \{X, \mathscr{R}, C\}$, thus providing a simple model of CRSs with tunable parameters $n$ and $p$.

### 4.2.2 Autocatalytic Sets

Given a CRS $Q = \{X, \mathscr{R}, C\}$ and a food set $F \subset X$, an *autocatalytic set* (or RAF set, in our terminology) is now defined as a subset $\mathscr{R}' \subseteq \mathscr{R}$ of reactions (and associated molecule types) which is:

1. *reflexively autocatalytic* (RA): each reaction $r \in \mathscr{R}'$ is catalyzed by at least one molecule type involved in $\mathscr{R}'$, and
2. *food-generated* (F): all reactants in $\mathscr{R}'$ can be created from the food set $F$ by using a series of reactions only from $\mathscr{R}'$ itself.

For a full mathematical definition of RAF sets, see Hordijk and Steel (2004); Hordijk et al. (2011), where we also introduced an efficient algorithm for detecting RAF sets in a general CRS $Q$.

Figure 4.1 shows an example of an RAF set that was found by our algorithm in an instance of the binary polymer model with $n = 5$. Black dots indicate molecule types, and white boxes indicate reactions. Solid black arrows indicate reactants going into and products coming out of a reaction, and dashed gray arrows indicate catalysis.

The RAF sets that are found by the RAF algorithm are called *maximal* RAF sets (maxRAFs). However, as we will show below, a maxRAF can often be decomposed into several smaller subsets which themselves are RAF sets (subRAFs). If such a subRAF cannot be reduced any further without losing the RAF property, it is called an *irreducible* RAF (irrRAF).

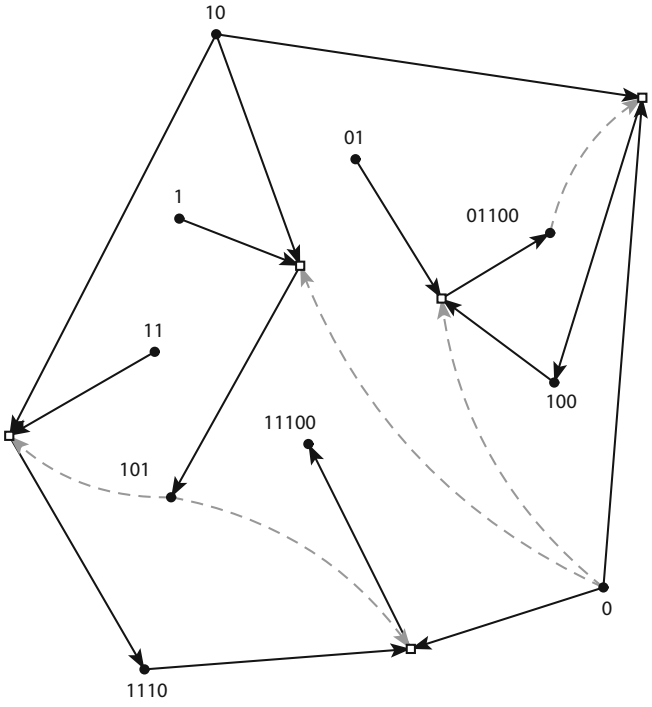



**Fig. 4.1** An example of an RAF set that was found in an instance of the binary polymer model with $n = 5$, and food set $F = \{0, 1, 00, 01, 10, 11\}$

## 4.3 The Emergence and Evolution of Autocatalytic Sets

To investigate the probability, structure, and dynamics of autocatalytic sets in CRSs, we used the RAF framework and the binary polymer model to perform and analyze computational simulations and derive mathematical properties and theorems. Here, we present a brief review of the main results and insights obtained so far.

### *4.3.1 The Probability of Autocatalytic Sets*

In Hordijk and Steel (2004) we showed computationally, and then proved mathematically in Mossel and Steel (2005), that at most a *linear* growth rate in the level of catalysis (with increasing maximum molecule length $n$) is sufficient for autocatalytic sets to exist with high probability. The level of catalysis is defined here as the average number of reactions catalyzed by any one molecule. Note that this result is in stark contrast with the required exponential growth rate in the original argument (Kauffman 1986, 1993). Furthermore, each molecule only needs to catalyze (on average) between one and two reactions to get a high probability of autocatalytic sets (for $n$ up to 50, at least) (Hordijk and Steel 2004; Hordijk et al. 2010). Thus, it seems that autocatalytic sets are highly likely, even for very moderate levels of catalysis.

One might argue that the binary polymer model is perhaps too simple to say much about reach chemical systems. However, in recent experiments, an almost literal chemical implementation of the binary polymer model was used, which has the potential to form autocatalytic sets (Taran et al. 2010). Furthermore, the theoretical linear upper bound on the growth rate in the required level of catalysis also holds for alphabets other than binary, e.g., polymers consisting of four building blocks (such as nucleotides), or 20 (such as amino acids), or any number. Finally, we also investigated several more realistic model extensions, such as *template-based catalysis*. The idea here is that a molecule must match a certain number of bits around the reaction site to be considered a candidate catalyst (similar, for example, to base-pair bonding in RNA molecules). In this case, the linear bound on the required growth rate in the level of catalysis still holds (Hordijk et al. 2011). Moreover, the required level of catalysis in the template-based case can be predicted mathematically from the purely random case (Hordijk and Steel 2012b).

### *4.3.2 The Structure of Autocatalytic Sets*

Having established the high likelihood of autocatalytic sets, we next investigated the structure of such sets. In particular, it turns out that RAF sets can often be decomposed into smaller RAF subsets (subRAFs) (Hordijk et al. 2012). For

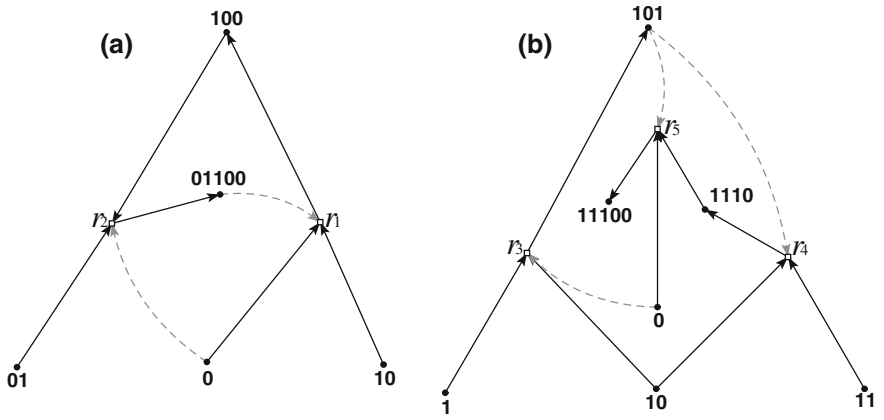



**Fig. 4.2** The two independent subRAFs $\mathscr{R}_1$ (*left*) and $\mathscr{R}_2$ (*right*) of which the maxRAF shown in Fig. 4.1 is composed

example, consider the maxRAF shown in Fig. 4.1. This RAF set, consisting of five reactions, can be decomposed into two smaller and independent subRAFs $\mathscr{R}_1$ and $\mathscr{R}_2$ of size two and three, respectively. These subRAFs are shown in Fig. 4.2.

Furthermore, the subRAF $\mathscr{R}_2$ itself contains even smaller subRAFs. In fact, any given RAF set can (potentially) contain an exponentially large number of subRAFs. All these subRAFs together form a partially ordered set, which can be visualized in a so-called *Hasse diagram*. Such a diagram of the subRAFs of the RAF set in Fig. 4.1 is shown in Fig. 4.3.

This decomposability of RAF sets provides a possible mechanism for evolution to happen in autocatalytic sets (Vasas et al. 2012; Hordijk et al. 2012). The process of combining and recombining different subRAFs into various (compartmentalized) reaction sets, giving rise to inheritance with variation, selection, and thus *evolvability*, was convincingly shown to work in Vasas et al. (2012). Furthermore, in Hordijk et al. (2012) we also argued how such a process can give rise to the emergence of higher-level structures (autocatalytic sets of autocatalytic sets). In that sense, the Hasse diagram shows the possible "trajectories" by which autocatalytic sets can emerge and evolve.

### *4.3.3 The Dynamics of Autocatalytic Sets*

To confirm these ideas about the emergence and evolution of autocatalytic sets, we also performed actual molecular flow simulations on RAF sets. Given a (max)RAF set $\mathscr{R}$ and starting with only molecules from the food set $F$, we used the well-known Gillespie algorithm (Gillespie 1976, 1977) to simulate the flow of molecules on the given reaction network. What these simulations show, is that indeed

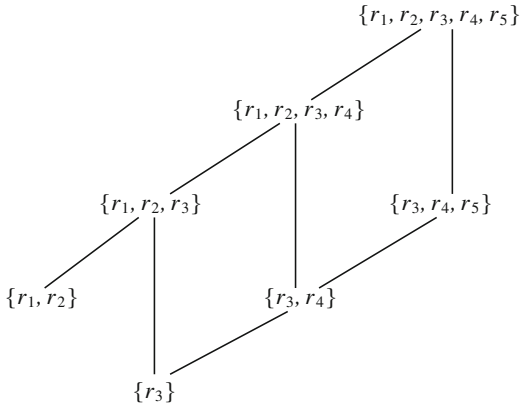



**Fig. 4.3** The Hasse diagram of all subRAFs contained within the RAF set shown in Fig. 4.1

the reaction system generally starts out with smaller (irreducible) RAF sets, which over time grow and combine into larger ones (Hordijk and Steel 2012a). Moreover, the existence of some (sub)RAFs can give rise to the coming into existence of other (sub)RAFs. For example, a particular subRAF $\mathscr{R}_B$ might be dependent on a molecule produced by another subRAF $\mathscr{R}_A$, and so $\mathscr{R}_B$ can only come into existence once $\mathscr{R}_A$ exists.

Figures 4.4 and 4.5 show a simple example [taken from Hordijk and Steel (2012a)]. Figure 4.4 shows another (maximal) RAF set that was found by the RAF algorithm in an instance of the binary polymer model. The colored boxes show the various subRAFs that exist within this maxRAF. Note that some boxes (subRAFs) are nested (such as the purple box within the red box), and some boxes depend on or form possible extensions of yet other boxes (such as the green box forming an extension of the blue box).

Figure 4.5 shows the results of a molecular flow simulation on the RAF network in Fig. 4.4. Starting with only food molecules, at first only the concentrations of the product molecules of the smallest (irreducible) subRAFs start to increase (the yellow and purple lines; line colors in Fig. 4.5 correspond to the colors of the boxes in Fig. 4.4). The blue subRAF actually needs to be "seeded", i.e., one of its three reactions needs to happen "spontaneously" (uncatalyzed) at least once for it to come into existence. Such an uncatalyzed reaction is always possible, but at a much lower rate than a catalyzed reaction. However, around time point 0.3 one such spontaneous reaction happens, and the blue subRAF comes into existence. After a while (around time point 0.5), the green extension of the blue subRAF (on which it depends) also comes into existence. Finally, the purple subRAF grows into the red subRAF after a last necessary spontaneous reaction happens around time point 0.55.

Furthermore, additional simulations on simple (collaborative) autocatalytic sets and equivalent but (selfish) non-autocatalytic sets of reactions confirm that autocatalytic sets do indeed have an advantage in terms of outperforming their non-autocatalytic competitors and in being more robust against perturbations.

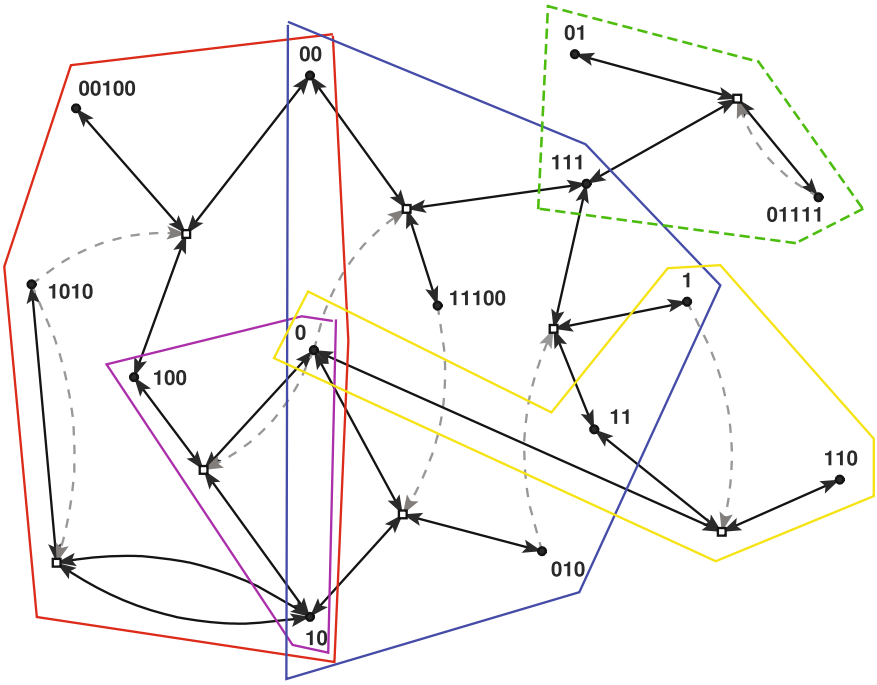



**Fig. 4.4** Another example of a (maximal) RAF set found by the RAF algorithm in an instance of the binary polymer model. The *colored boxes* indicate the various RAF subsets that are contained within the maxRAF

### *4.3.4 The Importance of Autocatalytic Sets*

In conclusion, using the formal RAF framework we have shown that autocatalytic sets are highly likely to emerge in CRSs, that they have the ability to evolve, and that they can outcompete, and are more robust than, equivalent non-autocatalytic sets. Note, however, that the examples shown here are deliberately kept small, so they can be visually inspected and understood. As a consequence, though, they are limited in their ability to show the full extent of possible structures, inter-dependencies, and dynamics that can (and do) occur in larger instances. But we trust that the simple examples provided here are sufficient to convey the main results and implications, at least conceptually.

Moreover, recent results show that the RAF framework can be applied directly and meaningfully to *real* CRSs exhibiting autocatalytic sets (Hordijk and Steel 2013) [in this case an RNA replicator system (Vaidya 2012)]. Not only does the model replicate many of the original experimental results, it also provides additional insights and predictions that would be difficult or impossible to obtain from experiments alone. This forms an important step toward merging experimental and theoretical lines of work on autocatalytic sets and their emergence and evolution.

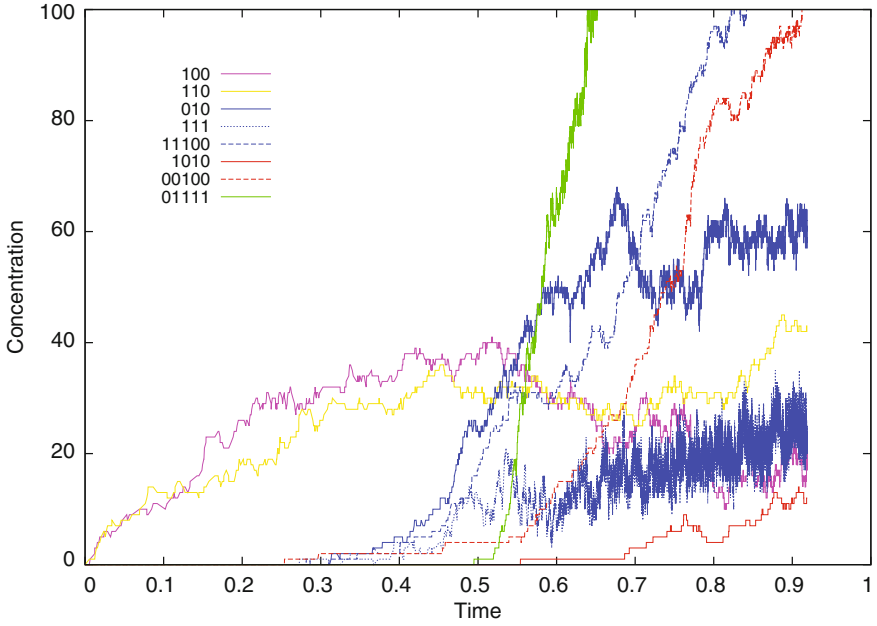



**Fig. 4.5** The results of a molecular flow simulation on the maxRAF shown in Fig. 4.4. The *lines* (colors correspond to the boxes in Fig. 4.4) show the concentrations of the reaction products over time, starting from an initial concentration of food molecules only

As mentioned in the introduction, the general idea behind autocatalytic sets is not new, as it was already developed during the 1970s. In fact, this idea of catalytic (or functional) closure and self-sustainability as a necessary property for living systems is incorporated in several similar formalisms, such as autopoietic systems (Varela et al. 1974; Maturana and Varela 1980), (M,R) systems (Rosen 1991), the Chemoton model (Gánti 1997), and various others. However, many of these formalisms remained mostly at a conceptual level, and have not generated the same richness and depth of results as the RAF framework. As summarized in this brief chapter, with this new framework we have obtained clear and concrete insights into the emergence, evolution, and structure of autocatalytic sets, which has direct relevance for a possible mechanism for the origin of life.

## 4.4 A Generalized Theory of Autocatalytic Sets

The original RAF algorithm is specifically formulated in the context of CRSs. However, it is possible to state the algorithm in a more general form. This is useful for extending it in further directions, both within the context of chemical systems as well as for other applications.

In Hordijk and Steel (2012a), we presented a generalized RAF algorithm, which we refer to as the "$gf-$algorithm". We then also showed that this algorithm can efficiently (i.e., in polynomial time) solve the HORN-SAT problem, which is a basic problem in propositional logic. HORN-SAT is of interest, as it is P-complete, which means that every problem in the computational complexity class P (i.e., every problem for which an efficient (polynomial-time) algorithm exists), can be reduced to HORN-SAT. Consequently, the $gf-$algorithm can efficiently solve any problem in the problem class P, and can thus be applied to a wide variety of problems other than finding autocatalytic sets in CRSs. One example was presented in Hordijk and Steel (2012a) where we described a simple toy problem in economics that can be solved with the $gf-$algorithm.

This generalization makes it clear that the RAF algorithm can be applied in a wider context outside of chemistry and beyond the origin of life. In a similar way, we would like to propose that the entire RAF framework, i.e., the notion of an autocatalytic set as a functionally closed, self-sustaining set, can be generalized. Given that autocatalytic sets of molecules and chemical reactions can grow, (re)combine, evolve, and potentially give rise to the emergence of higher-level functionally closed structures, we argue that a living cell could perhaps also be considered an autocatalytic (super)set, consisting of several inter-dependent autocatalytic subsets. Taking this a step further, what about collections of inter-dependent living cells, such as bacterial colonies or multicellular organisms? And finally, to take the argument the whole way, perhaps even an entire ecology of organisms can be viewed in this way.

Moreover, if interdependent groups of cells or multicellular organisms can be viewed as collections of interacting autocatalytic sets, then the same might hold for interdependent groups of people or organizations, i.e., an economy. For example, a "production function" (such as converting raw materials into products) can be viewed as the economic equivalent of a chemical reaction. The equivalent of catalysts are then products that facilitate these production functions, but are not "used up" by them, and which themselves are the result of other production functions. As a simple example, consider several pieces of wood and some nails as "reactants", hammering them together as a "reaction", and the result, say a table, as the "product", with a hammer being the catalyst, itself the product of some other reaction. In other words, an economy as a functionally closed, self-sustaining autocatalytic set, where the coming into existence of one subset (e.g., the Internet) enables the coming into existence of other, new subsets that could not have existed before (e.g., a company such as Google).

These arguments are still rather speculative, but as mentioned in the previous section, this idea of thinking about living (and perhaps even social) systems as funtionally closed, self-sustaining entities is not new. However, with the RAF framework, we claim to have a concrete, mathematically well-founded, and widely applicable formalism to investigate these ideas in a more direct, meaningful, and detailed manner. If a system of interest, be it chemical, biological, or social, can be defined within the context of the RAF framework, the tools we have developed can be applied directly to gain more insight into the emergence, evolution, and

structure of autocatalytic sets within these systems. Indeed, we are actively discussing and exploring these ideas with ecologists, economists, and social and cognitive scientists who are interested in finding ways to apply the RAF framework in their own fields of study. We truly hope that these combined efforts will eventually lead to a generalized theory of autocatalytic sets, and a better understanding of functional organization and emergence, not only in the context of the origin of life, but well beyond.

**Acknowledgments** This work was partly funded by the *Allan Wilson Centre for Molecular Ecology and Evolution* (AWCMEE).

## References

Ashkenasy G, Jegasia R, Yadav M, Ghadiri MR (2004) Design of a directed molecular network. PNAS 101(30):10872–10877

Braakman R, Smith E (2012) The emergence and early evolution of biological carbon-fixation. PLoS Comput Biol 8(4):e1002455

Dyson FJ (1982) A model for the origin of life. J Mol Evol 18:344–350

Dyson FJ (1985) Origins of life. Cambridge University Press, Cambridge

Eigen M, Schuster P (1977) The hypercycle: a principle of natural self-organization. Part A: emergence of the hyperccyle. Naturwissenschaften 64:541–565

Eigen M, Schuster P (1979) The hypercycle. Springer, Berlin

Gánti T (1997) Biogenesis itself. J Theor Biol 187:583–593

Gillespie DT (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J Comput Phys 22:403–434

Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. J Phys Chem 81(25):2340–2361

Hayden EJ, von Kiedrowski G, Lehman N (2008) Systems chemistry on ribozyme self-construction: Evidence for anabolic autocatalysis in a recombination network. Angewandte Chemie Int Ed 120:8552–8556

Hordijk W, Hein J, Steel M (2010) Autocatalytic sets and the origin of life. Entropy 12(7):1733–1742

Hordijk W, Kauffman SA, Steel M (2011) Required levels of catalysis for emergence of autocatalytic sets in models of chemical reaction systems. Int J Mol Sci 12(5):3085–3101

Hordijk W, Steel M (2004) Detecting autocatalytic, self-sustaining sets in chemical reaction systems. J Theor Biol 227(4):451–461

Hordijk W, Steel M (2012) Autocatalytic sets extended: dynamics, inhibition, and a generalization. J Syst Chem 3:5

Hordijk W, Steel M (2012) Predicting template-based catalysis rates in a simple catalytic reaction model. J Theor Biol 295:132–138

Hordijk W, Steel M (2013) A formal model of autocatalytic sets emerging in an RNA replicator system. J Syst Chem 4:3

Hordijk W, Steel M, Kauffman S (2012) The structure of autocatalytic sets: Evolvability, enablement, and emergence. Acta Biotheor 60(4):379–392

Kauffman SA (1971) Cellular homeostasis, epigenesis and replication in randomly aggregated macromolecular systems. J Cybern 1(1):71–96

Kauffman SA (1986) Autocatalytic sets of proteins. J Theor Biol 119:1–24

Kauffman SA (1993) The origins of order. Oxford University Press, Oxford

Kauffman SA (2011) Approaches to the origin of life on earth. Life 1:34–48

Kayala MA, Azencott C-A, Chen JH, Baldi P (2011) Learning to predict chemical reactions. J Chem Inf Model 51:2209–2222

Lane N (2009) Life ascending: the ten great inventions of evolution. Profile Books, Norton

Lee DH, Severin K, Ghadiri MR (1997) Autocatalytic networks: the transition from molecular self-replication to molecular ecosystems. Curr Opin Chem Biol 1:491–496

Letelier J-C, Soto-Andrade J, Abarzúa FG, Cornish-Bowden A, Cárdenas ML (2006) Organizational invariance and metabolic closure: analysis in terms of (M;R) systems. J Theor Biol 238:949–961

Lifson S (1997) On the crucial stages in the origin of animate matter. J Mol Evol 44:1–8

Martin W, Russel MJ (2007) On the origin of biochemistry at an alkaline hydrothermal vent. Philos Trans R Soc B 362:1887–1925

Maturana HR, Varela FJ (1980) Autopoiesis and cognition: the realization of the living. Reidel Publishing Co., Dordrecht

Morowitz HJ, Kostelnik JD, Yang J, Cody GD (2000) The origin of intermediary metabolism. PNAS 97(14):7704–7708

Mossel E, Steel M (2005) Random biochemical networks: the probability of self-sustaining autocatalysis. J Theor Biol 233(3):327–336

Orgel LE (2008) The implausibility of metabolic cycles on the prebiotic earth. PLoS Biology 6(1):5–13

Penny D (2005) An interpretive review of the origin of life research. Biol Philos 20:633–671

Rosen R (1991) Life itself. Columbia University Press, New York

Sievers D, von Kiedrowski G (1994) Self-replication of complementary nucleotide-based oligomers. Nature 369:221–224

Steel M (2000) The emergence of a self-catalysing structure in abstract origin-of-life models. Appl Math Lett 3:91–95

Taran O, Thoennessen O, Achilles K, von Kiedrowski G (2010) Synthesis of information-carrying polymers of mixed sequences from double stranded short deoxynucleotides. J Syst Chem 1(9)

Vaidya N, Manapat ML, Chen IA, Xulvi-Brunet R, Hayden EJ, Lehman N (2012) Spontaneous network formation among cooperative RNA replicators. Nature 491:72–77

Varela FJ, Maturana HR, Uribe R (1974) Autopoiesis: the organization of living systems, its characterization and a model. Biosystems 5(4):187–196

Vasas V, Fernando C, Santos M, Kauffman S, Sathmáry E (2012) Evolution before genes. Biol Direct 7:1

Vasas V, Szathmáry E, Santos M (2010) Lack of evolvability in self-sustaining autocatalytic networks constraints metabolism-first scenarios for the origin of life. PNAS 107(4): 1470–1475

Wächterhäuser G (1990) Evolution of the first metabolic cycles. PNAS 87:200–204

Wächterhäuser G (2007) On the chemistry and evolution of the pioneer organism. Chem Biodivers 4:584–602

# Chapter 5
# Interactions of Clay Minerals with RNA Components

**Hideo Hashizume, Sjerry van der Gaast and Benny K. G. Theng**

**Abstract** Because of their extensive surface area, layer structure, and surface charge characteristics, clay minerals can take up a wide range and variety of organic molecules. Further, clay minerals can shield these molecules from cosmic and ultraviolet radiation, and catalyze their polymerization. For these reasons, clay minerals might have played an important role in chemical evolution and the origins of life on Earth. The proposal that ribonucleic acid (RNA) can act as both a storehouse of genetic information and an enzyme-like catalyst in the primordial Earth, has stimulated research into the ability of clay minerals to catalyze the formation of RNA from its (activated) monomers. After outlining the probable role of clay minerals in chemical evolution and the origins of life, and summarizing clay minerals structures, we describe the interactions of clay minerals with nucleic acid bases, nucleosides, nucleotides, polynucleotides, and nucleic acids. These interactions are illustrated by selective experimental results from our laboratory and the literature.

## 5.1 Introduction

Oparin (1924, 1938) and Haldane (1929) proposed that the ocean of the prebiotic Earth contained a mixture of simple organic molecules arising from the action of an energy source on an atmosphere dominated by methane, ammonia, and hydrogen. In support of this proposal, Miller (1953), working in Urey's laboratory

H. Hashizume (✉)
National Institute for Materials Science, Tsukuba 305-0044, Japan
e-mail: HASHIZUME.Hideo@nims.go.jp

S. van der Gaast
Hugo de Grootstraat 37 2311XK Leiden, The Netherlands

B. K. G. Theng
Landcare Research, Palmerston North 4442, New Zealand

P. Pontarotti (ed.), *Evolutionary Biology: Exobiology and Evolutionary Mechanisms*,
DOI: 10.1007/978-3-642-38212-3_5, 

(Miller and Urey 1959), obtained a number of amino acids by subjecting a mixture of these primordial "reducing" gases to an electric discharge. Subsequently, Oró (1961) and Oró and Kimball (1961) reported the formation of adenine from hydrogen cyanide, while Sanchez et al. (1966) were able to synthesize pyrimidines from cyanoacetylene. The Oparin-Haldane hypothesis is further substantiated by the occurrence in meteorites of some of the amino acids formed in the Miller–Urey experiments together with nucleobases and related biomonomers (Orgel 1994; Burton et al. 2012). It seemed likely that these extraterrestrial organic molecules were protected through interaction with clay and mineral surfaces, enabling them to survive the high pressure and temperature conditions that prevailed during meteorite entry into the atmosphere and its subsequent collision with the Earth surface. The probability that meteorites (and comets) were instrumental in delivering biomolecules to the early Earth has been reviewed by Pizzarello (2004).

Bernal (1951) was the first to propose that clay minerals played an important role in chemical evolution and the origins of life on the basis that clay minerals could concentrate organic molecules, shield them against ultraviolet and cosmic radiation, and catalyze their polymerization. Cairns-Smith (1982, 2005) went so far as to suggest that clay minerals could act as a primitive genetic material, capable of replicating information stored in the form of structural irregularities, such as faults in layer stacking and cation distribution. At some geological time point, there was a 'genetic takeover' by a self-replicating organic species, such as ribonucleic acid (RNA). In this hypothetical 'RNA world' (Gilbert 1986; Orgel 1998), RNA acted as both a genetic material (without the aid of proteins) and an enzyme (capable of catalyzing protein synthesis). The discovery of catalytic RNA enzymes, termed 'ribozymes,' by Cech (1986) and S. Altman (Guerrier-Takada and Altman 1984) has focused attention on the role of RNA in the origins of life on Earth (Srivatsan 2004).

Strašák and Šeršeň (1991) have reported that $Cu^{2+}$-montmorillonite is capable of catalyzing the conversion of adenine and adenosine into hypoxanthine and inosine at room temperature. There is no experimental evidence to show, however, that clay minerals can catalyze the formation of nucleotides by combining nucleobases, ribose, and phosphate, let alone that of polynucleotides. But once an RNA oligomer has formed, long-chain RNA with up to 50-mers can be obtained in the presence of montmorillonite (Ferris and Ertem 1993; Ferris 2005). The finding by Biondi et al. (2007) that montmorillonite can protect ribozymes against degradation by ultraviolet radiation lends further support to the idea that clay minerals play an active role in the evolution of the RNA world (Gallori et al. 2006).

Here we review the interactions of clay minerals with nucleic acid bases, nucleosides, nucleotides, polynucleotides, and nucleic acids, present some experimental results, and summarize the ability of clay minerals in catalyzing the synthesis of RNA.

## 5.2 Clay Minerals: Structures and Surface Properties

Clay minerals, formed by weathering of primary rock-forming silicates through the action of water, air, and steam (Galan 2006; Theng and Yuan 2008), have been detected in early Archean rocks from Isua, West Greenland (Appel 1980) and the Pilbara, Western Australia (Cullers et al. 1993) that are older than 3.7 Ga when 'life' on Earth began to emerge (Mojzsis et al. 1996). Besides their likely presence in the primitive Earth and on Mars (Ehlmann et al. 2011), clay minerals are known for their ability to take up and intercalate a large variety of simple and polymeric organic compounds (Theng 1974, 2012; Lagaly et al. 2006). The reactivity of clay minerals toward external solute molecules, including bioorganic compounds, may be ascribed to their ultrafine ('nanoscale') particle size, extensive surface area, and the presence of surface defects and dislocations (Theng and Yuan 2008).

Clay minerals are layer silicates (phyllosilicates), the basic building block of which is an alumino-silicate layer comprising a silica tetrahedral sheet and an alumina octahedral sheet, joined together in certain proportions. Condensation in a 1:1 proportion gives rise to a T-O layer structure (as in kaolinite and serpentine). Similarly, the T-O-T layer structure arises from condensation in a 2:1 proportion with the octahedral sheet being sandwiched between two opposing tetrahedral sheets (as in talc and pyrophyllite). Isomorphous substitution of $Si^{4+}$ (in tetrahedral positions) and/or $Al^{3+}$ (in octahedral sites) by cations of lower valency, gives rise to a negatively charged layer structure. In the absence or near-absence of isomorphous substitution, the interlayer space is essentially empty, or may be occupied by water molecules (as in halloysite). For the majority of 2:1 type phyllosilicates, however, the negative layer charge is balanced by hydrated cations (as in montmorillonite, beidellite, hectorite, and vermiculite), non-hydrated cations (as in mica and illite), or a metal hydroxide sheet (as in chlorite), occupying interlayer positions .

As the name suggests, the basic building block of layered double hydroxides (LDH) is typically a magnesia octahedral layer—more correctly sheet—in which a portion of the $Mg^{2+}$ ions are replaced by $Al^{3+}$. Since the resultant positive layer charge is balanced by sorption of inorganic or organic anions (in the interlayer space), LDH have also been referred to as 'anionic clays.' The basic structural features of clay minerals and LDH are shown in Fig. 5.1.

Allophane is a non-planar phyllosilicate of short-range order whose unit particle is a hollow spherule with an outer diameter of 4.0–5.5 nm (Fig. 5.1). The spherule wall, with a thickness of 0.7–1.0 nm, is composed of an outer gibbsitic sheet to which $O_3SiOH$ groups are attached on the interior. Defects in the wall structure give rise to perforations of $\sim$0.3 nm in diameter, exposing $Al(OH_2)$ (OH) groups that can either acquire or lose protons. The point of zero net charge of allophane ranges from pH 5 to 6. Besides being at the source of the pH-dependent charge characteristics, these groups control the reactivity of allophane toward phosphate and bioorganic molecules, such as amino acids and adenine-5′-phosphate (Theng et al. 1982; Hashizume and Theng 1999, 2007; Hashizume et al. 2002).

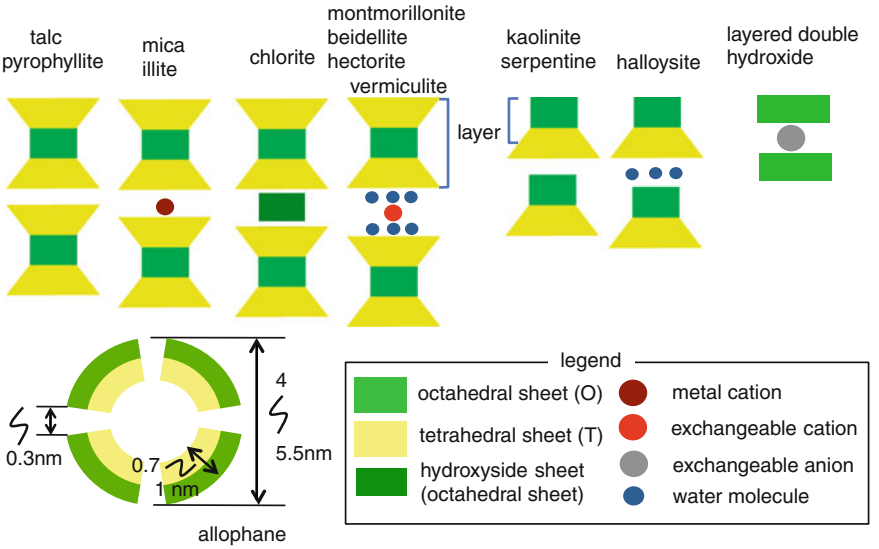



**Fig. 5.1** Diagram showing the basic structures of clay minerals

Because of its extensive interlayer surface area ($\sim$700 $m^2g^{-1}$), high pH-independent cation exchange capacity ($\sim$100 cmol $kg^{-1}$), and propensity for intercalating organic molecules, montmorillonite has held the limelight in clay-organic studies pertaining to chemical evolution and the origin of life. Being dispersible in the primordial ocean, montmorillonite would be able to concentrate dissolved organic solutes on basal and interlayer surfaces, and feasibly catalyze their oligomerization. We might also add that montmorillonite can be made organophilic (hydrophobic) by replacing the charge-balancing inorganic cations ('counterions') with long-chain alkyl ammonium and quaternary ammonium ions (Hedley et al. 2007). Such 'organoclays' have an enhanced affinity for extraneous organic solutes and pollutants (Churchman et al. 2006; Theng et al. 2008).

## 5.3 Nucleic Acid Bases

Lailach et al. (1968a, b) made an early attempt at investigating the interactions of purines, pyrimidines, and nucleosides with montmorillonite (and illite) as a function of pH, nature of the exchangeable cations, and molecular constitution. Figure 5.2 shows the adsorption of adenine and cytosine by $Li^+$-, $Na^+$-, $Mg^{2+}$-, and $Ca^{2+}$-exchanged montmorillonites as a function of solution pH. The extent of adsorption decreased slightly as pH increased from 2 to 6, then fell steeply between pH 6 and 8, and became hardly measurable above pH 8. This observation may largely be explained in terms of the basicity ($pK_a$) of the nucleic acid bases.

Adenine ($pK_a = 4.15$), for example, is positively charged at $pH < pK_a$, and hence was strongly adsorbed by electrostatic interactions (cation exchange) with the clay mineral surface. Since the molecule is essentially uncharged between pH 5 and 9, adsorption declined in this pH range. By the same token, very little adsorption occurred with thymine and uracil (not shown) because these molecules have no amino groups capable of acquiring protons, and hence are essentially uncharged even under highly acidic conditions (Perezgasga et al. 2005). Other contributing factors were proton transfer, complexation to the counterion, and molecular size and shape.

The pH-sorption curves for purines and pyrimidines with $Co^{2+}$-, $Ni^{2+}$-, and $Cu^{2+}$-montmorillonites in the acidic pH range were similar to those observed with the $Mg^{2+}$- and $Ca^{2+}$-exchanged counterparts (Fig. 5.2), indicative of a cation exchange process. In the weakly acidic to alkaline pH range, however, adsorption was primarily due to complex formation (coordination) with the exchangeable cations, probably involving the N-3 and N-9 positions (in the case of adenine, purine, and their substituted derivatives). Adsorption decreased in the order $Cu \gg Ni > Co$, reflecting the diminishing strength and stability of the respective complexes (Lailach et al. 1968b).

Benetoli et al. (2008) measured the adsorption of adenine, cytosine, uracil, and thymine by montmorillonite and kaolinite from (artificial) sea water at pH 2.0 and 7.2. All things being equal, more nucleic acid bases were adsorbed by montmorillonite than kaolinite. As would be expected, more nucleic acid base was adsorbed at pH 2 than pH 7.2 with larger amounts of adenine and cytosine being taken up in comparison with thymine and uracil. Fourier transform infrared (FT-IR) spectroscopy further indicated that at pH 2 all four nucleobases were adsorbed as the corresponding protonated species.

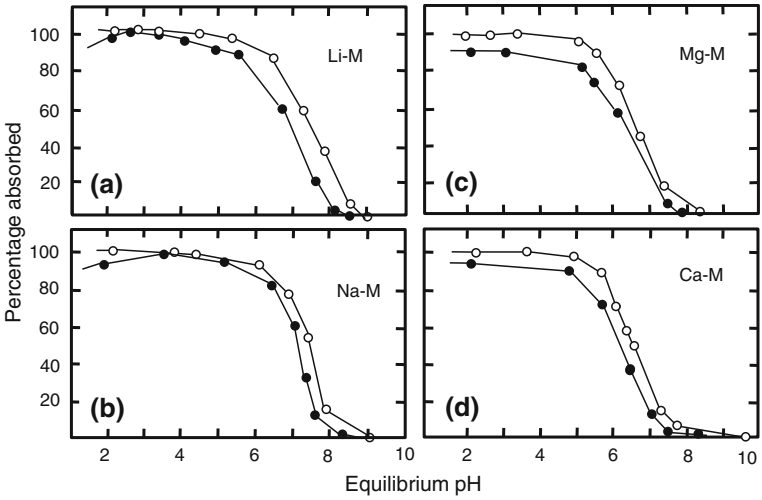



**Fig. 5.2** Adsorption of adenine *circle* and cytosine *filled circle* by **a** Li-; **b** Na-; **c** Mg-; and **d** Ca-montmorillonite as a function of solution pH (Lailach et al. 1968a)

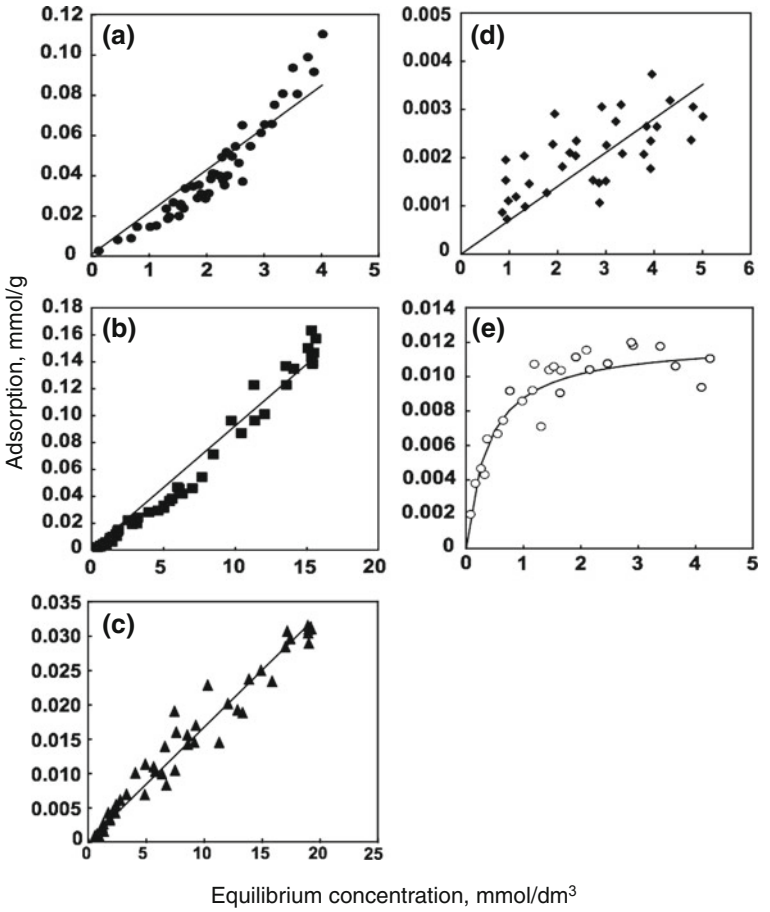



**Fig. 5.3** Isotherms for the adsorption of RNA compounds by Mg-montmorillonite. **a** adenine; **b** cytosine; **c** uracil; **d** ribose; and **e** phosphate (Hashizume et al. 2010)

More recently Hashizume et al. (2010) determined the adsorption of adenine, cytosine, and uracil by $Mg^{2+}$-montmorillonite at pH ∼ 8. The isotherms for adenine, cytosine, and uracil were of the C-type, the amount adsorbed increasing more or less linearly with solute concentration (Fig. 5.3a–c). At comparable solute concentrations, adsorption decreased in order adenine > cytosine > uracil. This observation was explained in terms of differences in their basicity ($pK_a$), molecular size, and aqueous solubility. They suggested that all three nucleic acid bases were adsorbed by coordination to $Mg^{2+}$ ions through a water bridge; that is, by H-bonding to water molecules in the primary hydration shell of interlayer $Mg^{2+}$ ions with the heterocyclic ring being oriented parallel to the interlayer surface. On the other hand, nucleobases can coordinate directly to transition metal ions, such as $Cu^{2+}$ and $Zn^{2+}$ (Weckhuysen et al. 1999).

At the same time, Hashizume et al. (2010) also measured the adsorption of ribose and phosphate. Figure 5.3d shows that the points for ribose were scattered and very little was adsorbed, presumably because this highly basic compound ($pK_a = 12.2$) had little tendency to coordinate to exchangeable $Mg^{2+}$ ions. The isotherm for phosphate was of the L- or Langmuir-type (Fig. 5.3e), giving a plateau adsorption of 0.012 mmol/g. The corresponding value for kaolinite was 0.007 mmol/g (Muljadi et al. 1966), while that for allophane was an order of magnitude greater (Theng et al. 1982). The result was consistent with a ligand exchange process, involving $Al(OH_2)$ (OH) groups exposed on the edge surface of montmorillonite particles: By the same token, ligand exchange (Eq. (5.1)) lies behind the large capacity of clay minerals, especially allophane, for taking up 5′-adenosine monophosphate (Graf and Lagaly 1980; Hashizume and Theng 2007).

$$Al(OH_2)(OH) + H_2PO_4^- \rightarrow Al(OH_2)(H_2PO_4) + OH^- \quad (5.1)$$

Using an acid-activated clay (montmorillonite-KSF), Carneiro et al. (2011) likewise observed that adsorption of adenine and cytosine was greater at pH 2 than at pH 7. At both pH 2 and 7, adsorption decreased in the order adenine $\approx$ cytosine > thymine > uracil. X-ray diffractometry showed that these biomolecules were able to penetrate the interlayer space of montmorillonite as Perezgasga et al. (2005) and Hashizume et al. (2010) had previously observed. Electron paramagnetic resonance (EPR) and Mössbauer spectroscopy further indicated that the intercalated nucleobases were able to oxidize structural $Fe^{2+}$ (in the octahedral sheet) to $Fe^{3+}$ although the underlying mechanism was not clearly defined. More importantly, intercalated nucleobases (e.g., cytosine) were highly stable against thermal degradation (Pucci et al. 2010).

Carneiro et al. (2011) also observed that montmorillonite-KSF that had been treated with $Na_2S$, could still adsorb appreciable amounts of adenine and cytosine. Since this treatment apparently caused the clay surface to become positively charged, they suggested that van der Waals forces played an important part in nucleobase adsorption. The ab initio study by Robinson et al. (2007), however, would indicate that H-bonding (of thymine and uracil) with surface oxygens (of dickite) is as important as, if not more so than, van der Waals interactions.

Earlier, Winter and Zubay (1995) measured the adsorption of adenine (and adenine-related compounds) by $Na^+$-montmorillonite in the presence of the disodium salt of piperazine-N,N′-bis-2-ethanesulfonic acid (PIPES buffer) at pH 6.7, or simulated sea salt solution at pH 8.6. More adenine was adsorbed from PIPES buffer than sea water because of the difference in pH. In both media, the isotherms for adenine were more or less linear as Hashizume et al. (2010) had reported for $Mg^{2+}$-montmorillonite. The extent of adsorption (at comparable solute concentration and pH), however, was markedly greater than what Hashizume et al. (2010) had observed (Fig. 5.3a). It would appear that the high ionic strength of the sea salt solution promoted adsorption by "screening" the negative surface charge on montmorillonite.

In contrast to the situation with montmorillonite, the isotherms for the adsorption of adenine by allophane decreased in the order pH 8 > pH 6 > pH 4 (Fig. 5.5a–c) (Hashizume and Theng 2007). This observation may be explained in terms of the pH-dependent charge characteristics of the mineral. Since the point of zero charge of the allophane sample used was close to 6 (Theng et al. 1982), the net surface charge was positive at pH 4, essentially zero at pH 6, and negative at pH 8. On the other hand, the positive charge on adenine decreased as pH increased from 4 to 8. Adenine adsorption (through electrostatic interactions) should, therefore, decrease in above-mentioned pH order. That the extent of adsorption was still appreciable at pH 4 would indicate that other attractive forces, such as van der Waals and H-bonding interactions, were operative.

## 5.4 Nucleosides

The adsorption of adenosine, guanosine, and cytidine by $Li^+$, $Na^+$, $Mg^{2+}$, and $Ca^{2+}$-exchanged montmorillonites as a function of solution pH was investigated by Lailach et al. (1968a). Figure 5.4 shows that the pH-sorption curves were similar in shape to those given by the corresponding nucleobases (Fig. 5.2) with adsorption falling steeply as pH increased from 3 to 8. The extent of nucleoside adsorption, however, was substantially smaller than that obtained with the nucleic acid bases. This observation may be ascribed to differences in molecular size and shape in that adenine and cytosine are planar whereas their respective nucleosides are non-planar due to the presence of ribose. The adsorption of nucleosides at pH < 5 by

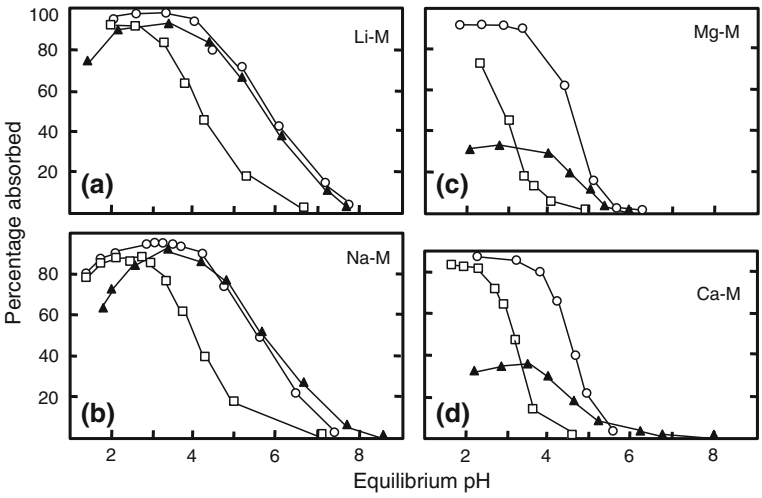



**Fig. 5.4** Adsorption of nucleosides by **a** Li-; **b** Na-; **c** Mg-; and **d** Ca-montmorillonite as a function of solution pH: *circle*, adenosine; *square*, guanosine; *filled triangle*, cytidine (Lailach et al. 1968a)

$Co^{2+}$-, $Ni^{2+}$-, and $Cu^{2+}$- montmorillonites resembled that obtained with the $Ca^{2+}$-exchanged clay, shown in Fig. 5.4d, while nucleoside adsorption at pH > 6 decreased in the order Cu ≫ Ni > Co. Adsorption also varied with nucleoside structure, decreasing in the order adenosine > cytidine ≫ guanosine ≫ inosine (Lailach et al. 1968b).

Figure 5.5d–f shows that the adsorption of adenosine by allophane resembled that of adenine in that uptake declined with solution pH, in the order pH 8 > pH 6 > pH 4 (Hashizume and Theng 2007). As for adenine, the explanation must be sought in the pH-dependent charge characteristics of both solute molecule and mineral surface. In this connection, we should also mention that at comparable solution concentration and pH, allophane can take up as much ribose (Fig. 5.5g–i) as adenine. On the other hand, very little ribose was taken up by montmorillonite as compared with adenine (Fig. 5.3).

## 5.5 Nucleotides

The interactions of nucleotides with clay minerals, notably montmorillonite, have been well documented. The early work by Rishpon et al. (1982) indicated that solution pH and the nature of the counterions (balancing the negative layer charge) had a marked effect on adsorption. Lawless et al. (1985) observed similarly for the adsorption of 5′-, 3′-, 2′-adenine monophosphates (AMP), and 5′-cytidine monophosphate (CMP) by $Na^{+}$-, $Mn^{2+}$-, $Fe^{3+}$-, $Co^{2+}$-, $Ni^{2+}$-, $Cu^{2+}$-, and $Zn^{2+}$-exchanged montmorillonites. They ascribed the influence of solution pH to the variation in nucleotide charge, ranging from cationic through zwitterionic to mono-, di-, and tri-anionic as pH increased from 2 to 12. At neutral pH, more 5′-AMP was adsorbed than 3′-AMP, 2′-AMP, and 5′-CMP. The counterions affected adsorption through their tendency to form nucleotide complexes of varying stability. Odom et al. (1979), for example, reported that the adsorption of 5′-AMP decreased in order $Zn^{2+} \approx Cu^{2+} > Fe^{3+} > Mn^{2+} > Ni^{2+} > Co^{2+}$, reflecting the high stability of 5′-AMP complexes with $Zn^{2+}$ and $Cu^{2+}$ as compared to those involving $Mn^{2+}$, $Fe^{3+}$, $Co^{2+}$, and $Ni^{2+}$ counterions. Similarly, Banin et al. (1985) observed that the adsorption of 5′-AMP (at a given pH) by mixed $Fe^{3+}$/$Ca^{2+}$-montmorillonites increased with the proportion of $Fe^{3+}$ ions occupying exchange sites.

Ferris et al. (1989a) also found that the binding of 3′- and 5′-AMP as well as 3′- and 5′-uridine monophosphates (UMP) was greatly enhanced when $Cu^{2+}$-montmorillonite was used instead of the $Na^{+}$-exchanged clay sample. Earlier, Ferris and Hagan (1986) reported that the affinity of 5′-AMP for $Zn^{2+}$-montmorillonite (as measured by the Langmuir adsorption coefficient), was three times greater in the presence of 0.2 M PIPES buffer than in its absence. In this instance, the enhancement in affinity was not due to the increase in ionic strength as this effect was not observed in the presence of 0.2 M sulfate. In combination with basal spacing measurements (by X-ray diffraction), the results were explained in terms

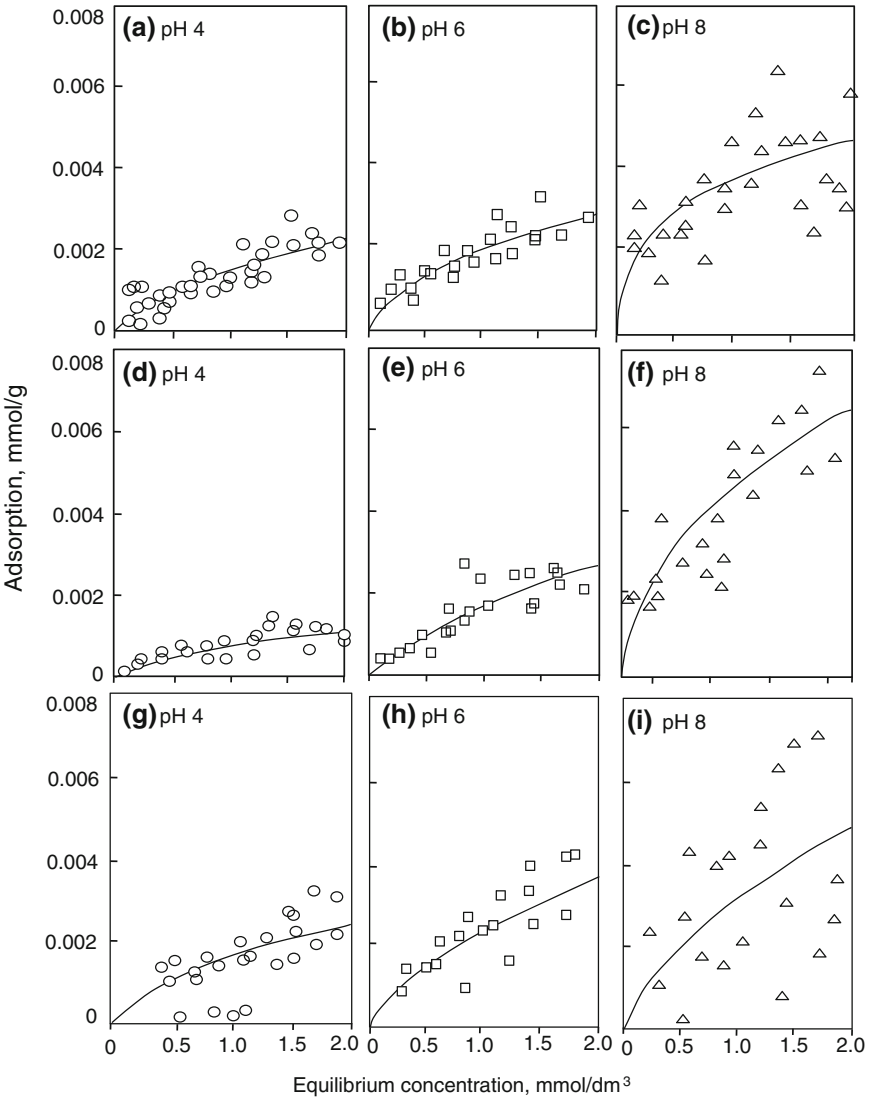



**Fig. 5.5** Isotherms for the adsorption of adenine, adenosine, and ribose by allophane: **a** adenine at pH 4; **b** adenine at pH 6; **c** adenine at pH 8; **d** adenosine at pH 4; **e** adenosine at pH 6; **f** adenosine at pH 8; **g** ribose at pH 4; **h** ribose at pH 6; and **i** ribose at pH 8: *circle* pH 4; *square* pH 6; *triangle* and pH 8 (Hashizume and Theng 2007)

of the formation of an interlayer Zn-PIPES complex, capable of mediating and promoting nucleotide adsorption.

The adsorption by $Na^+$-montmorillonite of "activated" mononucleotides (RNA monomers), formed by attaching an imidazole group to the phosphate moiety in adenosine (ImpA) or uridine (ImpU), was investigated by Wang and Ferris (2001).

The isotherms were of the L-type, giving rise to monolayer adsorption on the clay mineral surface. They suggested that the higher uptake of ImpA over ImpU reflected the stronger van der Waals interaction between the purine ring and the mineral surface. This suggestion was further supported by the finding that the binding and surface oligomerization of ImpA and ImpU were inhibited by $N^6$,$N^6$-dimethyladenine and $P^1$,$P^2$-dideoxyadenosine-5′-diphosphate ($dA^{5'}ppdA$).

More recently, Perezgasga et al. (2005) compared the adsorption (at pH 2) of AMP, adenine diphosphate (ADP), and adenine triphosphate (ATP) by $Na^+$-montmorillonite with that by the same clay mineral sample in which the interlayer sodium ions have largely been replaced by hexadecyltrimethyl ammonium (HDTMA) ions. The extent of adsorption by HDTMA-$Na^+$-montmorillonite was substantially less than by the original sodium-exchanged mineral, presumably because nucleotide intercalation into the organoclay was restricted due to high occupancy of the interlayer space by HDTMA ions.

Earlier, Graf and Lagaly (1980) measured the adsorption of 5′-AMP, 5′-ADP, and 5′-ATP by montmorillonite, beidellite, illite, kaolinite, quartz, and a synthetic carborundum. The clay minerals took up much more nucleotides than either quartz or carborundum because to former materials had a relatively large concentration of $Al(OH_2)(OH)$ groups at their particle edge surface to which nucleotides can bind through ligand exchange. Thus, prior treatment of the clay samples with sodium diphosphate led to a marked decrease in adsorption. On the other hand, OH groups attached to silicon on the surface of quartz and carborundum were relatively unreactive toward phosphate. The extent of adsorption by clay minerals decreased in the order 5′-ATP > 5′-ADP > 5′-AMP. Indeed, hardly any 5′-AMP was taken up although measurable adsorption occurred in the presence of 5′-ATP. In explanation, Graf and Lagaly (1980) suggested that the bulky adenosine moiety in AMP interfered with the ligand exchange process. As the phosphate group in 5′-ADP, and more so in 5′-ATP, is farther away from the adenosine moiety, steric hindrance would be less important in the case of the di- and tri-phosphate nucleotides.

We should point out, however, that Graf and Lagaly's (1980) measurements were made with clay minerals that had been coagulated using high concentrations of NaCl. At pH > 6, this pretreatment would induce edge(−)/face(−) and edge(−)/edge(−) particle associations (Lagaly and Ziesmer 2003). As a result, much of the clay particle edge surface would not be accessible to the nucleotides. Winter and Zubay (1995), for example, reported that $Na^+$-montmorillonite could take up small amounts of all three adenosine-5′-phosphates from PIPES-buffered solutions (pH 6.7). On the other hand, both flavin-adenine dinucleotide (FAD) and flavin mononucleotide (FMN) were adsorbed to greater than 70 % at pH ≤ 7. Interestingly, the positively charged nicotinamide adenine dinucleotide ($NAD^+$) was less strongly adsorbed than FAD.

Using $Cu^{2+}$-montmorillonite, Ferris et al. (1989a) observed that AMP and UMP were more strongly adsorbed than their respective nucleosides. This difference in adsorption behavior between nucleotide and nucleoside was much more pronounced in the case of allophane as shown by comparing Fig. 5.5d–f with Fig. 5.6

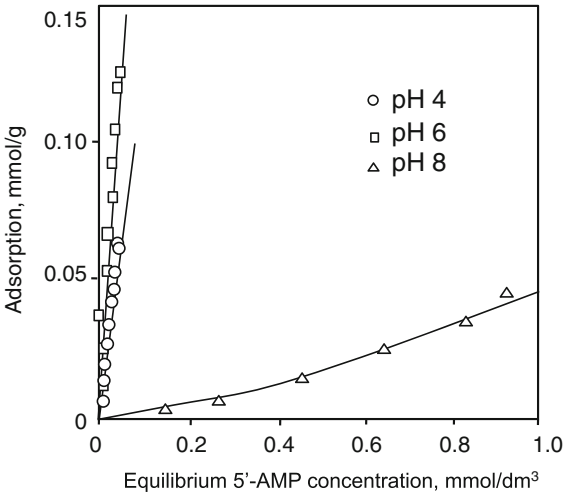



**Fig. 5.6** Isotherms for adsorption of 5′-AMP by allophane at pH 4 *circle*, 6 *square*, and 8 *triangle* (Hashizume and Theng 2007)

(Hashizume and Theng 2007). As already remarked on, the high affinity of 5′-AMP for allophane (at acidic pH) may be explained in terms of ligand exchange between the phosphate group of 5′-AMP and the hydroxyl of $Al(OH_2)(OH)$ groups, exposed at perforations on the surface of allophane spherules (Eq. 5.1). The substantial decrease in adsorption at pH 8 (Fig. 5.6) was due to solute-surface repulsion since both 5′-AMP and allophane were negatively charged under alkaline conditions.

## 5.6 Polynucleotides and Nucleic Acids

The interactions of clay minerals with polynucleotides have been widely investigated, while the formation and properties of clay-nucleic acid complexes have recently been reviewed by Theng (2012).

Perezgasga et al. (2005) reported that polyadenylic acid (poly-A) was strongly adsorbed by both $Na^+$- and HDTMA-$Na^+$-montmorillonite at pH 2 when nearly all of the polynucleotide in solution was removed. By comparison, only about 40 % of the total concentration of polyuridylic acid (poly-U) was taken up by $Na^+$-montmorillnite and HDTMA-$Na^+$-montmorillonite. Very little poly-A and poly-U, however, was adsorbed at pH 6 and 10. These observations may be explained in terms of differences in the structure and pH-dependent charge characteristics between the two polynucleotides. X-ray diffraction further indicated that poly-A failed to penetrate the interlayer space of montmorillonite, in common with the behavior of polyanions, in general (Theng 2012). They suggested that adsorption was largely effected through ligand exchange at the edge of montmorillonite particles (Eq. 5.1). In line with this suggestion, no measurable adsorption of poly-

A was observed with montmorillonite whose edge surface had been "blocked" by treatment with sodium tripolysphosphate.

Earlier, Franchi et al. (2003) investigated the effect of monovalent ($Na^+$) and divalent ($Mg^{2+}$, $Ca^{2+}$) cation concentrations (0.1–75 mM) on the adsorption by montmorillonite and kaolinite of a range of polynucleotides, including poly-A, poly-U, polydeoxyadenylic acid (poly-dA), polydeoxythymidylic acid (poly-dT), and chromosomal DNA ($DNA_{chr}$). They found that polynucleotide adsorption (at pH 5.0−5.5) increased steeply when the concentration of monovalent and divalent cations was ≥10 and 0.6 mM, respectively (Fig. 5.7). In the case of $DNA_{chr}$, much higher cation concentrations were required to obtain a similar effect. Likewise, Poly et al. (2000) and Cai et al. (2006) reported that DNA adsorption by clay minerals was enhanced in the presence of $CaCl_2$ and $MgCl_2$ (as compared with water). Taylor and Wilson (1979) also found that divalent cations were 100-fold more effective than monovalent cations in promoting RNA adsorption by allophane.

These observations may be explained in terms of the ability of polyvalent cation salts to "screen" the negative charge on DNA (Khan and Jönsson 1999). Further, polyvalent cations can act as a "bridge" between the phosphate groups of nucleic acids and the (negatively charged) clay surface. Indeed, cation-bridging would be

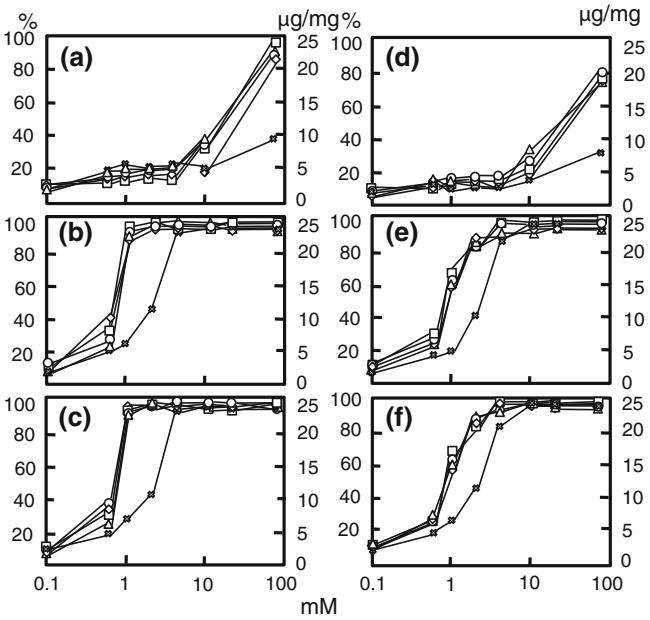



**Fig. 5.7** Adsorption of nucleic acids by montmorillonite and kaolinite in the presence of different concentrations (0.1–75 mM) of monovalent and divalent cations: **a** montmorillonite, $Na^+$; **b** montmorillonite, $Ca^{2+}$; **c** montmorillonite, $Mg^{2+}$: **d** kaolinite, $Na^+$; **e** kaolinite, $Ca^{2+}$; and **f** kaolinite, $Mg^{2+}$. *Square* poly-A; *diamond* poly-U; *triangle* poly-dA; *circle* poly-dT; and x, $DNA_{chr}$ (Franchi et al. 2003)

the dominant mode of interaction at pH > 5 when both the DNA molecule and the clay surface are negatively charged. This bonding mode also accounts for the large adsorption by $Ca^{2+}$-exchanged kaolinite and montmorillonite of linear, chromosomal DNA of high molecular weight in comparison with supercoiled plasmid DNA and rRNA of low molecular weight (Franchi et al. 1999).

Figure 5.7 shows that at comparable cation concentrations in solution, the extent of adsorption by montmorillonite was slightly higher than by kaolinite. That montmorillonite, in general, is more efficient in adsorbing DNA than kaolinite (under acid conditions) has since been corroborated by many investigators (Khanna and Stotzky 1992; Franchi et al. 1999; Pietramellara et al. 2001; Cai et al. 2006). This observation may be ascribed to the ability of montmorillonite to intercalate nucleic acids (at acid pH), whereas the interlayer space of kaolinite is inaccessible. As already mentioned, nucleic acids can also bind through ligand exchange. Although this mode of bonding is not as prominent as for small nucleotides, ligand exchange can account for the larger uptake of nucleic acids by kaolinite than montmorillonite at pH > 6 (Cai et al. 2007) because kaolinite has a relatively large edge-to-planar surface area, and intercalation by montmorillonite is restricted or absent in this pH range. On the other hand, ligand exchange appeared to be the dominant mechanism determining the interaction of nucleic acids with allophane (Taylor and Wilson 1979; Saeki et al. 2010).

The structural stability of nucleic acids intercalated into layered double hydroxides (LDH) has been examined by Thyveetil et al. (2008) using molecular dynamics simulation. Strong electrostatic interactions between LDH sheets (Fig. 5.1) and interlayer DNA (double-stranded, linear plasmid type) restricts polymer movement, and enhances its stability at elevated temperatures and pressures as compared to the free molecule (in bulk water). Their studies further suggest that hydration has a marked effect on the stability of intercalated DNA. In the anhydrous state, the Watson–Crick hydrogen bonds in the molecule are disrupted, while high hydration leads to interlayer expansion and less LDH sheet distortion. Using a similar approach, Swadling et al. (2012) found that LDH-intercalated DNA showed a greater ability to promote Watson–Crick hydrogen bonding than intercalated RNA and peptide nucleic acid (PNA), whereas the reverse trend was observed for the corresponding (free) compounds in bulk water.

## 5.7 Clay-Catalyzed Oligomerization and Polymerization of Nucleotides

The ability to clay minerals to catalyze mononucleotide dimerization was investigated by Ferris et al. (1989b) who found that adenyl-2,5′-adenosine-5-phosphate (2′,5′-$(pA)_2$), 3′,5′-$(pA)_2$, and $P_1$,$P_2$-diadenosine-5,5′-pyrophosphate (AppA) were formed when 5′-AMP was reacted with a carbodiimide (EDAC) in the presence of $Na^+$-montmorillonite. The formation of phosphodiester bonds, however, was not

observed with $Cu^{2+}$-montmorillonite or when diiminosuccinonitrile was used instead of EDAC.

Subsequent molecular dynamics simulation studies by Mathew and Luthey-Schulten (2010), using nucleotides activated by 1-methyladenine (1-Mead), also indicated that the interlayer space (of montmorillonite) was more effective in promoting dimerization than the external basal surface with the formation of 3′-5′ linkage being preferred over that of the 2′-5′ type.

Ferris and Ertem (1993) obtained oligomers (up to 11 unit long) of adenylic acid by self-condensation of ImpA (5′-phosphorimidazolide of adenosine) at pH 8 and room temperature in the presence of $Na^{+}$-montmorillonite. Again, the formation of 3′,5′ linked oligomers was preferred over the 2′,5′-linked species, while montmorillonite saturated with alkali metal ions were more efficient than the $Mg^{2+}$- and $Al^{3+}$-exchanged clays. The yield of RNA oligomers was enhanced in solution of high ionic strength (e.g., 1 M NaCl) and at pH 7–8 but decreased when the temperature was raised to 50 °C (Miyakawa et al. 2006). In this context, we might add that the montmorillonite-catalyzed formation of RNA oligomers from ImpA is but the first step toward preparing RNA with more than 40 monomers that are theoretically required for the initiation of the RNA world (Ferris 2002, 2005; Brack 2006).

Interestingly, Porter et al. (1999) found that $Cu^{2+}$-hectorite (a trioctahedral analogue of montmorillonite) could catalyze the formation of adenylic acid oligomers of up to 8 units in length by repeatedly exposing a film of the clay to a (15 mM) solution of ImpA at moderate ionic strength and pH 8. Electron-spin resonance and X-ray diffraction indicated that ImpA formed a complex with interlayer $Cu^{2+}$ ions. It would appear that the interlayer space (of hectorite) could initiate oligomerization although the operative mechanism remained unclear.

Ferris et al. (1996) were able to obtain long-chain (elongated) RNA using the "feeding" procedure; that is, by daily addition of ImpA to the decanucleotide (10-mer primer) adsorbed to $Na^{+}$-montmorillonite. Polynucleotides containing 40–50 mers were formed after 14 feedings, but only 25–30 mers were obtained using the 5′-phosphorimidazolide of uridine (ImpU) (Ferris 2002). More recently, Huang and Ferris (2006) obtained oligomers containing 40–50 mers by reacting the 5′-nucleotides of adenine, activated by 1-Mead (MeadpA), or those of uracil (MeadpU) for only one day at room temperature in the presence of montmorillonite with 61–74 % of the phosphodiester bonds being 3′,5′-linked.

## 5.8 Concluding Remarks

Bernal's (1951) hypothesis that clay minerals can act as selective adsorbents and concentrating agents of organic molecules (in the primordial Earth) has experimentally been verified. That clay minerals can also act as prebiotic catalysts for the polymerization of adsorbed organic biomonomers, however, has only partially been borne out by experiment (e.g., Theng 1974; Ponnamperuma et al. 1982). As

Brack (2006) has pointed out, the formation in water of nucleic acids (and proteins) from their monomers is not energetically favoured. These biopolymers can only form if chemical condensing agents are added or activating groups are attached to the corresponding monomers.

We have seen that montmorillonite can adsorb and intercalate individual components of RNA as well as nucleosides and nucleotides. The oligomerization of adsorbed RNA monomers, however, remains problematic in that these nucleotides need to be "activated" before they can condense into elongated polynucleotides. Even then the extent of chain elongation is often restricted (Ferris 2002). Thus, clay minerals promote polymer formation by providing a surface on which the biomonomer molecules can be concentrated and oriented in some particular configuration, rather than by initiating their polymerization *per se* as is the case with non-biological organic monomers (Theng and Walker 1970; Solomon and Hawthorne 1983).

## References

Appel PWU (1980) On the early Archaean Isua iron-formation, West Greenland. Precambrian Res 11:73–87

Banin A, Lawless JG, Massurco J, Church FM, Margulies L, Orenberg JB (1985) pH profile of the adsorption of nucleotides onto montmorillonite II. Adsorption and desorption of 5′-AMP in iron-calcium montmorillonite systems. Orig Life 15(2):89–101

Benetoli LO de B, de Santana H, Zaia CTBV, Zaia DAM (2008) Adsorption of nucleic acid bases on clays: an investigation using Langmuir and Freundlich isotherms and FT-IR spectroscopy. Monatsh Chem, 139: 753–761

Bernal JD (1951) The physical basis of life. Routledge and Kegan Paul, London

Biondi E, Branciamore S, Maurel M-C, Gallori E (2007) Montmorillonite protection of an UV-irradiated hairpin ribozyme: evolution of the RNA world in a mineral environment. BMC Evol Biol 7:1–7

Brack A (2006) Clay minerals and the origin of life. In: Bergaya F, Theng BKG, Lagaly G (eds) Handbook of clay science. Elsevier, Amsterdam, pp 379–391

Burton AS, Stern JC, Elsila JE, Glavin DP, Dworkin JP (2012) Understanding prebiotic chemistry through the analysis of extraterrestrial amino acids and nucleobases in meteorites. Chem Soc Rev 41:5459–5472

Cai P, Huang Q, Zhang X (2006) Microcalorimetric studies of the effects of $MgCl_2$ concentrations and pH on the adsorption of DNA on montmorillonite, kaolinite and goethite. Appl Clay Sci 32:147–152

Cai P, Huang Q, Lu Y, Chen W, Jiang D, Liang W (2007) Soil colloids-bound plasmid DNA: effect on transformation of *E. coli* and resistance to DNase degradation. Soil Biol Biochem 39:1007–1013

Cairns-Smith AG (1982) Genetic takeover and the mineral origins of life. Cambridge University Press, Cambridge

Cairns-Smith AG (2005) Sketches for a mineral genetic material. Elements 1:157–161

Carneiro CEA, Berndt G, de Souza Junior IG, de Souza CMD, Paesano A Jr, da Costa ACS, di Mauro E, de Santana H, Zaia CTBV, Zaia DAM (2011) Adsorption of adenine cytosine, thymine, and uracil on sulfide-modified montmorillonite: FT-IR, Mössbauer and EPR spectroscopy and X-ray diffractometry studies. Orig Life Evol Biosph 41:453–468

Cech TR (1986) A model for the RNA-catalyzed replication of RNA. Proc Natl Acad Sci USA 83:4360–4363
Churchman GJ, Gates WP, Theng BKG, Yuan G (2006) Clays and clay minerals for pollution control. In: Bergaya F, Theng BKG, Lagaly G (eds) Handbook of clay science. Elsevier, Amsterdam, pp 625–675
Cullers RL, DiMarco MJ, Lowe DR, Stone J (1993) Geochemistry of a silicified, felsic volcaniclastic suite from the early Archaean panorama formation, Pilbara block, Western Australia: an evaluation of depositional and post-depositional processes with special emphasis on the rare-earth elements. Precambrian Res 60:99–116
Ehlmann BL, Mustard JF, Murchie SL, Bibring J-P, Meunier A, Fraeman AA, Langevin Y (2011) Subsurface water and clay mineral formation during the early history of Mars. Nature 479:53–60
Ferris JP (2002) Montmorillonite catalysis of 30–50 mer oligonucleotides: laboratory demonstration of potential steps in the origin of the RNA world. Orig Life Evol Biosph 32:311–332
Ferris JP (2005) Mineral catalysis and prebiotic synthesis: montmorillonite-catalyzed formation of RNA. Elements 1:145–149
Ferris JP, Ertem G (1993) Montmorillonite catalysis of RNA oligomer formation in aqueous solution, a model for the prebiotic formation of RNA. J Am Chem Soc 115:12270–12275
Ferris JP, Hagan WJ Jr (1986) The adsorption and reaction of adenine nucleotides on montmorillonite. Orig Life 17:69–84
Ferris JP, Ertem G, Agarwal VK (1989a) The adsorption of nucleotides and polynucleotides on montmorillonite. Orig Life Evol Biosph 19:153–164
Ferris JP, Ertem G, Agarwal V (1989b) Mineral catalysis of the formation of dimers of 5'-AMP in aqueous solution: the possible role of montmorillonite clays in the prebiotic synthesis of RNA. Orig Life Evol Biosph 19:165–178
Ferris JP, Hill AR Jr, Liu R, Orgel LE (1996) Synthesis of long prebiotic oligomers on mineral surfaces. Nature 381:59–61
Franchi M, Bramanti E, Morassi Bonzi L, Luigi Orioli P, Vettori C, Gallori E (1999) Clay-nucleic acid complexes: characteristics and implications for the preservation of genetic material in primeval habitats. Orig Life Evol Biosph 29:297–315
Franchi M, Ferris JP, Gallori E (2003) Cations as mediators of the adsorption of nucleic acids on clay surfaces in prebiotic environments. Orig Life Evol Biosph 33:1–16
Galan E (2006) Genesis of clay minerals. In: Bergaya F, Theng BKG, Lagaly G (eds) Handbook of clay science. Elsevier, Amsterdam, pp 1129–1162
Gallori E, Biondi E, Branciamore S (2006) Looking for the primordial genetic honeycomb. Orig Life Evol Biosph 36:493–499
Gilbert W (1986) The RNA world. Nature 319:618
Graf G, Lagaly G (1980) Interaction of clay minerals with adenosine-5′-phosphates. Clays Clay Miner 28:12–18
Guerrier-Takada C, Altman S (1984) Catalytic activity of an RNA molecule prepared by transcription in vitro. Science 223:285–286
Haldane JBS (1929) The origin of life. Ration Annu 148:3–10
Hashizume H, Theng BKG (1999) Adsorption of DL-alanine by allophane: effect of pH and unit particle aggregation. Clay Miner 34:233–238
Hashizume H, Theng BKG (2007) Adenine, adenosine, ribose and 5′-AMP adsorption to allophane. Clays Clay Miner 55:599–605
Hashizume H, Theng BKG, Yamagishi A (2002) Adsorption and discrimination of alanine and alanyl-alanine enantiomers by allophane. Clay Miner 37:551–557
Hashizume H, van der Gaast S, Theng BKG (2010) Adsorption of adenine, cytosine, uracil, ribose, and phosphate by Mg-exchanged montmorillonite. Clay Miner 45:469–475
Hedley CB, Yuan G, Theng BKG (2007) Thermal analysis of montmorillonite modified with quaternary phosphonium and ammonium surfactants. Appl Clay Sci 35:180–188
Huang W, Ferris JP (2006) One-step regioselective synthesis of up to 50-mers of RNA oligomers by montmorillonite catalysis. J Am Chem Soc 128:8914–8919

Khan MO, Jönsson B (1999) Electrostatic correlations fold DNA. Biopolymers 49:121–125
Khanna M, Stotzky G (1992) Transformation of *Bacillus subtilis* by DNA bound on montmorillonite and effect of DNase on the transforming ability of bound DNA. Appl Environ Microbiol 58:1930–1939
Lagaly G, Ogawa M, Dékány I (2006) Clay mineral organic interactions. In: Bergaya F, Theng BKG, Lagaly G (eds) Handbook of clay science. Elsevier, Amsterdam, pp 309–377
Lagaly G, Ziesmer S (2003) Colloid chemistry of clay minerals: the coagulation of montmorillonite dispersions. Adv Colloid Interface Sci 100–102:105–128
Lailach GE, Thompson TD, Brindley GW (1968a) Adsorption of pyrimidines, purines, and nucleosides by Li-, Na-, Mg-, and Ca-montmorillonite (Clay-organic studies XII). Clays Clay Miner 16:285–293
Lailach GE, Thompson TD, Brindley GW (1968b) Absorption of pyrimidines, purines, and nucleosides by Co-, Ni-, Cu-, and Fe(III)- montmorillonite (Clay-organic studies XIII). Clays Clay Miner 16:295–301
Lawless JG, Banin A, Church FM, Mazzurco J, Huff R, Kao J, Cook A, Lowe T, Orenberg JB, Edelson E (1985) pH profile of the adsorption of nucleotides onto montmorillonite. I. Selected homoionic clays. Orig Life 15(1):77–88
Mathew DC, Luthey-Schulten Z (2010) Influence of montmorillonite on nucleotide oligomerization reactions: a molecular dynamics study. Orig Life Evol Biosph 40:303–317
Miller SL (1953) A production of amino acids under possible primitive earth conditions. Science 117:528–529
Miller SL, Urey HC (1959) Organic compound synthesis on the primitive earth. Science 130:245–251
Miyakawa S, Joshi PC, Gaffey MJ, Gonzalez-Toril E, Hyland C, Ross T, Rybij K, Ferris JP (2006) Studies in the mineral and salt-catalyzed formation of RNA oligomers. Orig Life Evol Biosph 36:343–361
Mojzsis SJ, Arrhenius G, McKeegan KD, Harrison TM, Nutman AP, Friend CRL (1996) Evidence for life on Earth before 3,800 million years ago. Nature 384:55–59
Muljadi D, Posner AM, Quirk JP (1966) The mechanism of phosphate adsorption by kaolinite, gibbsite and pseudoboehmite, Part I, the isotherms and the effect of pH on adsorption. J Soil Sci 17:212–228
Odom DG, Rao M, Lawless JG, Oró J (1979) Association of nucleotides with homoionic clays. J Mol Evol 12:365–367
Oparin AI (1924) Proiskhozhdenie zhizny. Izd. Moskovski Rabochii, Moscow
Oparin AI (1938) The origin of life. MacMillan, New York
Orgel LE (1994) The origin of life on the Earth. Sci Am, 53–61
Orgel LE (1998) The origin of life—a review of facts and speculations. Trends Biochem Sci 23:491–495
Oró J (1961) Mechanism of synthesis of adenine from hydrogen cyanide under possible primitive earth conditions. Nature 191:1193–1194
Oró J, Kimball AP (1961) Synthesis of purines under possible primitive earth conditions. I: adenine from hydrogen cyanide. Arch Biochem Biophys 94:217–227
Perezgasga L, Serrato-Díaz A, Negrón-Mendoza A, de Pablo Galán L, Mosqueira FG (2005) Sites of adsorption of adenine, uracil, and their corresponding derivatives on sodium montmorillonite. Orig Life Evol Biosph 35:91–110
Pietramellara G, Franchi M, Gallori E, Nannipieri P (2001) Effect of molecular characteristics of DNA on its adsorption and binding on homoionic montmorillonite and kaolinite. Biol Fertil Soils 33:402–409
Pizzarello S (2004) Chemical evolution and meteorites: an update. Orig Life Evol Biosph 34:25–34
Poly F, Chenu C, Simonet P, Rouiller J, Monrozier LJ (2000) Differences between linear chromosomal and supercoiled plasmid DNA in their mechanisms and extent of adsorption on clay minerals. Langmuir 16:1233–1238
Ponamperuma C, Shimoyama A, Friebele E (1982) Clay and the origin of life. Orig Life 12:9–40

Porter TL, Whitehorse R, Eastman MP, Bain ED (1999) Studies on the reaction of the 5'-phosphorimidazolide of adenosine with Cu(II)-exchanged hectorite. Appl Phys Lett 75:2674–2676

Pucci A, Branciamore S, Casarosa M, D'Aqui LP, Gallori E (2010) Implications for an RNA-clay world: Interaction of cytosine with clay minerals. J Cosmol 10:3398–3407

Rishpon J, O'Hara PJ, Lahav N, Lawless JG (1982) Interaction between ATP, metal ions, glycine, and several minerals. J Mol Evol 18:179–184

Robinson TL, Michalkova A, Gorb L, Leszczynski J (2007) Hydrogen bonding of thymine and uracil with surface of dickite: an ab initio study. J Mol Struct 844–845:48–58

Saeki K, Sakai M, Wada S-I (2010) DNA adsorption on synthetic and natural allophanes. Appl Clay Sci 50:493–497

Sanchez RA, Ferris JP, Orgel LE (1966) Cyanoacetylene in prebiotic synthesis. Sci 154:784–785

Solomon DH, Hawthorne DG (1983) Chemistry of pigments and fillers. Wiley, New York

Srivatsan SG (2004) Modeling prebiotic catalysis with nucleic acid-like polymers and its implications for the proposed RNA world. Pure Appl Chem 76:2085–2099

Strašák M, Šeršeň F (1991) An unusual reaction of adenine and adenosine on montmorillonite. Naturwissenschaften 78:121–122

Swadling JB, Coveney PV, Greenwell HC (2012) Stability of free and mineral-protected nucleic acids: implications of the RNA world. Geochim Cosmochim Acta 83:360–378

Taylor DH, Wilson AT (1979) The adsorption of yeast RNA by allophane. Clays Clay Miner 27:261–268

Theng BKG (1974) The chemistry of clay-organic reactions. Adam Hilger, London

Theng BKG (2012) Formation and properties of clay-polymer complexes, 2nd edn. Elsevier, Amsterdam

Theng BKG, Churchman GJ, Gates WP, Yuan G (2008) Organically modified clays for pollution uptake and environmental protection. In: Huang Q, Huang PM, Violante A (eds) Soil mineral-microbe-organic interactions: theories and applications. Springer, Berlin, pp 145–174

Theng BKG, Russell M, Churchman GJ, Parfitt RL (1982) Surface properties of allophane, halloysite, and imogolite. Clays Clay Miner 30:143–149

Theng BKG, Walker GF (1970) Interactions of clay minerals with organic monomers. Israel J Chem 8:417–424

Theng BKG, Yuan G (2008) Nanoparticles in the soil environment. Elements 4:395–399

Thyveetil M-A, Coveney PV, Greenwell HC, Suter JL (2008) Computer simulation study of the structural stability and materials properties of DNA-intercalated layered double hydroxides. J Am Chem Soc 130:4742–4756

Wang K-J, Ferris JP (2001) Effect of inhibitors on the montmorillontie clay-catalyzed formation of RNA: studies on the reaction pathway. Orig Life Evol Biosph 31:381–402

Weckhuysen BM, Leeman H, Schoonheydt RA (1999) Synthesis and spectroscopy of clay intercalate Cu(II) bio-monomer complexes: coordination of Cu(II) with purines and nucleotides. Phys Chem Chem Phys 1:2875–2880

Winter D, Zubay G (1995) Binding of adenine and adenine-related compounds to the clay montmorillonite and the mineral hydroxylapatite. Orig Life Evol Biosph 25:61–81

# Chapter 6
# Bottom–Up Protocell Design: Gaining Insights in the Emergence of Complex Functions

**Rafał Wieczorek, Michael C. Wamberg, Anders N. Albertsen, Philipp M. G. Löffler and Pierre-Alain Monnard**

**Abstract** All contemporary living cells are a collection of self-assembled molecular elements that by themselves are non-living but through the creation of a network exhibit the emergent properties of self-maintenance, self-reproduction, and evolution. Protocells are chemical systems that should mimic cell behavior and their emergent properties through the interactions of their components. For a functional protocell designed bottom-up, three fundamental elements are required: a compartment, a reaction network, and an information system. Even if the functions of protocell components are very simplified compared to those of modern cells, realizing a system with true inter-connection and inter-dependence of all the functions should lead to emergent properties. However, none of the currently studied systems have yet reached the threshold level necessary to be considered alive. This chapter will discuss the on-going research that aims at creating artificial cells assembled from a collection of smaller components, i.e., protocell systems from bottom-up designs.

## 6.1 Introduction

In recent years, the scientific understanding of chemistry and biology has reached a point at which the synthesis of life in the laboratory can be envisioned (Szostak et al. 2001; Rasmussen et al. 2004). Although it is surprisingly hard to agree upon a universal definition of life among scientist (Luisi 1998; Cleland and Chyba 2002), many groups of researchers have adopted various working definitions of life serving as criteria for their investigations toward creating life in the laboratory. At the FLinT Center at the University of Southern Denmark, we have adopted one

R. Wieczorek · M. C. Wamberg · A. N. Albertsen · P. M. G. Löffler · P.-A. Monnard (✉)
Center of Fundamental Living Technology (FLinT), Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, Campusvej 55 DK-5230 Odense M, Denmark
e-mail: monnard@sdu.dk

P. Pontarotti (ed.), *Evolutionary Biology: Exobiology and Evolutionary Mechanisms*,
DOI: 10.1007/978-3-642-38212-3_6, 

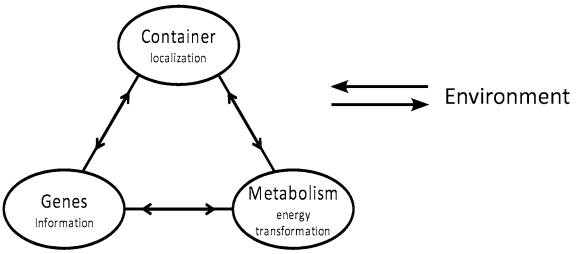



**Fig. 6.1** The view of life as a feature emerging from the mutual dependency of three components: compartment, information and metabolism. The network of those components will be influenced by and will influence the environment

such operative definition, which allowed us to envision a very concrete design of a molecular machine (Rasmussen et al. 2003).

At the foundation of the systemic protocell design is the interconnectedness between the three main components of a living entity: a compartment, a reaction network, and an information system (Fig. 6.1). If all those parts and by extension their functions are present in a chemical system, it should exhibit the emergent properties of self-maintenance, self-replication, and evolvability. The embodiment of such system could thus be considered alive.

The notion of emergence is very important for the present proposition of assembling a living system from non-living parts. Unfortunately, 'emergent properties' is a term widely used and can have different, often contradictory meanings depending on the discipline, which leads to frequent misunderstandings. Balazs and Epstein (2009) contend that the emergent properties of chemical systems are simply complexities that are not yet understood, and that greater understanding will disqualify a part of them. Other scholars differ in their understanding of emergence (Bersini et al. 2012). Luisi (Appendix in Bersini et al. 2012) states that emergence is arising from novel features due to the interactions of the components. The issue of predictability of those novel features is not crucial. Rather the fact that those properties are not intrinsic to the components is the essence of emergence.

A cell—the unit of life—is the best example of a system with emergent properties. None of the myriads of biomolecules inside a cell are alive by themselves. It is only together that they can exhibit properties like homeostasis or adaptation. A protein by itself cannot be alive in the same way as one surface of a cube cannot have a volume. Several connected surfaces are needed to make a three dimensional structure that has a volume. Analogously, we need a separated system of interacting chemicals to yield a "living" entity.

Even the simplest contemporary cell is much too complicated to envision its total synthesis. The most simplified design of an artificial cell is likely still much too complicated to realize it at first go. Therefore, researchers approach the creation of artificial life in the laboratory in a step-wise fashion: First when systems that possess some of the features of life can be made, additional features can be

incorporated while attempting to interconnect them. Such bottom-up systems, designed to be *en route* toward life, are called protocells.

A protocell mimics certain functions of cells. Even with a simplified design, some emergent properties have already been observed, arising from interactions between its various components (DeClue et al. 2009; Maurer et al. 2011; Lazzerini-Ospri et al. 2012; Mann 2012). However, the ultimate goal must be to answer the following questions: Can a simple chemical system be designed to contain within defined boundaries (container) its own set of catalytic reactions (metabolism) that could under the control of a form of stored chemical information produce its own building blocks? Would such a system be functionally equivalent to modern cells?

Protocells, besides satisfying our basic scientific curiosity, have potential applications in biotechnology. Artificial cells could be used to synthesize biotechnological products, coupling high efficiency, and level of tailoring with mass scale as well as lower production costs (Pohorille and Deamer 2002). Instead of using liposome simply as drug carriers, they could also be used as nano-factories that produce drugs inside the body regulated by cellular signals (Pasparakis et al. 2010; Stano et al. 2011). Clearly, the idea of protocells also represents a new technological concept and as such we can imagine that, once matured, it will bring about some revolutionary ways of processing (bio-) chemicals. Due to the emergent characteristics of such systems, it might even allow for novel applications that lie beyond our current vision.

## 6.2 Protocell Designs

We will now briefly review several possible ways to construct protocells, each having advantages and drawbacks depending intimately on the researcher's purposes.

### *6.2.1 Classical Enclosed Models*

The most common approaches to constructing a protocell are models based on the traditional architecture of a liposome (or vesicle depending on its building blocks) with its aqueous lumen. Reacting biochemical species are trapped inside the space bounded by the membrane of the vesicle (Luisi et al. 2006). This strategy seems to be the most natural starting point, as a cell, in a very simplified form, can be viewed as a lipid vesicle bag enclosing a set of mutually interacting molecules. Using this design, many groups have demonstrated that various cell-like processes can be transplanted and studied inside liposomes. The first attempts at this task included enzymatic production of RNA chain (poly(A)) from ADP substrate (Chakrabarti et al. 1994; Walde et al. 1994) and very soon resulted in the full PCR reaction realized inside a liposome (Oberholzer et al. 1995). Many similar works

followed involving various liposome types and various templates and different enzymes. Those early experiments involving nucleic acid polymerase enzymes have been reviewed by Monnard (2003).

The next step in this line of research was to express a protein inside a liposome compartment. Pioneering work in this direction was made by Oberholzer et al. (1999), who expressed a poly(Phe) peptide inside a liposome. The advent of molecular biology tools such as GFP-encoding plasmids allowed for a new level of investigations and the expression of a whole functional protein inside a lipid vesicle (Yu et al. 2001). Successive experiments showed gene expression within cell-sized vesicles (Nomura et al. 2003) and a cascading genetic network in which expression of one gene triggered expression of the second one (Ishikawa et al. 2004).

Within this framework, the expression of a pore protein inside the vesicle has established that it is possible to obtain a selective permeability for nutrients, which resulted in protein synthesis lasting for several days (Noireaux and Libchaber 2004). That is, the 7 identical subunits of an alpha-hemolysin pore were not only synthesized, but could also assemble spontaneously into a functional protein structure. A Japanese group (Kita et al. 2008) has shown that an expressed protein, Qbeta replicase, could catalyze the amplification of its own RNA template. Finally, Kuruma et al. (2009) proved that proteins (GPAT and LPAAT) expressed inside a liposome could carry out the final steps of phospholipid synthesis, a significant step toward true autonomy. Indeed, if one intends to synthesize living systems, not only the reproduction of the internal parts, but also the membranous container itself has to be provided. All these experiments highlight the potential of enclosed systems to carry out complex biochemistry and, in certain cases, permit the observations of higher functions through interconnection of the protocell components, e.g., the container promoting the function of the metabolism.

A model focusing more on the Origin of Life is being pursued by the Szostak group in Boston (Szostak et al. 2001). In their approach the compartments are assembled from fatty acids, which are membrane building blocks that were plausibly available on the early Earth (Monnard and Deamer 2002). Fatty acid membranes are more permeable than the usually employed phospholipid ones, enabling passive nutrient uptake from the external volume. Moreover, the genetic as well as the enzymatic components are envisioned to be both built out of RNAs instead of DNA and proteins, respectively (Szostak et al. 2001). This last premise originates from the RNA-world theory that envisions RNA as the information depository and catalytic molecules of a primitive life (Woese 1967; Crick 1968; Orgel 1968; Gilbert 1986).

The Szostak group has explored various mixtures of fatty acids in order to find one that allows for the best permeation of activated RNA monomers through the protocell membrane (Mansy et al. 2008). The authors were then able to detect efficient polymerization on an RNA primer/template system encapsulated in the vesicles (template-directed RNA polymerization, which is the basis for information replication). The group has also advanced studies on the "feeding" of the fatty acid vesicles with fatty acid micelles, which allowed for the demonstration of

vesicle growth and division (Hanczyc et al. 2003; Chen et al. 2004; Zhu and Szostak 2009).

### 6.2.2 Other Models

In recent years, several studies have brought up "unorthodox" ideas about protocell designs based on the self-assembly of inorganic chemicals. At the University of Glasgow, the Cronin group has showed that inorganic polyoxometalates can self-assemble into many unexpected supramolecular structures (Miras et al. 2012). Surprisingly, those inorganic-based materials could be engineered to self-assemble into vesicle-like structures (Pradeep et al. 2009). Thus, such structures were suggested for building an alternative protocell system, called Inorganic Chemical Cells (iCHELLs) by the authors (Cooper et al. 2011).

Another unusual concept involving inorganic membranes is pursued in the Mann group (Mann 2013). Hydrophilic spherical silica nanoparticles with diameter of 20–30 nm have been partially coated with a hydrophobic sililation agent (Li et al. 2011). Based on this system, protocells with bilayer made out of amphiphilic mineral particles could be produced to encapsulate cell-free expression system and express GFP protein. An artificial hydrogel-based cytoskeleton within these bio-inorganic membranes could even provide the protocell with a distinct physical and chemical organization in the internal environment of the protocell (Kumar et al. 2013).

### 6.2.3 Interfacial Model

The interfacial model used for the FLinT protocell has some of the features of both other strategies and living systems. It uses vesicles built from fatty acids, as main membrane constituents, and DNA as a molecule storing "catalytic information". The main unorthodox feature of this model is the fact that all reactions during the life-cycle of the protocell occur at the interface of the vesicle, i.e., the fatty acid bilayer, with the outside environment instead of being processed in its aqueous lumen.

The key processes of this protocell—energy uptake, membrane production, and information replication—all happen within or on the surface of the membrane vesicle.

## 6.3 Components of the Proposed Interfacial Protocell

The interfacial protocell is based on bilayers of fatty acid vesicle (container). A ruthenium (II) complex captures energy from visible light and carries out the production of membrane building blocks (metabolism). The information differs

from the biological one—instead of using codons coding for a function, the sequence is directly related to the phenotype (interaction with metabolism).

All of the elements were thought of as to establish a minimal life form. Complexity is reduced to a theoretical minimum that should still yield the final product that can be considered a "living system", while allowing us to assemble it in the laboratory.

### 6.3.1 Container and Reaction Networks

The container of the proposed protocell is a vesicle composed of decanoic acid (DA) molecules. Such vesicles are polydisperse in size and morphology and can be multilamellar. Their size varies from 50 nm to few μm in diameter (Cape et al. 2011). Within the membrane of the vesicles, the principal elements of the reaction system are tethered/anchored. The central element of the metabolism is the photosensitizer ($[Ru(II)(bpy)_3]^{2+}$) (Fig. 6.2a) (Maurer et al. 2011), which can be excited by visible light. Upon excitation and after reductive quenching by 8-oxoguanine (OxoG), the metal complex attains an intermediate state that can initiate an electron transfer reaction with picolyl decanoate (pL) as the acceptor (DeClue et al. 2009) (Fig. 6.2a).

The interdependency between metabolism and the compartment "body" of the protocell has been established by studying the effect the influence of co-localization of photosensitizer (Ru), reductive quencher (OxoG) in the protocell container on the production of DA. Four conformations of OxoG and ruthenium were investigated: (1) Ru and OxoG in bulk solution, (2) covalently linked Ru-OxoG, (3) covalently linked Ru-OxoG with membrane anchor, and (4) Ru and OxoG, both derivatized with membrane anchors (Fig. 6.2b). For each setup amphiphile production was measured in the presence and absence of DA vesicles (containers). It was so demonstrated that setups (2) and (4) were equally effective in the presence of the container. That is, the increase in reaction efficiency due to a covalent linkage between Ru and OxoG could be reproduced by co-localizing the two molecules on the container surface (by interaction of hydrophobic chain moieties with the membrane) (Maurer et al. 2011).

The picolinium ester, which is not an amphiphile (i.e., it will not self-assemble into vesicles), is provided as "food" from the environment, which extensively interacts with preformed bilayers. As a result, picolinium ester will only undergo reductive cleavage producing DA (L)- and a picolyl radical ($p(e^-)$)- in the protocell container (Fig. 6.2c). This property and the measured rates can therefore allow for the protocell replication by growth and division (Fig. 6.2d).

### 6.3.2 Information Component

The OxoG molecule is necessary for conversion of amphiphile precursor and its replacement by guanine leads to a drop in DA production to background level

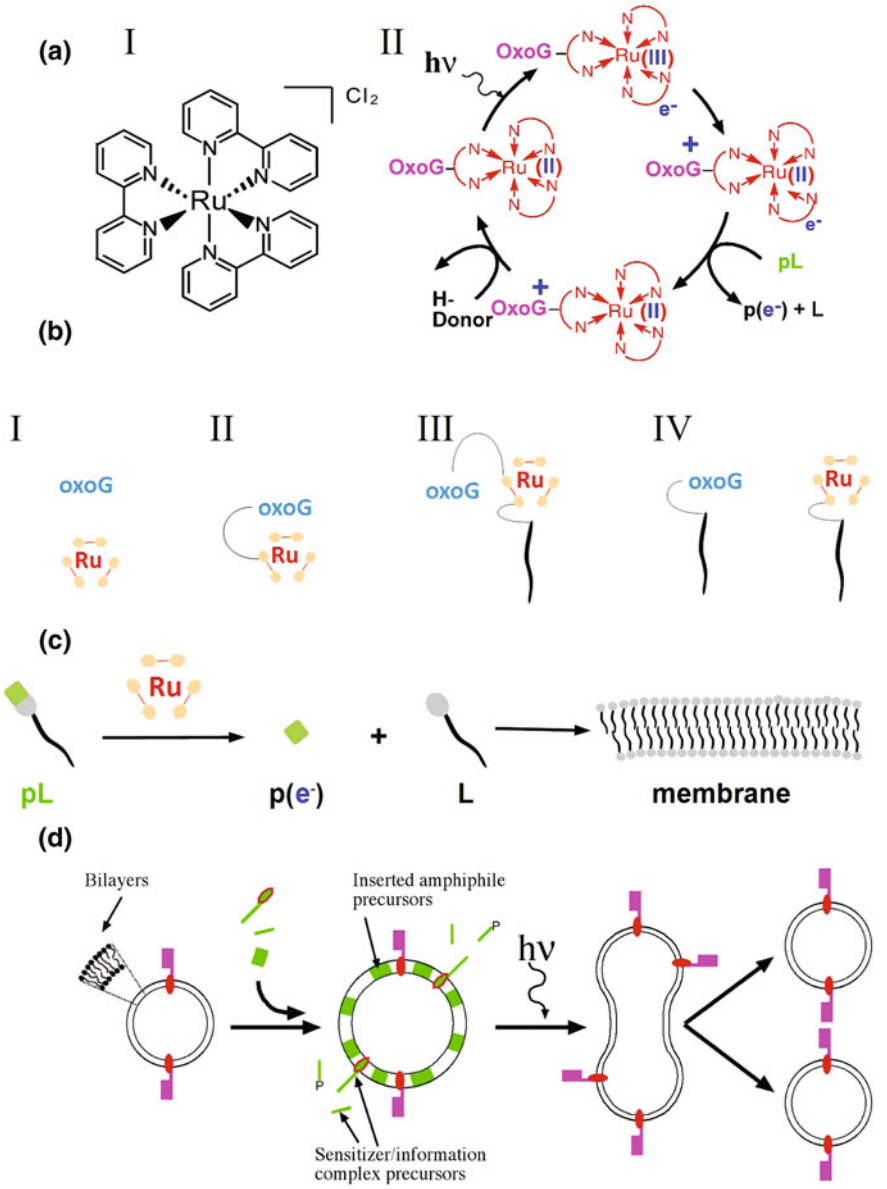



**Fig. 6.2** **a** (*I*) Chemical structure of $[Ru(II)(bpy)_3]^{2+}$ complex responsible for the harvesting of light energy, (*II*) Ruthenium complex can be excited by visible light. In the presence of oxoguanine (*oxoG*) it can initiate electron transfer reaction; **b** Four types of studied assemblies of oxoG and ruthenium complex: (*I*) water soluble intermolecular interactions (*II*) water soluble intramolecular interactions (*III*) lipid anchored intramolecular interactions (*IV*) lipid anchored intermolecular interactions; for the detailed chemical structures see (Maurer et al. 2011); **c** Precursor lipid (*pL*) is transformed by the ruthenium complex into an amphiphilic "lipid" molecule (*L*)—in our case decanoic acid, which can form membranes; **d** Light induced growth of the membrane promotes division and reproduction of the protocell

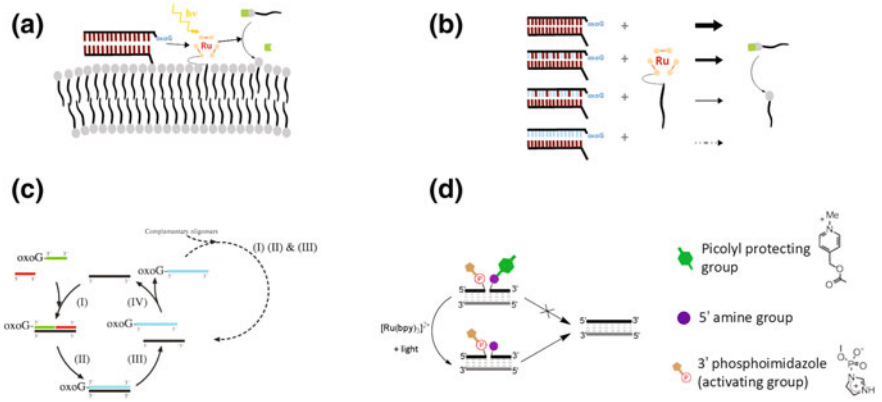



**Fig. 6.3** **a** 8-Oxoguanine present in one of the oligomers associated with the membrane of the protocell is assisting the ruthenium complex in the light-induced transformation of precursor lipid into an amphiphilic fatty acid; **b** The metabolism rates of the production of amphiphilic fatty acid depend on the sequence complementarity of the two oligomers (in *red* complementary bases and in *blue* the mismatches); **c** The two oligomers act as mutual templates. Each of them is formed from two smaller parts that are ligated on the complementary strand; **d** The mechanism of ligation involves catalysis by the light excited ruthenium complex. This process establishes the dependency of information replication with the protocell metabolism rate

(alkaline hydrolysis) (DeClue et al. 2009). Any other canonical nucleobase (A, C, T and U) will also not sustain the reaction because they are even harder to oxidize than guanine. Thus, the couple OxoG/ruthenium can be considered as a specific form of chemical information; however, which cannot be replicated by the system.

The advantage of using a nucleobase co-factor for the metabolism is obvious as OxoG can be incorporated in a DNA oligomer (8-oxoG-oligomer, either as a residue of the sequence or as a moiety attached to it) that would be replicable (Fig. 6.3a). The chemical activity of OxoG at the surface of the vesicle will depend on the level of complementarity of the 8-oxoG-oligomer to the anchor-oligomer. This results in a situation where the metabolic activity (phenotype) depends on the sequence of the information molecule (genotype) (Fig. 6.3b). Such joining of genes and metabolism although much simpler than in the extant life is in its essence performing the same function.

An additional link between those two principal functions of the protocell is the mechanism by which the genetic information is replicated. Both oligomers (8-oxoG-oligomer and anchor-oligomer) can be synthesized from two precursors provided as "food" using the ruthenium complex as catalyst, as shown in Fig. 6.3c and d. Their ligation into fully functional information molecules follows the scheme: the picolinium carbamate from the 5′-end of one fragment is cleaved as in the case of the amphiphile production and undergoes spontaneous decarboxylation, exposing the highly nucleophilic amine group. This group then performs a

nucleophilic attack at the imidazole activated 3′-phosphate group of the other DNA chain resulting in the synthesis of desired product chains (Cape et al. 2012).[1]

### 6.3.3 Expected Interactions

Not all of the designed pieces have yet been synthesized and incorporated into the experimental system. The final interactions between all of the components and the emergence, out of those interactions, of a living entity are still a proposed outcome and not a realized one. However, several aspects show validity of the undertaken path as well as the emergence of functions out of interacting pieces of our protocell have already been demonstrated, which augur well for the coming protocell construction steps. For example, the promoting effect of the container on the photochemical production of amphiphiles once the ruthenium trisbipyridine and the OxoG are anchored separately but simultaneously in the DA bilayers (Maurer et al. 2011) allow us to envision a similar outcome in the case of a DNA promoted reaction. Indeed, provided a DNA template is stably inserted into the container, this DNA should then form a double-stranded system with a complementary OxoG strand on the surface of the protocell vesicle. That is, such a construct would bring, in a sequence-specific manner, the OxoG in close proximity to an amphiphilic ruthenium complex, allowing the two photochemical reactions (amphiphile production and information replication by ligation) to occur, demonstrating one more time the interconnectedness of all the elements. The photochemical reactions are proceeding inefficiently without the container interconnection (Maurer et al. 2011).

The successful delivery of the amphiphile precursor to preformed protocells in conjunction with the observed growth of protocell upon the photo-conversion in all the functional conformations of the OxoG and ruthenium complexes also foreshadow the possibility of observing life cycles with the proposed design. However, our predictions of the behavior of certain molecular systems should interact are often proven wrong (Ludlow and Otto 2008), it is therefore important to note the elasticity of this design. Many of our initial ideas (Rasmussen et al. 2003; Rasmussen et al. 2004) had to be profoundly revised because of chemical problems or technical impracticality. It is more than likely that the current design presented in this chapter will be subjected to further modifications when subsequent experimental problems will surface.

The final goal, however, will remain to be the implementation of all of the pieces and processes simultaneously, which should result in the emergence of a synthetic life form.

---

[1] OxoG needs to be present to initiate the reaction. For example, both 8-oxo-oligomer hybridized on anchor-oligomer would need as original information component of the protocell of 1st generation while new oligomers are being formed on the surface of the vesicle.

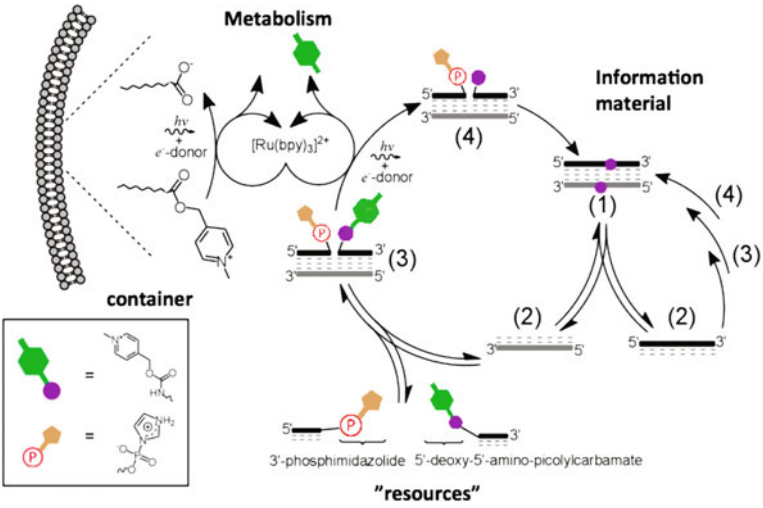



**Fig. 6.4** Scheme of the main chemical reactions of the FLinT protocell with emphasis on the replication of genetic information. Arabic numerals in brackets denotes different stages of information replication: (*1*) ligation of "resource" oligomers; (*2*) strand dissociation; (*3*) hybridization of "resource" oligomers with the templating oligomer; (*4*) ruthenium catalysed deprotection of picolinium carbamate protected "resource" oligomer

### *6.3.4 Environment and Energy Uptake*

The current overall design of FLinT protocell information replication cycle and precursor lipid metabolism is presented in Fig. 6.4 and illustrates few environmental requirements to form and maintain a functional system. The environmental conditions, such as pH, ionic strength, will significantly influence the formation of the protocell. Indeed, fatty acid bilayers are stabilized by van der Waals and hydrophobic interactions between hydrocarbon chains of DA and most importantly by hydrogen bonds between protonated/deprotonated headgroups, which means that the pH range for the container self-assembly is centered on the $pK_a$ of DA (7.2). Obviously, a high ionic strength can compromise the interactions between the headgroups.

Many of the substrates of the protocell are provided from the environment. They are absorbed and built in by the protocell without any further processing. The principal element of this pathway is the ruthenium complex ($[Ru(II)(bpy)_3]^{2+}$) which cannot be produced by the protocell. This situation is an analogue to that of vitamins that have to be ingested by humans and then used in their organisms as cofactors without any further processing. Simultaneously, the four oligomers that are needed to replicate the information protocell component are delivered as building blocks that require formation of bonds (Watson–Crick base-pairing).

Other molecules are absorbed as substrates and then broken-down into more useful chemicals: For example, the picolinium ester of DA used as a precursor of decanoic acid. An analogue of this process could be a human consumption of

starch. Humans take up starch because we can break it down into glucose, which is a useful and desired product.

Besides "food" to produce its own building blocks, the protocell needs a source of energy that drives its metabolism. This protocell is a photo-heterotrophic system as it is directly captures the energy of photons, but needs rather complex precursor molecules. One could therefore wonder whether a protocell would be able to produce complex molecules from small precursors, $CO_2$, $N_2$, or short alkanes, perhaps by using photosensitizers such as polycyclic aromatic hydrocarbons (Cape et al. 2011).

Thus, the environment will play an essential role in many aspects of the protocell assembly and functions and should be considered its fourth component.

## 6.4 Emergent Properties and Evolution

Until now modest emergent properties arising from the interactions of implemented elements have been demonstrated. For example, vesicular structures of the protocell can emerge out of the interaction of oxoguanine, $[Ru(II)(bpy)_3]^{2+}$, a cold light source, and picolinium ester of DA (DeClue et al. 2009). However, more complex behaviors can be expected while the number of interacting molecular components rises. This protocell should be able to demonstrate self-maintenance, an ability to withstand entropic decay through the self-production of its components. Self-reproduction should be a next step among the new emerging functionalities. If protocell-associated production of DA is efficient enough to exceed the natural loss of this molecule (which should be negligible small under our conditions and time-frame), the protocell will grow and led to the formation of new daughter protocells.

The final emergent property to be achieved is evolution, which might be attainable if the ability to reproduce is linked to the quality of genetic information controlling the metabolism. In Fig. 6.3b we have presented how this feature might be realized in the model of our protocell. Since the activity of ruthenium complex is dependent on the availability of OxoG whose availability in turn depends on the DNA sequence of the single protocell gene, one can hypothesize that a simple evolution principle allows for natural selection. Protocells with anchoring oligomers that have higher complementarity to the OxoG bearing oligomer than others will reproduce more efficiently, and should soon become the dominant protocell species. Thus, monitoring protocell population over time might lead to the observation of evolution in so far as individual protocells remain distinct and do not exchange their surface-bound chemicals readily.

## 6.5 Conclusions

The prospect of artificially created living systems often evokes a spectre of horror-like scenario in general public (Bedau and Parke 2009). The researchers have the

duty to inform about and explain their research as clearly as possible so that the laymen can assess whether a given research path presents any danger. The dependence of the protocell on a very well-parametrized environment should preclude, even at its final stage of construction, its "escape from the laboratory" or even its survival in any form outside of the reaction chamber.

Synthesis of life in the laboratory is no longer a domain of science-fiction authors, but a serious endeavour that modern science can and will undertake. Although many chemical and biological laboratories all over the world are performing research in this direction, few of them have proposed a well-formed plan covering all aspects of the design. In this chapter, we intended to show that a systemic protocell design can be completed and is on its way to realization. As already mentioned, it is likely that the approach described here will evolve together with the experimental progress and the final goal, if achieved, will differ in some details to those processes here described. Nevertheless, we have a clear goal and a reasonably outlined plan to achieve it. The results so far support the validity of the undertaken strategy (bottom-up systemic approach) and already show the emergence of some interesting features resulting from the interactions of various protocell components. Although, life has not been created yet, it will one day. This achievement will likely happen in the more pessimistic scenarios in the next couple of decades or in more optimistic ones much earlier. Artificial life will have a profound impact on the way humans view natural sciences and will challenge our preconceptions, thereby forcing us to reconsider the very basics of both biology and evolution.

**Acknowledgments** The authors are grateful for many fruitful discussion and suggestions from colleagues at the University of Southern Denmark and Los Alamos National Laboratory, where early stages of the systemic design described in this chapter were called Los Alamos Bug.

The work described in this chapter has been supported by the Danish National Research Foundation and by the University of Southern Denmark which jointly support the Center for Fundamental Living Technology (FLinT, Director Prof. S. Rasmussen) as well by the NASA Exobiology Program, grant # NNH08AI881 and by funding from the European Community's Seventh Framework Programme (FP7/2007–2013) under grant agreement no. 249032 (MATCHIT), and from the European Community's Seventh Framework Programme (FP7/2007–2013) under Grant Agreement no. 222422, ECCell—Electronic Chemical Cell (a project in the EU FP7-IST-FET Open Initiative).

## References

Balazs AC, Epstein IR (2009) Chemistry. Emergent or just complex? Science 325(5948): 1632–1634

Bedau MA, Parke EC (eds) (2009) The prospect of Protocells: moral and social implications of creating life in the laboratory. MIT Press, Cambridge

Bersini H, Stano P, Luisi PL, Bedau M (2012) Philosophical and scientific perspectives on emergence. Synthese 185:165–169

Cape JL, Monnard PA, Boncella JM (2011) Prebiotically relevant mixed fatty acid vesicles support anionic solute encapsulation and photochemically catalyzed trans-membrane charge transport. Chem Sci 2(4):661–667

Cape JL, Edson JB, Spencer LP, DeClue MS, Ziock HJ, Maurer SE, Rasmussen S, Monnard PA, Boncella JM (2012) Phototriggered DNA phosphoramidate ligation in a tandem 5′-amine deprotection/3′-imidazole activated phosphate coupling reaction. Bioconjug Chem 23(10): 2014–2019

Chakrabarti AC, Breaker RR, Joyce GF, Deamer DW (1994) Production of RNA by a polymerase protein encapsulated within phospholipid vesicles. J Mol Evol 39:555–559

Chen IA, Roberts RW, Szostak JW (2004) The emergence of competition between model protocells. Science 305(5689):1474–1476

Cleland CE, Chyba CF (2002) Defining life. Orig Life Evol Biosph 32(4):387–393

Cooper GJ, Kitson PJ, Winter R, Zagnoni M, Long DL, Cronin L (2011) Modular redox-active inorganic chemical cells: iCHELLs. Angew Chem Int Ed Engl 50(44):10373–10376

Crick FH (1968) The origin of the genetic code. J Mol Biol 38:367–379

DeClue MS, Monnard PA, Bailey JA, Maurer SE, Collis GE, Ziock HJ, Rasmussen S, Boncella JM (2009) Nucleobase mediated, photocatalytic vesicle formation from an ester precursor. J Am Chem Soc 131(3):931–933

Gilbert W (1986) Origin of life: the RNA world. Nature 319:618

Hanczyc MM, Fujikawa SM, Szostak JW (2003) Experimental models of primitive cellular compartments: encapsulation, growth, and division. Science 302(5645):618–622

Ishikawa K, Sato K, Shima Y, Urabe I, Yomo T (2004) Expression of a cascading genetic network within liposomes. FEBS Lett 576:387–390

Kita H, Matsuura T, Sunami T, Hosoda K, Ichihashi N, Tsukada K, Urabe I, Yomo T (2008) Replication of genetic information with self-encoded replicase in liposomes. ChemBioChem 9:2403–2410

Kumar RK, Li M, Olof SN, Patil AJ, Mann S (2013) Artificial cytoskeletal structures within enzymatically active bio-inorganic protocells. Small 9(3):357–362

Kuruma Y, Stano P, Ueda T, Luisi PL (2009) A synthetic biology approach to the construction of membrane proteins in semi-synthetic minimal cells. Biochim Biophys Acta 1788(2):567–574

Lazzerini-Ospri L, Stano P, Luisi P, Marangoni R (2012) Characterization of the emergent properties of a synthetic quasi-cellular system. BMC Bioinformatics 13(Suppl 4):S9

Li M, Green DC, Anderson JLR, Binks BP, Mann S (2011) In vitro gene expression and enzyme catalysis in bio-inorganic protocells. Chem Sci 2:1739–1745

Ludlow RF, Otto S (2008) Systems chemistry. Chem Soc Rev 37(1):101–108

Luisi PL (1998) About various definitions of life. Orig Life Evol Biosph 4–6:613–622

Luisi PL, Ferri F, Stano P (2006) Approaches to semi-synthetic minimal cells: a review. Naturwissenschaften 93:1–13

Mann S (2012) Systems of creation: the emergence of life from nonliving matter. Acc Chem Res 45(12):2131–2141

Mann S (2013) The origins of life: old problems, new chemistries. Angew Chem Int Ed Engl 52(1):155–162

Mansy SS, Schrum JP, Krishnamurthy M, Tobé S, Treco DA, Szostak JW (2008) Template-directed synthesis of a genetic polymer in a model protocell. Nature 454(7200):122–125

Maurer SE, DeClue MS, Albertsen AN, Dörr M, Kuiper DS, Ziock H, Rasmussen S, Boncella JM, Monnard PA (2011) Interactions between catalysts and amphiphilic structures and their implications for a protocell model. ChemPhysChem 12(4):828–835

Miras HN, Yan J, Long DL, Cronin L (2012) Engineering polyoxometalates with emergent properties. Chem Soc Rev 41(22):7403–7430

Monnard PA (2003) Liposome-entrapped polymerases as models for microscale/nanoscale bioreactors. J Membr Biol 191(2):87–97

Monnard PA, Deamer DW (2002) Membrane self-assembly processes: steps toward the first cellular life. Anat Rec 268(3):196–207

Noireaux V, Libchaber A (2004) A vesicle bioreactor as a step toward an artificial cell assembly. In: Proceedings of the national academy of sciences of the United States of America, vol 101(51). pp 17669–17674

Nomura SM, Tsumoto K, Hamada T, Akiyoshi K, Nakatani Y, Yoshikawa K (2003) Gene expression within cell-sized lipid vesicles. ChemBioChem 4(11):1172–1175

Oberholzer T, Albrizio M, Luisi PL (1995) Polymerase chain reaction in liposomes. Chem Biol 2:677–682

Oberholzer T, Nierhaus KH, Luisi PL (1999) Protein expression in liposomes. Biochem Biophys Res Commun 261:238–241

Orgel LE (1968) Evolution of the genetic apparatus. J Mol Biol 38:381–393

Pasparakis G, Krasnogor N, Cronin L, Davis BG, Alexander C (2010) Controlled polymer synthesis–from biomimicry towards synthetic biology. Chem Soc Rev 39(1):286–300

Pohorille A, Deamer D (2002) Artificial cells: prospects for biotechnology. Trends Biotechnol 20(3):123–128

Pradeep CP, Misdrahi MF, Li FY, Zhang J, Xu L, Long DL, Liu T, Cronin L (2009) Synthesis of modular "inorganic-organic-inorganic" polyoxometalates and their assembly into vesicles. Angew Chem Int Ed Engl 48(44):8309–8313. http://onlinelibrary.wiley.com/doi/10.1002/anie.200903070/abstract

Rasmussen S, Chen L, Nilsson M, Abe S (2003) Bridging nonliving and living matter. Artif Life 9(3):269–316

Rasmussen S, Chen L, Deamer DW, Krakauer DC, Packard NH, Stadler PF, Bedau MA (2004) Transitions from nonliving to living matter. Science 303:963–965

Stano P, Carrara P, Kuruma Y, de Souza TP, Luisi PL (2011) Compartmentalized reactions as a case of soft-matter biotechnology: synthesis of proteins and nucleic acids inside lipid vesicles. J Mater Chem 21(47):18887–18902

Szostak JW, Bartel DP, Luisi PL (2001) Synthesizing life. Nature 409(6818):387–390

Walde P, Goto A, Monnard PA, Wessicken M, Luisi PL (1994) Oparin's reactions revisited: enzymic synthesis of poly (adenylic acid) in micelles and self-reproducing vesicles. J Am Chem Soc 116(17):7541–7547

Woese CR (1967) The genetic code: The molecular basis for genetic expression. Harper & Row, New York

Yu W, Sato K, Wakabayashi M, Nakatshi T, Ko-Mitamura EP, Shima Y, Urabe I, Yomo T (2001) Synthesis of functional protein in liposome. J Biosci Bioeng 92:590–593

Zhu TF, Szostak JW (2009) Coupled growth and division of model protocell membranes. J Am Chem Soc 131(15):5705–5713

# Chapter 7
# From Life to Exolife: The Interdependence of Astrobiology and Evolutionary Biology

**Jack T. O'Malley-James and Stefanie Lutz**

**Abstract** The field of astrobiology is a relatively new discipline, which draws from a wide range of existing fields to search for life beyond Earth and to answer questions about how life originated and what its future in the Universe may be. This chapter explores the links between evolutionary biology and astrobiology and how these two disciplines can be mutually beneficial to one another.

## 7.1 Introduction

Astrobiology aims to answer the questions of how life originated on Earth, whether there is life elsewhere in our solar system and beyond, and what the future holds for life (Des Marais et al. 2008). These questions have been asked throughout much of recorded history, at least as far back in time as the Atomist philosophers of Ancient Greece (Crowe 1997) who believed there were "other worlds in other parts of the Universe, with races of different men and different animals" (Lucretius, 1 BC). It was not until the twentieth century that science had progressed to such a point that realistic attempts to take these long-held ideas and investigate them could be made. Astrobiology (originally under the title of "exobiology") has been an active field of research since the 1960s, but it was only in the last decade that the inter-disciplinary, collaborative discipline was born, becoming a scientific focus of both the European and American space programmes (Kanik and Russell 2010).

The search for life beyond Earth (with the exception, perhaps, of the debatable and contested findings of the Mars Viking lander in the 1970s) has, to date, been

J. T. O'Malley-James (✉)
School of Physics and Astronomy, University of St Andrews, North Haugh, St Andrews, Fife, Scotland KY16 9SS, UK
e-mail: jto5@st-andrews.ac.uk

S. Lutz
School of Earth and Environment, University of Leeds, Leeds LS2 9JT, UK

P. Pontarotti (ed.), *Evolutionary Biology: Exobiology and Evolutionary Mechanisms*,
DOI: 10.1007/978-3-642-38212-3_7, 

unfruitful. However, recent research suggests that life can exist in a much wider range of environments than previously thought, suggesting that, if life can adapt to such a range of conditions, then there is a chance it may have independently arisen under a wide variety of planetary environments. Such potentially habitable environments are already being found within our own solar system while the hunt for Earth-like planets around other stars is drawing ever closer to its goal.

All astrobiological theories and experiments are based on the evolutionary and biological principles gleaned from studying life on Earth. Evolutionary biology studies the evolutionary processes that have given rise to the diversity of life we see on Earth today. Life on Earth helps us define what life is, what it can tolerate and what it cannot and what features seem to be universal to all life as we know it. Moreover, it informs us of the possible sequence of events that led to the origin of life and the development of more complex organisms. This constrains astrobiological speculations, informing the next steps that are taken in the search for extraterrestrial life. Similarly, the multidisciplinary nature of astrobiology can aid research into the evolutionary laws and principles, which, to rank alongside the laws of physics for example, need to be shown to apply to everything across the Universe (Haywood 2007).

## 7.2 The Link Between Evolutionary Biology and Astrobiology

The processes that give rise to the diversity of life on Earth can tell us how life might arise beyond Earth and what that life might be like. Natural laws and processes which we observe on Earth now, have always operated on Earth and anywhere in the Universe. First formulated in the eighteenth century by the geologist James Hutton (1788), this assumption is known as uniformitarianism and has been a key principle in all fields of science allowing a close link between evolutionary biology and astrobiology, which tries to understand the origin, evolution, distribution and future of life in the Universe. Since life elsewhere in the Universe has not yet been discovered, astrobiologists need to use known evolutionary laws on Earth in their search, assuming that they apply anywhere in the Universe. These laws have been extensively studied by evolutionary biologists in order to understand the origin, evolution, distribution and future of life on Earth. The study of extant species is the key to understanding extinct life on Earth and life elsewhere in the Universe and is accompanied by a variety of questions shared by both sciences. How did life evolve? What circumstances favoured the evolution of life, the evolution of the atmosphere or the origin of water? What are evolutionary trends and what factors influence them? What allows such a diversity of life on Earth? What survival and adaptation mechanisms has life evolved? How does life survive harsh periods (impact events, global glaciation periods) and what would cause complete extinction? How can such events favour the evolution of new species?

Only the highly interdisciplinary scientific approach of evolutionary biology and astrobiology can reveal the answers to these questions. In order to understand the emergence of life and its path afterwards, biologists need to understand the prevailing framework conditions, which can be derived by consulting related sciences such as geology, geochemistry, astronomy and planetology.

Many theories in evolutionary biology can be applied to astrobiology. One such example is island biogeography which tries to understand the filters that hinder the transfer of organisms to and from isolated communities. An "island" is defined as a habitable area surrounded by unsuitable habitats. On Earth, this can either be a true island surrounded by water or, for example, a waterhole in the desert and filters could, for example be mobility potential. On a grander scale this term refers to a potentially habitable planet surrounded by an expanse of uninhabitable space which presents barriers that life would have to overcome in order to be successfully transferred between planets, such as atmospheric exit, interplanetary transfer, atmospheric entry and inoculation of the new planetary surface (Cockell et al. 2007).

Evolutionary biology and astrobiology have much in common. They are investigating very similar questions from slightly different angles and therefore, can contribute to and learn from each other's work in order to find answers to these questions.

## 7.3 How Does Life on Earth Inform the Search for Life Beyond Earth?

What we know about life on Earth has shown us the wide variety of habitats that can support life and the physical extremes that life can tolerate. This has shaped and defined where we look for extraterrestrial life and what we look for.

### *7.3.1 Habitable Environments*

A habitable environment is one in which life as we know it could survive, but is not necessarily inhabited by life. The definition of what constitutes a habitable environment has changed with time as more extreme and varied forms of life have been discovered on Earth. Dole (1964) considered a definition of habitability based on human tolerances to heat and cold, defining a planet to be habitable if at least 10 % of its surface had a mean temperature of 0–30 °C, with extremes not exceeding a minimum of −10 °C and a maximum of 40 °C (Dole 1964; Kasting et al. 1993). More recently, the ability of a planet to permit liquid water to exist on the surface was thought to be the major criterion for determining a planet's habitability, leading to the definition of the liquid water habitable zone (HZ) around a star, i.e. the region around a star where the mean surface temperature of a planet (taking account of greenhouse warming effects and the reflection of incoming

radiation by cloud and ice cover) allows water to exist in the liquid state (Kasting et al. 1993).

Today, the concept of what delineates a habitable environment is criticised for being too restrictive. A planet is considered habitable if it falls within a certain mass range (between 1 and 10 times the mass of Earth)—too massive and it would have a very dense atmosphere; too small and it would not be able to hold on to a substantial enough atmosphere. It also needs to have formed with the necessary metallicity (abundance of elements heavier than hydrogen and helium) for the chemical processes of life as we know it to take place. However, this definition neglects many environments in the solar system that could support some of the forms of life we see on Earth.

Beyond the outer boundary of the classical HZ, there are many environments that could support life below the surface. For example, Europa and Titan (icy moons of the gas giants Jupiter and Saturn respectively) lie within what has been termed the sub-surface habitability zone (SSHZ), an extended definition of the classical HZ that encompasses the temperature at different depths below the surface of a planet or moon (McMahon et al. 2013). There is evidence that Europa harbours a briny ocean beneath its 150 km thick icy crust (Carr et al. 1998; Marion et al. 2003). Therefore, while its surface would be hostile to life due to the ionising radiation within Jupiter's magnetosphere (the region around the planet encompassing its magnetic field) and the very low surface temperatures ($\sim$130 K), life may have found a foothold within the subsurface ocean (Marion et al. 2003). Similarly, Saturn's moon Titan provides an environment dominated by organic chemistry and may even resemble the pre-biotic Earth (Trainer et al. 2006). While any life on Titan would be unlikely to resemble Earth-based life, the free energy obtained from photochemical reactions on Titan could theoretically support exotic forms of life that use liquid hydrocarbons as solvents (Tokano 2009).

To complicate the situation, the HZ does not remain static with time, but moves further out from a star over time due to the increasing luminosity (or power output) of stars as they age. Hence, to host a complex biosphere, a planet would have to be positioned within the HZ over a geologically long period of time. These positions fall within what is known as the continuously habitable zone (Kasting et al. 1993).

## *7.3.2 The Limits of Life*

### 7.3.2.1 Physical Limits

Studying life on Earth, especially those that inhabit "extreme environments", has shown us the biophysical limits to life as we know it (as summarised in Table 7.1), which can be considered to define a parameter space of habitable conditions.

All these "extreme" properties have been acquired through highly adapted genes which make these species resistant to extreme temperatures, pH, pressure,

**Table 7.1** Overview of the currently known limiting factors for life on Earth

| Factor | Value | Source |
|---|---|---|
| Temperature | Maximum: 122 °C | Takai et al. (2008) |
| | Minimum: −20 °C | Canganella and Wiegel (2011) |
| Salinity | Maximum: ~30 % | DasSarma and DasSarma (2012) |
| Pressure | In nature: ~110 MPa | Yayanos (1995) |
| | Theoretically: 1,600 MPa | Sharma et al. (2002) |
| pH | Maximum: 11 | Horikoshi (1999) |
| | Minimum: <1 (*Fp. acidarmanus*) | Macalady et al. (2004) |
| Water activity (*i.e. effective water content of a solution expressed as its mole fraction*) | Minimum: 0.61 | Grant (2004) |
| Radiation | Maximum: 30,000 Gy (*Thermococcus gammatolerans*) | Jolivet et al. (2003) |
| Chemical extremes: | | Rothschild and Mancinelli (2002) |
| Gases | *C. caldarium* (pure $CO_2$) | |
| Metals | *Ferroplasma acidarmanus* (Cu, As, Cd, Zn); Ralstonia sp. $CH_{34}$ (Zn, Co, Cd, Hg, Pb) | |

salinity or irradiation. Evolutionary biologists can tell us about how these genes have evolved and have been passed on, not only vertically within a species, but also horizontally between different species, which can occur through horizontal gene transfer. Life elsewhere in the Universe might have acquired adaptations to several "extremes" and therefore might be highly adapted (in comparison to life on Earth) to the very harsh conditions in outer space and on other planets.

#### 7.3.2.2 Temporal Limits

Another limiting factor for life is time. Just as every organism has a finite lifespan and every species has a finite lifetime before it becomes extinct, habitable planets only remain habitable for a finite amount of time. This can be due to changes in environmental conditions caused by the evolution of the planet's sun (stars become more luminous as they age over Gyr (1 Gyr = $10^9$ years) timescales, which results in greater power output and therefore, higher surface temperatures on orbiting planets), but can also be due to a particular planet's size, atmospheric composition or its distance from its star.

Our own solar system hosts three examples of planetary environments that are at varying stages in their habitable lifetimes. Venus is an example of a planet that is long past the end-point of its habitable lifetime. When Venus first formed, the planet received a similar inventory of water to that of Earth and would likely have had oceans of liquid water on the surface. Temperatures on early Venus would have been mitigated by a younger, fainter Sun than today, providing surface

conditions that may have fallen within the biophysical limits for life (Nisbet and Sleep 2001). However, Venus is closer to the Sun than Earth and, as the Sun's luminosity increased over time, surface temperatures began to rise to the extent that Venus' oceans evaporated, initiating a runaway greenhouse effect that has resulted in the extremely hot planet we see today (Kasting and Catling 2003).

A second example is Mars. This is a good example of a planet that has recently (in geological terms) reached (or is possibly still in the process of reaching) the end of its habitable lifetime, not through rising temperatures, but due to the loss of its atmosphere and decreased surface temperatures. Early Mars had a thicker, carbon dioxide-rich atmosphere with higher surface temperatures than those observed today (Fairén et al. 2010). However, this original atmosphere has been significantly depleted due to sequestration in the ground or loss to space via interactions with the solar wind (Kass and Jung 1995), especially after the planet's core solidified, severely weakening its protective magnetic field. Mars is not as tectonically active as Earth and, as plate tectonics are important for recycling carbon from the crust into a planet's atmosphere, the carbon from the original atmosphere remains locked up in the crust. This vastly reduced the levels of greenhouse gases that maintain a habitable surface temperature and increased the biologically damaging ultraviolet radiation levels reaching the planet's surface. While some parts of the Martian surface still fall within the parameter space for life (Jones et al. 2011; Ulrich et al. 2012), as defined in Table 7.1, the question of whether Mars does host life of any kind remains unanswered.

Finally, Earth is not only an example of a planet that is currently habitable (and inhabited), but also provides us with the tools and information to make informed speculations about what happens to an inhabited planet as it nears the end of its habitable lifetime. For Earth, this is still a few Gyrs into the future. In approximately 1 Gyr from now, the luminosity of the Sun will have increased by 10 % from its present-day value, which increases surface temperatures and therefore, surface water evaporation rates on Earth (Goldblatt and Watson 2012). As water vapour is a greenhouse gas, this begins a positive feedback loop of increasing temperatures. When temperatures cross the threshold ($\sim$330 K) that allows water vapour (which is normally trapped in the lower atmosphere (troposphere) by an atmospheric temperature inversion) to rapidly enter the upper atmosphere (stratosphere) the planet enters what is known as a moist greenhouse state. Eventually, the planet enters a runaway greenhouse state when water vapour becomes the dominant component of the atmosphere. Once water vapour is in the upper atmosphere, it can be photo-dissociated by ultraviolet radiation into hydrogen and oxygen. As hydrogen can easily escape the planet's atmosphere, all of the planet's water inventory is lost to space in a further billion years (Caldeira and Kasting 1992).

Combining astrophysics, climate science and geology in this way give us clues about what future environments will be present on Earth. What we know about life on Earth, both at present and in the geological past, allows us to make informed speculations about what life could persist on the far-future Earth. The slowing of plate tectonics caused by the cooling planet and the loss of water (which lubricates

plate movements) slows carbon recycling. As more carbon gets trapped in the crust, atmospheric carbon dioxide levels fall. Knowing the minimum carbon dioxide concentration for plants, combined with the time taken for atmospheric carbon dioxide levels to drop this low, gives an estimate of the end of plant life—approximately 0.9 Gyr from now for higher plants (Caldeira and Kasting 1992) and 1 Gyr from now for microbial photosynthesisers (O'Malley-James et al. 2012b). The end of plant life is followed closely by the end of animal life that depends on the oxygen produced by photosynthesis, leaving behind a hot, dry microbial future Earth. Microbial life could persist within low-temperature refuges for (at most) 2.8 Gyr from present (O'Malley-James et al. 2012b), but, as temperatures continue to rise, the threshold at which biochemical reactions breakdown is crossed and the entire planet ceases to be habitable.

While the evolving Sun eventually leads to the end of life in the inner solar system, it could lead to a new beginning for life in the outer solar system. Saturn's moon Titan, with its methane lakes and abundant organic molecules, would be warmed by the ageing Sun, which could be the catalyst needed to begin a new origin of life (Lorenz et al. 1997).

### 7.3.3 Informed Speculations About Life Beyond Earth

Just as life on Earth today can inform speculations about its future, it can also provide astrobiologists with a baseline from which to extrapolate the possible attributes of life in earth-like environments beyond our solar system. Biologists on Earth study species interactions and interdependencies and determine important primary colonisers after temporary local extinction events, such as impact events and volcanic eruptions or glacial retreats (Cockell and Lee 2007; Nemergut et al. 2007). Based on this, astrobiologists can speculate about which species might be present on other planetary bodies and which niches they might occupy and use this information for their search for life in the Universe. An example of the thought processes behind such speculations is illustrated below.

### 7.3.4 An Example: Extraterrestrial Photosynthesis

The majority of life on Earth uses the Sun as its primary energy source via photosynthesis. The most abundant and accessible energy source on an Earth-like planet would likely be the radiation from that planet's host star; hence one would expect life on such a planet to evolve some form of photosynthesis to exploit this energy source (Wolstencroft and Raven 2002; Raven and Cockell 2006). Photosynthetic life on Earth uses light harvesting machinery that is honed to the radiation emission properties of the Sun, compensating for attenuation by the atmosphere (and water for aquatic organisms). As the rate of photosynthesis is

determined more by the number of photons, rather than photon energy, generally, wavelengths between 400–700 nm are used.

Different stars have different properties (temperature, mass, etc.) compared to those of the Sun and therefore, the outgoing flux of photons from any given star will peak at different wavelengths: M dwarf stars are much cooler and less massive than our Sun and have a photon flux that peaks in the infrared part of the electromagnetic spectrum (~1,000 nm), whereas O stars are much hotter and more massive than the Sun with a photon flux that peaks in the blue part of the spectrum (~70 nm). A different type of star would therefore result in different spectral properties (colours) of photosynthetic life on a life-bearing earth-like planet orbiting that star (Kiang 2007).

Such speculations can be taken a step beyond the conventional solar system-type by considering environments in which a planet has two or more suns of differing star types (O'Malley-James et al. 2012a). Scenarios like this may actually be more common than single star systems like our own solar system. More than 50 % of stars in our galaxy are found in binary or multi-star systems and planets are already being found on stable orbits within such systems, most recently the Kepler-16 (Doyle et al. 2011) and Kepler-47 systems (Orosz et al. 2012).

Two different light sources, present throughout a planet's habitable lifetime, provoke interesting evolutionary questions. For example: would photosynthetic life evolve to use one source of radiation, or would it specialise itself to exploit the most available and abundant radiation source? There are a variety of potentially stable orbits (orbits in which the distance between a planet and its host star(s) does not vary by too much, allowing a reasonably stable radiation and surface environment to persist for geological time periods) that are permissible in multi-star systems. Some result in a planet with a radiation environment dominated by one particular star in the system, whereas others result in a radiation environment in which the dominant light source changes over the course of the planet's orbit. For the latter case, photosynthetic life would be more likely to be honed to the dominant light source; however, for the former case the answer becomes more complicated.

It may be that separate forms of photosynthetic machinery develop; each adapted to a different light source, with organisms perhaps going into a dormant state for the periods of time during which their preferred light source is not present. Observations on Earth suggest that this would certainly be a successful scenario for cases where resources are abundant, as specialist organisms tend to be more successful than generalists given a large supply of resources (Ferry-Graham et al. 2002). However, there are many forms of microbial photosynthesisers that house photosynthetic machinery for different light sources (Rothschild 2010), which suggests the possibility of generalist organisms that are honed to make the best use of all the available radiation in their environment.

However, solar radiation is not the only possible energy source on Earth used by life to survive and thrive. Although life would not be as diverse without photosynthetic energy gain (Jakosky and Shock 1998), chemotrophy (the oxidation of inorganic or organic electron donors) plays an important role in habitats where

solar radiation is not available, such as in the deep subsurface biosphere [hydrothermal vent communities, for example (Fisher et al. 2007)], or potentially where surface radiation is lethal to carbon based life as is the case on Mars (Benner et al. 2000). Together, phototrophs and chemotrophs developed an incredible metabolic diversity, which allows life on Earth to colonise almost every environment.

## 7.4 How Does the Search for Extraterrestrial Life Benefit Evolutionary Biology?

Answering the questions that are important to astrobiology helps to develop a greater understanding of fundamental topics in evolutionary biology, especially questions such as where does life come from and where is it going?

### *7.4.1 Where Does Life Come From?*

The origin of life has been an enigma for humans since antiquity and several abiogenesis models have been developed including the RNA world (Gilbert 1986) and the iron-sulphur world hypotheses (Wächtershäuser 1988).

Another model which does not explain the origin of life in general, but its emergence on Earth is panspermia. First mentioned in the fifth century BC by the Greek philosopher Anaxagoras (O'Leary 2008) and revived by several modern scientists (Thompson 1871; Arrhenius 1903), it hypothesises that life exists throughout the Universe and gets propagated between planets (including Earth). This theory supposes life forms that can survive space conditions and even atmospheric entry, which has been extensively studied by astrobiologists over the last decade (de la Torre et al. 2010; Horneck et al. 2010; Olsson-Francis and Cockell 2010; Rettberg et al. 2002; Fajardo-Cavazos et al. 2005). These studies suggest that terrestrial life is able to survive space travel (if protected by an inorganic layer of substantial mass within a rocky material during travel), either in a viable or a dormant spore state until conditions allow reactivation (Horneck et al. 2001; Mileikowsky et al. 2000).

Although this theory is not meant to solve the question of how life began, findings can help evolutionary biologists get closer to answering the question of how life has emerged on Earth. In order to understand its evolution one also needs to know the conditions on the early Earth which are the focus of many studies (Lunine 2006; Jortner 2006; Shock et al. 2000). Astrobiologists use this data to search for planets with past or present conditions similar to early Earth such as Mars or Titan respectively (Jakosky 1998; Trainer et al. 2006). By combining and applying measured and modelled conditions astrobiologists try to reveal how life might have evolved under these conditions. Current and upcoming rover missions such as NASA's *Mars Science Laboratory* or ESA's *ExoMars* (planned to be

launched in 2016) and the possible outcomes of these missions in terms of life detections will immensely contribute to our understanding of life's evolution.

Even the absence of any traces of life would teach us about the limits for the emergence of life as we know it. However, the case of positive life detection, either extinct or extant, would give us a second example (separate from that on Earth) of the origin and evolution of life, finally giving us a sample number greater than one ($n > 1$) for exploring evolutionary theories.

### *7.4.2 Where Is Evolution Going?*

The most straightforward way to search for extraterrestrial life would be to go to the place of interest, or at least to send robotic probes on sample return missions. Neither of these options is currently very feasible due to the detrimental and in many cases lethal effects of outer space on the human body (Horneck and Comet 2006) and budget constraints respectively. Therefore, the remaining options left to astrobiologists are to use either remote sensing of biosignatures or to use analogue environments and species on Earth.

Studying life in (from an anthropocentric view) extreme environments on Earth can tell us about life's capabilities to adapt to these environments (Cavicchioli 2002). These capabilities can further be examined by exposing these extremophiles to simulated scenarios (e.g., irradiation, temperature, etc.) such as for the early Earth, Mars or outer space (Cockell et al. 2005; Paulino-Lima et al. 2011; Bauermeister et al. 2011; Dartnell 2011). This is also of particular interest for evolutionary biology since the increased luminosity of an ageing Sun will result in increased surface temperatures and irradiation on Earth. Environmental conditions will become more extreme and even hostile and life will have to adapt to these conditions in order to persist. Models for conditions on a far-future Earth can inform us about which route evolution would need to take in order to develop resistant species.

In most irradiation studies the viability of the exposed organisms decreases very fast and only a low percentage survive a short exposure time. However, the presence of protective material (e.g., meteorite dust) positively affects viability, which suggests a high survival potential for endolithic (rock dwelling) organisms.

Although these studies give us clues about the survival potential of life in space and on a far-future Earth, we need to take into consideration that (microbial) life will have time to slowly adapt to changing conditions on Earth. More long-term small-scale evolution studies need to be carried out to investigate this adaptation potential over several generations, such as a study by Wassmann et al. (2011), who found that a *Bacillus subtilis* strain showed increased UVC radioresistance after several generations.

### 7.4.3 What Will Ultimately Terminate Life?

Since the emergence of life on Earth there have been many catastrophic events in its history: heavy impact events, large-scale volcanic eruptions and dramatic changes in atmospheric composition and climate including the complete freezing of Earth's surface, a theory known as Snowball Earth (Hoffman and Schrag 2002). Although these catastrophic events led to global mass extinctions, life always managed to find a way to survive, which shows that events of a much greater magnitude are needed to completely sterilise a planet and terminate all microbial life. The search for life elsewhere in the Universe could give us information about the causes of terminal extinction. By finding biosignatures of extinct life elsewhere in the Universe we could make assumptions about what makes a formerly habitable planet or moon uninhabitable.

One known cause is the ageing of the Sun, which will make it impossible for life to persist on Earth after a maximum of 2.8 Gyr from the present due to temperatures exceeding life's maximum tolerance accompanied by the vanishing of liquid surface water (O'Malley-James 2012b). However, this may not be the end of life as we know it. As discussed above, there are species which could survive escape from Earth (via rockets or rocks ejected by impact events) and conditions in space. Therefore, some life could escape from the dying Earth, survive transit through space and, following this, fertilise a new habitable planet (Gladman et al. 2005).

## 7.5 Conclusion

Answering the questions, originally posed by the ancient Greek philosophers, about where life came from and how widespread it may, or may not be in the Universe is beyond the abilities of any single scientific discipline. Collaborations between biologists, astronomers, geologists, chemists and other scientists from a diverse range of fields are necessary to begin to find the answers to these questions. Such collaborations do already exist and result in new thoughts and ideas that not only help to answer the questions posed by astrobiologists, but also help to shed new light and provide new perspectives on the research within individual scientific fields. However, astrobiology is still a very young field, despite its ancient roots and more collaborations and better mutual understanding between scientific disciplines are needed to push forward our understanding of life, how it came to be and where we might find it beyond Earth.

## References

Arrhenius S (1903) Die Verbreitung des Lebens im Weltenraum. Die Umschau 7:481–485

Bauermeister A, Moeller R, Reitz G, Sommer S, Rettberg P (2011) Effect of relative humidity on *Deinococcus radiodurans*' resistance to prolonged desiccation, heat, ionizing, germicidal, and environmentally relevant UV radiation. Microb Ecol 61:715–722

Benner SA, Devine KG, Matveeva LN, Powell DH (2000) The missing organic molecules on Mars. Proc Natl Acad Sci 97:2425–2430

Caldeira C, Kasting JF (1992) The lifespan of the biosphere revisited. Nature 360:721–723

Canganella F, Wiegel J (2011) Extremophiles: from abyssal to terrestrial ecosystems and possibly beyond. Naturwissenschaften 98:253–279

Carr MH, Belton MJS, Chapman CR, Davies ME, Geissler P, Greenberg R, McEwen AS, Tufts BR, Greeley R, Sullivan R, Head JW, Pappalardo RT, Klaasen KP, Johnson TV, Kaufman J, Senske D, Moore J, Neukum J, Schubert G, Burns JA, Thomas P, Veverka J (1998) Evidence for a subsurface ocean on Europa. Nature 391:363–365

Cavicchioli R (2002) Extremophiles and the search for extraterrestrial life. Astrobiology 2:281–292

Cockell CS, Brack A, Wynn-Williams DD, Baglioni P, Brandstätter F, Demets R, Edwards HGM, Gronstal AL, Kurat G, Lee P, Osinski GR, Pearce DA, Pillinger JM, Roten C-A, Sancisi-Frey S (2007) Interplanetary transfer of photosynthesis: an experimental demonstration of a selective dispersal filter in planetary island biogeography. Astrobiology 7:1–9

Cockell CS, Lee P (2007) The biology of impact craters—a review. Biol Rev 77:279–310

Cockell CS, Schuerger AC, Billi D, Friedmann EI, Panitz C (2005) Effects of a simulated martian UV flux on the cyanobacterium, *Chroococcidiopsis* sp. 029. Astrobiology 5:127–140

Crowe M (1997) A history of the extraterrestrial life debate. Zygon 32:147–162

Dartnell LR (2011) Ionizing radiation and life. Astrobiology 11:551–582

DasSarma S, DasSarma P (2012) Halophiles. In: eLS. Wiley, Chichester

de La Torre R, Sancho LG, Horneck G, Ríos ADL, Wierzchos J, Olsson-Francis K, Cockell CS, Rettberg P, Berger T, de Vera J-PP, Ott S, Frías JM, Melendi PG, Lucas MM, Reina M, Pintado A, Demets R (2010) Survival of lichens and bacteria exposed to outer space conditions—results of the *Lithopanspermia* experiments. Icarus 208:735–748

Des Marais DJ, Nuth JA, Allamandola LJ, Boss AP, Farmer JD, Hoehler TM, Jakosky BM, Meadows VS, Pohorille A, Runnegar B, Spormann AM (2008) The NASA astrobiology roadmap. Astrobiology 8:715–730

Dole SH (1964) Habitable planets for man. Blaisdell, New York

Doyle LR, Carter JA, Fabrycky DC, Slawson RW, Howell SB, Winn JN, Orosz JA, Prsa A, Welsh WF, Quinn SN, Latham D, Torres G, Buchhave LA, Marcy GW, Fortney JJ, Shporer A, Ford EB, Lissauer JJ, Ragozzine D, Rucker M, Batalha N, Jenkins JM, Borucki WJ, Kock D, Middour CK, Hall JR, McCauliff S, Fanelli MN, Quintana EV, Holman MJ, Caldwell DA, Still M, Stefanik RP, Brown WR, Esquerdo GA, Tang S, Furesz G, Geary JC, Berlind P, Calkins ML, Short DR, Steffen JH, Sasselov D, Dunham EW, Cochran WD, Boss A, Haas MR, Buzasi D, Fischer D (2011) Kepler 16: a transiting circumbinary planet. Science 333:1602–1606

Fairén AG, Davila AF, Lim D, Bramall N, Bonaccorsi R, Zavaleta J, Uceda ER, Stoker C, Wierzchos J, Dohm JM, Amilis R, Andersen D, McKay CP (2010) Astrobiology through the ages of Mars: the study of terrestrial analogues to understand the habitability of Mars. Astrobiology 10:821–843

Fajardo-Cavazos P, Link L, Melosh HJ, Nicholson WL (2005) *Bacillus subtilis* spores on artificial meteorites survive hypervelocity atmospheric entry: implications for lithopanspermia. Astrobiology 5:726–736

Ferry-Graham LA, Bolnick DI, Wainwright PC (2002) Using functional morphology to examine the ecology and evolution of specialization. Integr Comp Biol 42:265–277

Fisher CR, Takai K, Le Bris N (2007) Hydrothermal vent ecosystems. Oceanography 20:14–23

Gilbert W (1986) Origin of life: the RNA world. Nature 319:618

Gladman B, Dones L, Levison HF, Burns JA (2005) Impact seeding and reseeding in the inner solar system. Astrobiology 5:483–496

Goldblatt C, Watson JA (2012) The runaway greenhouse: implications for future climate change, geoengineering and planetary atmospheres. Phil Trans R Soc A 370:4197–4216

Grant WD (2004) Life at low water activity. Phil Trans R Soc Lond B 359:1249–1267

Haywood S (2007) The laws of evolution and derived lawlike principles. Hagenia, Oxford, UK

Hoffman PF, Schrag DP (2002) The snowball Earth hypothesis: testing the limits of global change. Terra Nova 14:129–155
Horikoshi K (1999) Alkaliphiles: some applications of their products for biotechnology. Microbiol Mol Biol Rev 63:735–750
Horneck G, Rettberg P, Reitz G, Wehner J, Eschweiler U, Strauch K, Panitz C, Starke V, Baumstark-Khan C (2001) Protection of bacterial spores in space, a contribution to the discussion on panspermia. Orig Life Evol Biosph 31:527–547
Horneck G, Comet B (2006) General human health issues for Moon and Mars missions: results from the HUMEX study. Adv Space Res 37:100–108
Horneck G, Klaus DM, Mancinelli RL (2010) Space microbiology. Microbiol Mol Biol Rev 74:121–156
Hutton J (1788) Theory of the Earth; or, an investigation of the laws observable in the composition, dissolution and restoration of land upon the globe. R Soc Edinb 1(2):209–304
Jakosky BM, Shock EL (1998) The biological potential of Mars, the early Earth, and Europa. J Geophys Res 103:19359–19364
Jolivet E, L'Haridon S, Corre E, Forterre P, Prieur D (2003) *Thermococcus gammatolerans* sp. nov., a hyperthermophilic archaeon from a deep-sea hydrothermal vent that resists ionizing radiation. Int J Syst Evol Microbiol 53:847–851
Jones EG, Lineweaver CH, Clarke JD (2011) An extensive phase space for the potential Martian biosphere. Astrobiology 11:1017–1033
Jortner J (2006) Conditions for the emergence of life on the early Earth: summary and reflections. Phil Trans R Soc B 361:1877–1891
Kanik I, Russell MJ (2010) Integrating astrobiology research for exploration of icy bodies, astrobiology science conference 2010. Available via http://www.lpi.usra.edu/meetings/abscicon2010/pdf/5416.pdf. Cited 17 Nov 2010
Kass DM, Yung YL (1995) Loss of atmosphere from Mars due to solar wind-induced sputtering. Science 268:697–699
Kasting JF, Catling D (2003) Evolution of a habitable planet. Annu Rev Astron Astrophys 41:429–463
Kasting JF, Whitmire DP, Reynolds RT (1993) Habitable zones around main sequence stars. Icarus 101:108–128
Kiang N, Segura A, Tinetti G, Govindjee X, Blankenship RE, Cohen M, Siefert J, Crisp D, Meadows VS (2007) Spectral signatures of photosynthesis II: coevolution with other stars and the atmosphere on extrasolar worlds. Astrobiology 7:252–274
Lorenz RD, Lunine JI, McKay CP (1997) Titan under a red giant sun: a new kind of "habitable" moon. Geophys Res Lett 24:2905–2908
Lunine JI (2006) Physical conditions on the early Earth. Phil Trans R Soc B 361:1721–1731
Macalady JL, Vestling MM, Baumler D, Boekelheide N, Kaspar CW, Banfiel JF (2004) Tetraether-linked membrane monolayers in *Ferroplasma* spp a key to survival in acid. Extremophiles 8:411–419
Marion GM, Fritsen CH, Eicken H, Payne MC (2003) The search for life on Europa: limiting environmental factors, potential habitats, and Earth analogues. Astrobiology 3:785–811
McMahon S, O'Malley-James JT, Parnell J (2013) Circumstellar habitable zones for deep biospheres. 44th Lunar and Planetary Science Conference, held March 18-22, 2013 in The Woodlands, Texas. LPI Contribution No. 1719, p.2931
Mileikowsky C, Cucinotta FA, Wilson JW, Gladman B, Horneck G, Lindegren L, Meloshg J, Rickmanh H, Valtonesi M, Zheng JQ (2000) Natural transfer of viable microbes in space: 1. From Mars to Earth and Earth to Mars. Icarus 145:391–427
Nemergut DR, Anderson SP, Cleveland CC, Martin AP, Miller AE, Seimon A, Schmidt SK (2007) Microbial community succession in an unvegetated, recently deglaciated soil. Microb Ecol 53:110–122
Nisbet EG, Sleep NH (2001) The habitat and nature of early life. Nature 409:1083–1091
O'Leary MR (2008) Anaxagoras and the origin of panspermia theory. iUniverse, Bloomington, IN, USA

Olsson-Francis K, Cockell CS (2010) Experimental methods for studying microbial survival in extraterrestrial environments. J Microbiol Methods 80:1–13

O'Malley-James JT, Raven JA, Cockell CS, Greaves JS (2012a) Life and light: exotic photosynthesis in binary and multiple-star systems. Astrobiology 12:115–124

O'Malley-James JT, Greaves JS, Raven JA, Cockell CS (2012b) Swansong biospheres: refuges for life and novel microbial biospheres on terrestrial planets near the end of their habitable lifetimes. Int J Astrobiology (in press). doi:10.1017/S147355041200047X

Orosz JA, Welsh WF, Carter JA, Fabrycky DC, Cochran WD, Endl M, Ford EB, Haghighipour N, MacQueen PJ, Mazeh T, Sanchis-Ojeda R, Short DR, Torres G, Agol E, Buchhave LA, Doyle LR, Isaacson H, Lissauer JJ, Marcy GW, Shporer A, Windmiller G, Barclay T, Boss AP, Clarke BD, Fortney J, Geary JC, Holman MJ, Huber D, Jenkins JM, Kinemuchi K, Kruse E, Ragozzine D, Sasselov D, Still M, Tenenbaum P, Uddin K, Winn JN, Koch DG, Borucki WJ (2012) Kepler-47: a transiting circumbinary multiplanet system. Science 337:1511–1514

Paulino-Lima IG, Janot-Pacheco E, Galante D, Cockell C, Olsson-Francis K, Brucato JR, Baratta GA, Strazzulla G, Merrigan T, McCullough R, Mason N, Lage C (2011) Survival of *Deinococcus radiodurans* against laboratory-simulated solar wind charged particles. Astrobiology 11:875–882

Raven JA, Cockell CS (2006) Influence on photosynthesis of starlight, moonlight, planetlight and light pollution (reflections on photosynthetically active radiation in the Universe). Astrobiology 6:668–675

Rettberg P, Eschweiler U, Strauch K, Reitz G, Horneck G, Wänke H, Brack A, Barbier B (2002) Survival of microorganisms in space protected by meteorite material: results of the experiment 'EXOBIOLOGIE' of the PERSEUS mission. Adv Space Res 30:1539–1545

Rothschild LJ (2010) The evolution of photosynthesis…again? Phil Trans R Soc B 363:2787–2801

Rothschild LJ, Mancinelli RL (2002) Life in extreme environments. Nature 409:1092–1101

Sharma A, Scott JH, Cody GD, Fogel ML, Hazen RM, Hemley RJ, Huntress WT (2002) Microbial activity at gigapascal pressures. Science 295:1514–1516

Shock EL, Amend JP, Zolotov MY (2000) The early Earth vs. the origin of life. Orig Earth Moon 1:527–543

Takai T, Nakamura K, Toki T, Tsunogai U, Miyazaki M, Miyazaki J, Hirayama H, Nakagawa S, Nunoura T, Horikoshi K (2008) Cell proliferation at 122 °C and isotopically heavy $CH_4$ production by a hyperthermophilic methanogen under high-pressure cultivation. PNAS 105:10949–10951

Thompson W (1871) Presidential address to the British association. In: Popular lectures and addresses. MacMillan and Company, England, pp 132–205

Tokano T (2009) Limnological structure of Titan's hydrocarbon lakes and its astrobiological implication. Astrobiology 9:147–164

Trainer MG, Pavlov AA, DeWitt HL, Jimenez JL, McKay CP, Toon OB, Tolbert MA (2006) Organic haze on Titan and the early Earth. PNAS 103:18035–18042

Ulrich M, Wagner D, Hauber E, de Vera J-P, Schirrmeister L (2012) Habitable periglacial landscapes in Martian mid-latitudes. Icarus 219:345–357

Wächtershäuser G (1988) Before enzymes and templates: theory of surface metabolism. Microbiol Rev 52:452–484

Wassmann M, Moeller R, Reitz G, Rettberg P (2011) Growth phase-dependent UV-C resistance of *Bacillus subtilis*: data from a short-term evolution experiment. Arch Microbiol 193:823–832

Wolstencroft RD, Raven JA (2002) Photosynthesis: likelihood of occurrence and possibility of detection on Earth-like planets. Icarus 157:535–548

Yayanos AA (1995) Microbiology to 10,500 meters in the deep sea. Annu Rev Microbiol 49:777–805

# Chapter 8
# The Possible Roles of Water in the Prebiotic Chemical Evolution of DNA: An Approach by Single Molecule Studies

**Shuxun Cui**

**Abstract** There is no doubt that water is pivotal to life. Yet, as the emergence of life is still a big challenge in science, the detailed involvement of water in that process is not well recognized. Following the clues provided by recent single-molecule studies on DNA, we attempt to elucidate the possible roles of water in the prebiotic chemical evolution. Water has been long recognized as an important reactant in the Miller–Urey experiment and then as the only solvent of the primitive soup. Besides these, water also played a vital role in the prebiotic chemical evolution: water is the important criterion in the combinatorial library screening for self-assembling macromolecules. With this notion, the uniformity of biochemistry for all terrestrial life may be explained. A possible roadmap from inorganic world to the origin of life is also discussed.

## 8.1 Introduction

Since 1960s, many spacecraft have been launched to explore Mars. Generally speaking, the main mission of these ambitious, extremely high cost projects is to search a "simple" thing—water on our neighbor planet. Indeed, the viewpoint, water is essential for life, is generally accepted for many years. Yet, how important is water for life? With the great efforts by numerous scientists for many years, the answer, though still under developing, is more and more clear. Here, we would like to outline the current understanding to this question, which can be summarized into the following two aspects.

First, water is the "life's solvent". Although this is already a common sense, the following fact is still interesting: some thermophil species prefer to live in boiling

S. Cui (✉)
Key Laboratory of Advanced Technologies of Materials (Ministry of Education), Southwest Jiaotong University, Chengdu 610031, China
e-mail: cuishuxun@swjtu.edu.cn

P. Pontarotti (ed.), *Evolutionary Biology: Exobiology and Evolutionary Mechanisms*,
DOI: 10.1007/978-3-642-38212-3_8, 

water, but none of them can survive in other solvents, such as ethanol. Water may not be the unique solvent for life, but this never detracts the great contribution of water to the terrestrial life.

Second, beyond the previous level, water is an active constituent in cell biology (Ball 2008). During the past two decades, more and more scientists realized that water is deeply involved in the structures and functions of the biomacromolecules. This viewpoint is supported by numerous experimental and computer simulation results. The related research field is developing so fast that a complete review is almost impossible. Nonetheless, a few typical facts should be mentioned to exemplify the importance of water to the biological process: The bilayer lipid membrane is fabricated via the driving force of hydrophobic effect; the water molecular wire in the protein channel is a necessary condition for a proton pump; the ligand binding to the protein pocket and even the folding of the protein chain itself are often mediated by water molecules (Ball 2008; Levy and Onuchic 2006).

In the view of evolution, the extant biological systems are developed from their ancestors. Although it can be imagined that water must have influenced the prebiotic systems as well, the related study is surprisingly scarce. Here in this perspective, we attempt to elucidate the possible roles of water in the prebiotic chemical evolution.

## 8.2 The Prebiotic Chemical Evolution and the Possible Roles of Water Therein

The origin of life is a fundamental issue with long history and broad interest. However, the related study is not scientifically serious until the historic work by Charles Darwin and others in the nineteenth century (Darwin 1859). Since then, it was gradually accepted by scientists that higher level organism, including human being ourselves, are results of the natural evolution from lower level microorganisms. More chemically, on the other hand, the work by many chemists revealed that each organism cell is fabricated by various biomacromolecules, such as DNA/RNA and proteins. Then the problem arose naturally: where and how did such biomolecules come from, especially before the emergence of life?

### *8.2.1 The Prebiotic Chemical Synthesis of Monomer Building Blocks and the Possible Roles of Water Therein*

The related research on this crucial topic can be traced back to the work by Wöhler in 1828, which represented the first synthesis of an organic compound (urea) from inorganic starting materials (Wöhler 1828). Later, alanine and sugar were

synthesized in chemical lab. By the end of the nineteenth century, an extensive of research had been reported on the organic compounds using electric discharges with various gas mixtures (Rabinowitch 1945). Among others, Löb reported the synthesis of glycine by exposing wet formamide (CHO–NH2) to a silent discharge (Löb 1913). However, the work of Löb did not evoke sufficient attention, because most of the chemists at that time had no interest in the question of how life began on Earth, or in the synthesis of organic molecules under possible prebiotic conditions (Bada and Lazcano 2003; Lazcano and Bada 2003). This situation was changed in the 1950s, when Calvin et al. attempt to simulate the synthesis of organic compounds under primitive Earth conditions with high-energy radiation sources. The results showed that only formic acid was generated after the irradiation of $CO_2$ solutions with the cyclotron (Garrison et al. 1951).

In 1953, Miller published a milestone paper on the synthesis of amino acids and other organic molecules under possible prebiotic conditions (Miller 1953). Miller had applied an electric discharge to a mixture of $CH_4$, $NH_3$, $H_2$, and $H_2O$—believed at the time to be the atmospheric composition of early Earth. The "Miller-Urey experiment" to produce the "primitive soup" was corroborated by many independent groups later (Hough and Rogers 1956; Sutherland and Whitfield 1997). However, the contemporary geoscientists believe that the primitive atmosphere is more likely to be neutral (such as $N_2$ and $CO_2$), rather than the reducing one used by Miller in 1950s (Tang 2007). The importance of the "primitive soup" to the emergence of life was doubted for that when the gas mixture becomes less reducing (i.e., less $H_2$, $CH_4$, or $NH_3$), the yields of organic compounds decrease drastically. It was proposed that the meteorite delivery may be a more important source of the biological relevant organic compounds. Yet, all the amino acids detected in the Murchison meteorite had been already found in the primitive soup, suggesting that the organic compounds were synthesized by reactions similar to those in the Miller experiment (Kvenvolden et al. 1970; Miller 1953; Ring et al. 1972).

Some recent studies supported the theory of primitive soup. In 2005, one theoretical study suggested that the early atmosphere could contain up to 30 % of reducing component (hydrogen) (Tian et al. 2005). In 2008, the Bada–Miller group successfully synthesized significant amounts of amino acids from neutral gas mixtures with the treatment of electric discharge (Cleaves et al. 2008). They also found the reason of very low yield in the previous reports using neutral gas mixtures: the produced organic compounds can be oxidized and hydrolyzed by nitrate and nitrite, which are also the production in the reactions. Thus, the yield of amino acids is greatly increased when oxidation inhibitors, such as ferrous iron, are added prior to hydrolysis. Experimental results showed that the addition of traces of $CH_4$ and/or $H_2$ to the neutral atmosphere would further enhance the production of amino acids (Schlesinger and Miller 1983).

The contributions of water in the abiotic synthesis of monomer building blocks are straightforward: First, water is one of the important reactants in the Miller–Urey experiment. Second, water is the good solvent for the resultants and many kinds of salts as well. The primitive soup could be a mixture of numerous

compounds. Many cations are catalyst for the possible reactions in the soup. The fact that water is a potent solvent indeed facilitates further evolution of the soup.

Although the primitive soup theory is still under debate, (Catling 2006; Yockey 1992, 1995a, b, 1997, 2000, 2005) there is little doubt that the "Miller-Urey experiment" is a reliable route from inorganic world to organic world. It is now time to envisage, again, the link from the "Miller-Urey experiment" to the origin of life. So far, a large assortment of biologically relevant organic compounds can be synthesized using a variety of gas mixtures and energy sources (Miller 1998; Robertson and Miller 1995; Sutherland and Whitfield 1997). Which direction should this "primitive soup" move in the next step of evolution?

### *8.2.2 Prebiotic Chemical Evolution Toward Higher Level Ordered Structures and the Possible Roles of Water Therein*

Even the most primitive organism should be fabricated by some biomacromolecules. Yet, all these organic molecules in the primitive soup can be considered to be only the monomers of the biomacromolecules. It is already shown that clays, metal cations, and some organic compounds may have catalyzed polymerization reactions between the monomers (Bujdak and Rode 1999; Ertem and Ferris 1997; Wills and Bada 2000). By the surface-bound template polymerization, nucleotides of more than 50 repeating units can be synthesized (Ferris et al. 1996; Ferris 2002). Here, let's skip all the kinetics and thermodynamics of the possible reactions and presume that the copolymerization between all the reactive organic compounds in the soup are possible. Then, there should have numerous species of macromolecules in the soup. Which species have more potential to contribute more to the primitive organism (Eschenmoser 1999)?

Organism is not merely a cluster of molecules. The information embedded in an organism is much larger than that of a cluster of common inorganic molecules. The richness of information is a result of self-recognition or more generally, self-organization (Piqueira and Roberto 2008; Prokopenko et al. 2009). The propagation or the self-replication of the organism itself certainly needs precise supramolecular processes, which is necessary to pass on gene or other unambiguous biological information. It should be noted here that the precise self-organization is not an inherent property of all macromolecules. One fact is that most of the man–made polymers only form random coils in their aqueous solutions (Zhang and Zhang 2003). The ability of precise self-organization is one important divide between the nonbiological macromolecules and the biological ones. Then, which kinds of macromolecules can form precise supramolecular structures in their aqueous solutions?

Since the symbol event of Nobel Prize in 1987, supramolecular chemistry has been developed for more than 20 years (Lehn 1988). Plenty of wonderful ordered supramolecular structures have been fabricated via weak intermolecular interactions, such as H-bonding, coordinate bond, and pi–pi stacking (Lawrence et al.

1995; Lehn 1990, 1993, 1995). Among others, H-bonding, especially multiple H-bonding, is highly specific and thus is most useful in fabricating a precise supramolecular structure. To date, however, most of the processes of artificial precise self-organization have been taken place in organic solvents with relative low polarity (Lawrence et al. 1995). Highly polar solvents, including water, have been shown to be strong disturber to the H-bonding between the assembling moieties, which often finally destroy the precise self-assemblies (Zou et al. 2005).

It is worth noting that water is the most abundant component in the "primitive soup". Water is also the main component inside the active cell of an extant organism, which is reasonably applicable for the primitive organism. The possible primitive organism has to face this situation: water is not avoidable. The water environment is a barrier that has to be overcome by the prebiotic system, on the route to the primitive organism.

The "primitive soup", which has numerous possibilities in evolution, is actually a library in the viewpoint of combinatorial chemistry (Landweber 1999). Here, water can be regarded as the screening criterion of the combinatorial library: only those macromolecules that can form precise supramolecular structures in aqueous solutions will dominate in the competition. The self-organization of a random coiled macromolecule into a specific stable structure will effectively prevent the chemical species from hydrolysis or other decomposition effect. It has been shown that most of the organic compounds would be hydrolyzed or oxidized into abio-molecules in the aqueous solution, when oxidation inhibitors are absent (Cleaves et al. 2008).

Alternatively, the prebiotic library screening can be regarded as a Darwinian fashion natural selection (Landweber 1999). Water is the surrounding environment of the prebiotic chemical system. To survive, the prebiotic system must adapt the aqueous environment. The H-bonding directed self-organization in aqueous solution is rather difficult, but still possible. It is expected that after a very long time natural selection by the water environment, some species of macromolecules that can form precise supramolecular structures were selected, which entered the next stage for the further evolution.

If the above speculation is truly happened in the history, the extant biomacromolecules should also have the feature of adaptability to the water environment (Conrad 1977). This environmental adaptability implies that the primary molecular structures of the biomacromolecules are specially designed (or selected) to fit the aqueous environment and the potential of the precise self-organization of the biomacromolecules are not dissipated by water.

## 8.3 Quantitative Tests for the Adaptability of DNA to the Aqueous Environment

In spite of the vast diversity in appearance, all the terrestrial organisms employ the same RNA/DNA system as gene carrier. Compared with RNA, the higher level structures of DNA are more uniform for all organisms. This feature entitles DNA the

most appropriate object in a pilot study to test the water environment adaptability, which may be also helpful in understanding the chemical etiology of DNA (Beier et al. 1999; Eschenmoser 1999; Schoening et al. 2000). Recently, by utilizing single-molecule force spectroscopy, (Hugel and Seitz 2001; Janshoff et al. 2000; Zhang and Zhang 2003). Cui et al. carried out series investigations on the effects of the aqueous environment on the structures and functions of DNA, which present supportive evidences for the hypothesis (Cui et al. 2006, 2007, 2009; Cui 2009, 2010).

### 8.3.1 The Effects of Environment on the Formation of the DNA Duplex

The self-organization from single-stranded DNA (ssDNA) to double-stranded DNA (dsDNA) is usually formularized below, where ssDNA′ denotes the complimentary chain of ssDNA

$$\mathrm{ssDNA} + \mathrm{ssDNA}' \rightleftharpoons \mathrm{dsDNA} \tag{8.1}$$

It has long been known that both ssDNA and dsDNA are hydrated in the aqueous solution (Ball 2008; Brovchenko et al. 2008). It is also clear that dsDNA has less binding sites with water than that of the sum of the two free ssDNA chains. Therefore, a partial dehydration process should occur prior to the self-organization of ssDNA. Thus, Eq. 8.1 should be revised into a more rigorous form as follows:

$$\mathrm{ssDNA}\cdot \mathrm{xH_2O} + \mathrm{ssDNA}'\cdot \mathrm{xH_2O} \rightleftharpoons \mathrm{dsDNA}\cdot \mathrm{yH_2O} + (2\mathrm{x} - \mathrm{y})\mathrm{H_2O} \tag{8.2}$$

Whether the supramolecular self-organization described in Eq. 8.2 can occur is hinged on the free energy change ($\Delta G_2$) of the process. A typical value measured by differential scanning calorimetry is −4.3 kJ/(mol bp) (Breslauer et al. 1986). Like other complex reactions, Eq. 8.2 can be separated into the following two simpler steps:

$$\mathrm{ssDNA}\cdot \mathrm{xH_2O} \rightleftharpoons \mathrm{ssDNA}\cdot \mathrm{y}/2\mathrm{H_2O} + (\mathrm{x} - \mathrm{y}/2)\mathrm{H_2O} \tag{8.3}$$

$$\mathrm{ssDNA}\cdot \mathrm{y}/2\mathrm{H_2O} + \mathrm{ssDNA}'\cdot \mathrm{y}/2\mathrm{H_2O} \rightleftharpoons \mathrm{dsDNA}\cdot \mathrm{yH_2O} \tag{8.4}$$

$$\Delta \mathrm{G_2} = 2 \times \Delta \mathrm{G_3} + \Delta \mathrm{G_4} \tag{8.5}$$

All the water rearrangement is completed in Eq. 8.3, whereas all the assembly between ssDNA chains occurs in Eq. 8.4. It is reasonable to assume that the partial dehydration (Eq. 8.3) is a nonspontaneous process (i.e., $\Delta G_3 > 0$), whereas the self-organization of the two ssDNA chains (Eq. 8.4) is a spontaneous one (i.e., $\Delta G_4 < 0$). Given that the total process in Eq. 8.2 is a spontaneous one, the Gibbs free energy necessary to partially remove the hydration shell of the ssDNA

(Eq. 8.3) must be compensated (or overcompensated) by the Gibbs free energy of the base pairing (Eq. 8.4). However, $\Delta G_3$ (or $\Delta G_4$) is neither a ready data in references, nor a value can be measured easily by traditional ensemble methods.

Recently, Cui et al. have simulated the process of the partial dehydration (Eq. 8.3) by manipulating an individual ssDNA chain in aqueous solutions and organic nonsolvents, respectively (Cui et al. 2006, 2009). The interactions between the nonpolar organic solvent molecules and the solute molecules are van der Waals interactions in general, which are the weakest intermolecular interactions. In this condition, it is to be expected that the solute molecules' behavior is close to that in the vacuum condition. Therefore, we were able to obtain the inherent elasticity of a single ssDNA chain by stretching it in organic nonsolvents. The experimental inherent elasticity is in good agreement with the theoretical value that obtained from *ab-initio* calculations (Cui et al. 2006, 2009; Hugel et al. 2005). Similarly, the elasticity of ssDNA was obtained in aqueous environments. For both cases, single ssDNA chain was elongated under different stretching velocity, but no evident difference was observed. This finding is consistent with the fact that the lifetime of the H-bonds between ssDNA and water as well the thermal fluctuation of the ssDNA chain is much shorter than the time scale in force measurement (0.2–1 ms), which implies that the single molecule experiments were carried out close to the thermodynamic equilibrium (Cui et al. 2006). A comparison between the mechanical behaviors of ssDNA in the two environments reveals a remarkable difference, see Fig. 8.1.

There are many functional groups containing H-bonding acceptors and donors in each ssDNA chain. SsDNA is soluble in water, and can form H-bonds with water molecules in various combinations. It is possible that one water molecule can form two H-bonds with one ssDNA molecule and create a water "bridge" around the ssDNA chain (Liu et al. 2005; Oesterhelt et al. 1999). To meet the length requirement of the water "bridge", the ssDNA chain may be shortened in some extent. Upon stretching by external force, the ssDNA chain is lengthened and

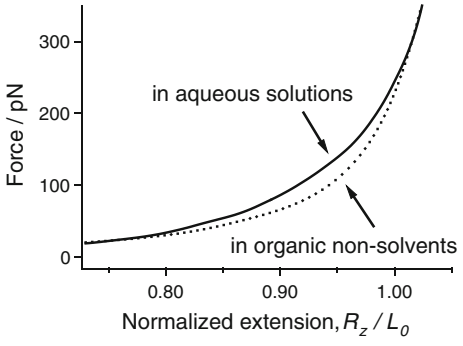



**Fig. 8.1** The comparison of normalized force-extension curves of ssDNA obtained in different environments, i.e., aqueous solutions (*upper curve*) and organic nonsolvents (*lower curve*). The experimental force curves are smoothed and zoomed to show a clearer deviation between them. Reproduced from Cui (2010) with permission

the length requirement no longer meets, which will break the water "bridges" consequently. The rearrangement of water molecules around the ssDNA chain will consume additional energy besides those contributions stored in the "pure" elastic behavior. In this way, the elongation of ssDNA in water will consume more energy (0.58 $k_BT$/base, or 1.4 kJ/mol base) than that consumed in organic solvents, as reflected by the deviation between the force curves shown in Fig. 8.1.

The rearrangement of the water molecules around the ssDNA chain upon stretching can be regarded as partial dehydration, see Eq. 8.6.

$$\mathrm{ssDNA{\cdot}xH_2O} \rightleftharpoons \mathrm{ssDNA \cdot zH_2O + (x - z)H_2O} \tag{8.6}$$

$\Delta G_6 = 1.4$ kJ/mol base

It can be observed that Eqs. 8.6 and 8.3 are very similar. Although the values of $z$ and y/2 are not necessarily to be identical, the free energy changes of the two equations are expected to be very close, i.e., $\Delta G_3 \approx \Delta G_6$.

Then by Eq. 8.5, we can estimate that $\Delta G_4 \approx -7.1$ kJ/mol base. Note that the value of $\Delta G_4$ is not far from zero, making the reaction in Eq. 8.2 rather sensitive to $\Delta G_3$. Comparing with other systems having water rearrangement upon stretching, it is found that ssDNA costs the lowest energy among them (Cui et al. 2006; Liu et al. 2005; Oesterhelt et al. 1999). The rather small value of $\Delta G_3$ may be crucial to life: If $\Delta G_3$ is a much larger value like the poly(N-vinyl-2-pyrrolidone)/water system (13.0 kJ/mol unit) (Liu et al. 2005) or the poly(ethylene–glycol)/water system (7.2 kJ/mol unit), (Oesterhelt et al. 1999) $\Delta G_2$ would be larger than zero, and the self-organization from ssDNA to dsDNA will not be a favorable process. We may conclude that it is the weak disturbance of water molecules on ssDNA that makes the self-organization possible at physiological condition.

### 8.3.2 *The Effects of Environment on the Dissociation of the Duplex*

To read the gene information encoded in dsDNA, the strands of the double helix need to be separated. There are several traditional methods to dissociate the duplex in vitro, including thermal-induced melting and reagent denaturation. However, it is expected that a different mechanism should be exploited to dissociate the duplex in vivo (Lohman and Bjornson 1996).

Previous extensive studies showed that in aqueous environment, the DNA duplex was maintained by several weak intermolecular interactions, including hydrogen bond, pi–pi stacking, van der Waals interactions, and hydrophobic force (Bloomfield et al. 2000). Differ from the covalent bond, these weak intermolecular interactions are known to be environment dependent, (Fu et al. 2002; Sukhishvili and Granick 2000; Zou et al. 2005) see Table 8.1. It is expected that if a different environment, e.g., a nonpolar one, is provided for dsDNA, the total binding force within the duplex will be weakened. Previous ensemble measurements showed that

**Table 8.1** Intermolecular interactions involved in the dsDNA assembly

| | In aqueous solutions | In nonpolar organic solvents |
|---|---|---|
| H-bonding | √ | √ |
| pi–pi stacking | √ | √ |
| vdW | √ | √ |
| solvophobic effect | √ | × |

(Bloomfield et al. 2000) Reproduced from Cui (2010) with permission

nonpolar organic solvents are poor solvents for dsDNA, where dsDNA condensated or even precipitated (Feng et al. 1999; Montesi et al. 2004; Pereira and Williams 2001). Other detailed information of DNA in poor solvents is still unclear. The ensemble measurement methods encountered limitations here.

Recent single-molecular studies by Cui et al. showed that when DNA is dragged into a nonpolar organic solvent, the mechanical behavior of a dsDNA molecule is undistinguishable to that of a ssDNA molecule, see Fig. 8.2 (Cui et al. 2007). Since it was already known that the single-molecular mechanical behavior can be used as a fingerprint to identify the stranding status of DNA, (Clausen-Schaumann et al. 2000; Rief et al. 1999; Smith et al. 1996) the above result implies that dsDNA is denatured in a nonpolar organic solvent. Control experiments indicated that different nonpolar organic solvents have the same effect on dsDNA. Possible influences from the substrate were also excluded, for that similar results could be obtained from different type of substrates. These results strongly suggest that the nonpolar organic solvents have similar effects that denature dsDNA into ssDNA. This hypothesis was supported by molecular dynamics (MD) simulations, which showed that dsDNA underwent a strand separation at the interface of the two kinds of liquids (Cui et al. 2007).

The solvents can be classified into two types, according to the stranding status of DNA, see Table 8.2 (Bonner and Klibanov 2000; Cui et al. 2007). It can be seen that the dsDNA structure is only maintained in water and water-like solvents, which have multi-hydroxyl groups. Other solvents, despite the large difference in polarity, are denaturant for dsDNA actually. Note that for most time, DNA exists in its dsDNA status. When needed, dsDNA is unwound into ssDNA by enzymes. Another fact is that water acts as the usual environment for DNA, while the

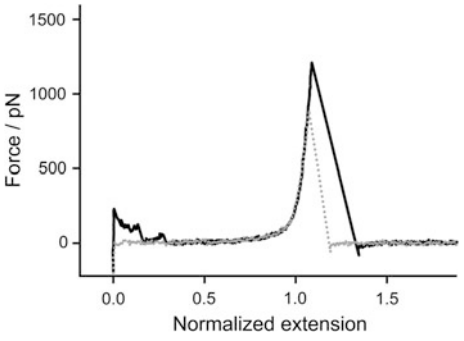



**Fig. 8.2** Normalized force curves of denatured dsDNA (*dotted line*) and ssDNA (*solid line*) both obtained on amino-modified substrate in 1-propanol. Reproduced from Cui (2010) with permission

**Table 8.2** The stranding status of DNA and the corresponding liquid environments

| Stranding status of DNA | Mono-solvent liquid environment |
|---|---|
| Double stranded | Water, glycerol, and ethylene glycol |
| Single stranded | DEBenzene, octane, 1-propanol, DMSO, and formamide |

(Bonner and Klibanov 2000; Cui et al. 2007) Reproduced from Cui (2010) with permission

enzymes seem to provide an unusual environment. Those local environments, which are helpful to remove the water shell around DNA, are expected to destabilize the supramolecular structure and unwind the duplex at the binding position. If this mechanism is truly happening in vivo, the phenomenon of the diversity of helicases may also be explained, (Lohman and Bjornson 1996) since various micro environments, ranging broadly from very apolar to very polar (see Table 8.2), can lead to strand separation.

### *8.3.3 The Interpretation of Single-Molecule Results of DNA in the Framework of Aqueous Environment Adaptability*

As discussed in the previous sections, the aqueous environment plays a crucial role in the regulation of supramolecular structure of DNA, see Fig. 8.3. On one hand, the small energy cost of the water rearrangement upon the duplex formation ensures the stability of dsDNA in the aqueous environment. On the other hand, by shaving off the water molecules binding to the duplex, dsDNA is destabilized and tends to be unwound. Water is known to be a special solvent: it usually interferes the hydrogen-bonding directed self-organization; at the same time, liquid water provides hydrophobic forces that often drive the self-organization of organic molecules (Ball 2008; Chandler 2005; Lawrence et al. 1995; Lum et al. 1999; Zhang et al. 2003). There are many kinds of water soluble macromolecules, but few can form stable dsDNA structure. This fact implies that DNA is somewhat special in the molecular structure. To achieve a precise self-assembly, there should be a subtle balance for the structure design of the biomacromolecules that utilizes

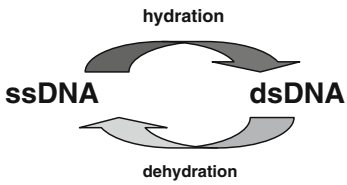



**Fig. 8.3** The hydration/dehydration processes regulate the supramolecular structures of DNA. Reproduced from Cui (2010) with permission

the hydrophobic force from water. The H-bonding sites are located in the hydrophobic core of the helix, where disturbance from water is screened. The weak intermolecular interactions involved in the duplex formation seem to be handled carefully by a delicate "design" of the primary structure of DNA, so that the self-organization between the complementary ssDNA chains can be reversible under a facile regulation (see Fig. 8.3). One can say that the aqueous environment is crucial to the proper structure and function of DNA. In the other way around, we believe, it is more likely that the primary structure of DNA has been selected via the prebiotic chemical evolution from the ancestor molecules in a Darwinian fashion (Cairns-Smith 1975; Eigen 1971; Landweber 1999; Lauterbur 2008; Leemhuis et al. 2009; Lehn 2002, 2007) to adapt to the aqueous environment of our planet. After all, the aqueous environment is an elementary condition of our planet and exists prior to any biological systems.

## 8.4 Concluding Remarks

As implied by the pilot studies, DNA could be a final result of long time prebiotic natural selection. The primary molecular structure of DNA was selected, because such structure enabled the further precise self-organization in the aqueous environment. If this really occurred in the prebiotic evolution, the possible route from inorganic world to the emergence of life can be conceived as described below.

According to the corresponding complexity, we attempt to fractionize the whole possible process of the emergence of life into six levels, as shown in Fig. 8.4. The first level is the inorganic world of the primitive Earth, in which the chemical elements and energy needed are ready for the further evolution. Then, by the Miller experiment and other possible contribution, such as meteorite delivery, the inorganic compounds can be converted into the organic monomers (Level 2). From here, various macromolecules (Level 3) can be generated by polymerization reactions catalyzed by clays, etc. Note that these reactions (from Level 1 to Level 3) are in a

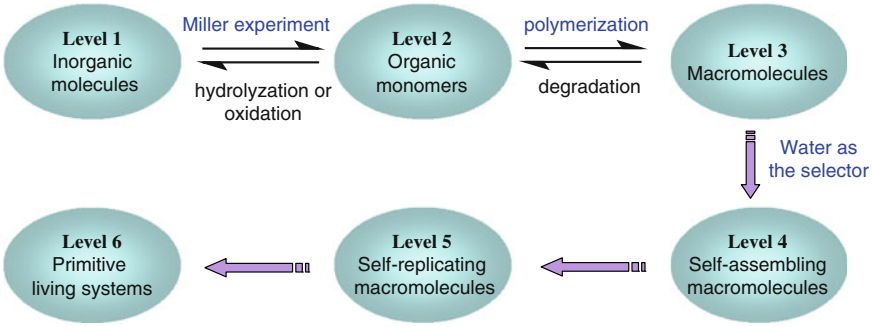



**Fig. 8.4** A possible roadmap from inorganic molecules to the primitive living systems. Reproduced from Cui (2010) with permission

dynamic equilibrium: The macromolecules can be degraded into monomers, and the monomers can be hydrolyzed or oxidized into inorganic compounds. If a certain kind of macromolecule can form stable supramolecular structures (Level 4) in the aqueous solution, the equilibrium would be moved toward higher levels and such kind of macromolecule would survive in the prebiotic chemical evolution.

In the further natural selection, it is expected that the self-replicating macromolecules (Level 5) will be the winner in the competition simply because of preponderance from self-replication. The RNA world may be one typical system in the Level 5 status (Bielski and Tencer 2007; Gilbert 1986). From this level, the gap between the prebiotic system and the primitive life is much narrower, which greatly enhanced the possibility for the emergence of life.

Water has been long recognized as an important reactant in the Miller–Urey experiment and then as the only solvent of the primitive soup. Besides these, water also played a vital role in the prebiotic chemical evolution: Water is the important criterion in the combinatorial library screening for self-assembling macromolecules. There could be numerous species in the Level 3 stage in the prebiotic systems. However, most of the macromolecule species were screened out by the water environment in the process to Level 4. With this notion, the uniformity of biochemistry for all terrestrial life may be explained.

Water is the primary environment of terrestrial life, which supports life and at the same time, restricts the form of life. Therefore, it is water that defines and shapes the life. With many pending problems, the origin of life is still a riddle (Tang 2007; Wächtershäuser 2000; Yockey 2000). Yet, it is exciting to find some clues implying that the Darwin's theory is also effective in the prebiotic screening of primary molecular structures of the biomacromolecules. We believe that this notion will cast new light on the origin of life in the future.

## References

Bada JL, Lazcano A (2003) Prebiotic Soup–revisiting the miller experiment. Science 300:745

Ball P (2008) Water as an active constituent in cell biology. Chem Rev 108:74–108

Beier M, Reck F, Wagner T, Krishnamurthy R, Eschenmoser A (1999) Chemical etiology of nucleic acid structure: comparing Pentopyranosyl-(2′-4′) oligonucleotides with RNA. Science 283:699–703

Bielski R, Tencer M (2007) A possible path to the rna world: enantioselective and diastereoselective purification of ribose. Orig Life Evol Biosph 37:167–175

Bloomfield VA, Crothers DM, Tinoco I (2000) Nucleic acids: structures, properties, and functions. University Science Books, Sausalito

Bonner G, Klibanov AM (2000) Structural stability of DNA in nonaqueous solvents. Biotech Bioeng 68:339–344

Breslauer KJ, Frank R, Bloecker H, Marky LA (1986) Predicting DNA duplex stability from the base sequence. Proc Natl Acad Sci USA 83:3746–3750

Brovchenko I, Krukau A, Oleinikova A, Mazur AK (2008) Ion dynamics and water percolation effects in DNA polymorphism. J Am Chem Soc 130:121–131

Bujdak J, Rode BM (1999) The effect of clay structure on peptide bond formation catalysis. J Mol Cataly A 144:129–136

Cairns-Smith AG (1975) Ambiguity in the interpretation of abiotic synthesis. Orig Life 6:265–267
Catling DC (2006) Comment on "A hydrogen-rich early Earth atmosphere". Science 311:38
Chandler D (2005) Interfaces and the driving force of hydrophobic assembly. Nature 437:640–647
Clausen-Schaumann H, Rief M, Tolksdorf C, Gaub HE (2000) Mechanical stability of single DNA molecules. Biophys J 78:1997–2007
Cleaves HJ, Chalmers JH, Lazcano A, Miller SL, Bada JL (2008) A reassessment of prebiotic organic synthesis in neutral planetary atmospheres. Orig Life Evol Biosph 38:105–115
Conrad M (1977) Evolutionary adaptability of biological macromolecules. J Mol Evol 10:87–91
Cui S, Albrecht C, Kühner F, Gaub HE (2006) Weakly bound water molecules shorten single-stranded DNA. J Am Chem Soc 128:6636–6639
Cui S, Yu J, Kühner J, Schulten K, Gaub HE (2007) Double stranded DNA dissociates into single strands when dragged into a poor solvent. J Am Chem Soc 129:14710–14716
Cui S (2009) Have the primary structures of biomacromolecules been selected in a Darwinian fashion to adapt to the surrounding environments of our planet? IUBMB Life 61:860–863
Cui S, Yu Y, Lin Z (2009) Modeling single chain elasticity of single-stranded DNA: a comparison of three models. Polymer 50:930–935
Cui S (2010) The possible roles of water in the prebiotic chemical evolution of DNA. Phys Chem Chem Phys 12:10147–10153
Darwin C (1859) The origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. John Murray, London
Eigen M (1971) Selforganization of matter and the evolution of biological macromolecules. Naturwissenschaften 58:465–523
Ertem G, Ferris JP (1997) Template-directed synthesis using the heterogeneous templates produced by montmorillonite catalysis. A possible bridge between the prebiotic and RNA worlds. J Am Chem Soc 119:7197–7201
Eschenmoser A (1999) Chemical etiology of nucleic acid structure. Science 284:2118–2124
Feng Y, Spisz TS, Hoh JH (1999) Ethanol-induced structural transitions of DNA on mica. Nucleic Acids Res 27:1943–1949
Ferris JP, Hill AR Jr, Liu R, Orgel LE (1996) Synthesis of long prebiotic oligomers on mineral surfaces. Nature 381:59–61
Ferris JP (2002) Montmorillonite catalysis of 30–50 mer oligonucleotides: laboratory demonstration of potential steps in the origin of the RNA world. Orig. Life Evol, Biosph 32
Fu Y, Bai SL, Cui SX, Qiu DL, Wang ZQ et al (2002) Hydrogen-bonding-directed layer-by-layer multilayer assembly: reconformation yielding microporous films. Macromolecules 35:9451–9458
Garrison WM, Morrison DC, Hamilton JG, Benson AA, Calvin M (1951) Reduction of carbon dioxide in aqueous solutions by ionizing radiation. Science 114:416
Gilbert W (1986) The RNA world. Nature 319:618
Hough L, Rogers AF (1956) J Physiol (London) 132:28p
Hugel T, Seitz M (2001) The study of molecular interactions by AFM force spectroscopy. Macromol Rapid Commun 22:989–1016
Hugel T, Rief M, Seitz M, Gaub HE, Netz RR (2005) Highly stretched single polymers: atomic-force-microscope experiments versus ab-initio theory. Phys Rev Lett 94:048301
Janshoff A, Neitzert M, Oberdorfer Y, Fuchs H (2000) Force spectroscopy of molecular systems-single molecule spectroscopy of polymers and biomolecules. Angew Chem Int Ed 39:3212–3237
Kvenvolden KA, Lawless J, Pering K, Peterson E, Flores J et al (1970) Evidence for extraterrestrial amino-acids and hydrocarbons in the Murchison meteorite. Nature 228:923–926
Löb W (1913) Ueber das Verhalten des Formamids unter der Wirkung der stillen Entlandung. Ein Beitrag zur Frage der Stickstoff-Assimilation. Ber Dtsch Chem Ges 46:684–697
Landweber LF (1999) Experimental RNA evolution. Trends Ecol Evol 14:353–358

Lauterbur PC (2008) The spontaneous development of biology from chemistry. Astrobiology 8:3–8
Lawrence DS, Jiang T, Levett M (1995) Self-assembling supramolecular complexes. Chem Rev 95:2229–2260
Lazcano A, Bada JL (2003) The 1953 Stanley L. Miller experiment: fifty years of prebiotic organic chemistry. Orig Life Evol Biosph 33:235–242
Leemhuis H, Kelly RM, Dijkhuizen L (2009) Directed evolution of enzymes: library screening strategies. IUBMB Life 61:222–228
Lehn J-M (1988) Supramolecular chemistry—scope and perspectives molecules, supermolecules, and molecular devices (Nobel Lecture). Angew Chem Int Ed 27:89–112
Lehn J-M (1990) Perspectives in supramolecular chemistry—from molecular recognition towards molecular information-processing and self-organization. Angew Chem Int Ed 29:1304–1319
Lehn J-M (1993) Supramolecular chemistry. Science 260:1762–1763
Lehn J-M (1995) Supramolecular chemistry: concepts and perspectives. Wiley-VCH, Weinheim
Lehn J-M (2002) Toward complex matter: supramolecular chemistry and self-organization. Proc Natl Acad Sci USA 99:4763–4768
Lehn J-M (2007) From supramolecular chemistry towards constitutional dynamic chemistry and adaptive chemistry. Chem Soc Rev 36:151–160
Levy Y, Onuchic JN (2006) Water mediation in protein folding and molecular recognition. Annu Rev Biophys Biomol Struct 35:389–415
Liu C, Cui S, Wang Z, Zhang X (2005) Single chain mechanical property of Poly(N-vinyl-2-pyrrolidone) and interaction with small molecules. J Phys Chem B 109:14807–14812
Lohman TM, Bjornson KP (1996) Mechanisms of helicase-catalyzed DNA unwinding. Annu Rev Biochem 65:169–214
Lum K, Chandler D, Weeks JD (1999) Hydrophobicity at small and large length scales. J Phys Chem B 103:4570–4577
Miller SL (1953) A production of amino acids under possible primitive earth conditions. Science 117:528
Miller SL (1998) The endogenous synthesis of organic compounds. In: Brack A (ed) The molecular origins of life: assembling pieces of the puzzle. Cambridge University Press, Cambridge, pp 59–85
Montesi A, Pasquali M, MacKintosh FC (2004) Collapse of a semiflexible polymer in poor solvent. Phys Rev E 69:021916
Oesterhelt F, Rief M, Gaub HE (1999) Single molecule force spectroscopy by AFM indicates helical structure of poly(ethylene-glycol) in water. New J Phys 1(6):1
Pereira GG, Williams DRM (2001) Toroidal condensates of semiflexible polymers in poor solvents: adsorption, stretching, and compression. Biophys J 80:161–168
Piqueira C, Roberto J (2008) A mathematical view of biological complexity. Commun Nonlinear Sci Numer Simul 16:2581–2586
Prokopenko M, Boschietti F, Ryan AJ (2009) An information-theoretic primer on complexity, self-organization, and emergence. Complexity 15:11–28
Rabinowitch E (1945) Photosynthesis, vol 1, pp 61–98. Interscience, New York
Rief M, Clausen-Schaumann H, Gaub H (1999) Sequence-dependent mechanics of single DNA molecules. Nat Struct Biol 6:346–349
Ring D, Wolman Y, Friedmann N, Miller SL (1972) Prebiotic synthesis of hydrophobic and protein amino acids. Proc Natl Acad Sci USA 69:765–768
Robertson MP, Miller SL (1995) An efficient prebiotic synthesis of cytosine and uracil. Nature 375:772
Schlesinger G, Miller SL (1983) Prebiotic synthesis in atmospheres containing CH4, CO and CO2. I. Amino acids. J Mol Evol 19:376–382
Schoening K-U, Scholz P, Guntha S, Wu X, Krishnamurthy R et al (2000) Chemical etiology of nucleic acid structure: the a-Threofuranosyl-(3′-2′) oligonucleotide system. Science 290:1347–1351

Smith S, Cui Y, Bustamante C (1996) Overstretching B-DNA: the elastic response of individual double stranded and single stranded DNA molecules. Science 271:795–799
Sukhishvili SA, Granick S (2000) Layered, erasable, ultrathin polymer films. J Am Chem Soc 122:9550–9551
Sutherland JD, Whitfield JN (1997) Prebiotic chemistry: a bioorganic perspective. Tetrahedron 53:11493–11527
Tang BL (2007) Emergence of life–how and where? -An update based on recent ideas. Prog Nat Sci 17:500–510
Tian F, Toon OB, Pavlov AA (2005) A hydrogen-rich early Earth atmosphere. Science 308:1014–1017
Wächtershäuser G (2000) Origin of life: life as we don't know it. Science 289:1307–1308
Wöhler F (1828) Ueber das organische synthese. Ann Physik 12:253
Wills C, Bada J (2000) The spark of life: Darwin and the primeval soup, Cambridge
Yockey HP (1992) Information theory and molecular biology. Cambridge University Press, Cambridge
Yockey HP (1995a) Information in bits and bytes. Reply to Lifson's review of 'information theory and molecular biology'. BioEssays 17:85–88
Yockey HP (1995b) Comments on "let there be life: thermodynamic reflections on biogenesis and evolution" by Avshalom C. Elitzur J Theor Biol 176:349–355
Yockey HP (1997) Walther Lob, Stanley L. Miller and prebiotic "building blocks" in the silent electrical discharge. Perspect Biol Med 41:125–131
Yockey HP (2000) Origin of life on earth and Shannon's theory of communication. Comput Chem 24:105–123
Yockey HP (2005) Information theory, evolution and the origin of life. Cambridge University Press, Cambridge
Zhang G, Winnik FM, Wu C (2003) Structure of a collapsed polymer chain with stickers: a single- or multiflower? Phys Rev Lett 90:035506
Zhang WK, Zhang X (2003) Single molecule mechanochemistry of macromolecules. Prog Polym Sci 28:1271–1295
Zou S, Schoenherr H, Vancso GJ (2005) Stretching and rupturing individual supramolecular polymer chains by AFM. Angew Chem Int Ed 44:956–959

# Part IV
# Evolutionary Mechanisms

# Chapter 9
# Domain Architecture Evolution of Metazoan Proteins

**László Patthy**

**Abstract** Appearance of novel multidomain proteins with novel domain architectures (DAs) is closely associated with biological innovations therefore there is a growing interest in genome-scale analysis of the evolutionary history of multidomain proteins. A prerequisite of these studies, however, is that the protein sequences compared are correct and that their evolutionary relationship is correctly defined. We have shown that in the case of most Metazoan proteomes the contribution of mispredicted sequences to apparent DA differences of orthologous and paralogous proteins is greater than the contribution of true gene rearrangements and that analyses of DA evolution frequently suffer from confusing paralogous multidomain proteins (that evolved via gene duplication) with epaktologous multidomain proteins (that are related only through the independent acquisition of the same domain types). Since these methodological errors preferentially increase apparent terminal DA change they may lead to the erroneous conclusion that gene fusion played a dominant role in DA evolution of Metazoan proteins, whereas the contribution of exon shuffling was negligible. In contrast with such conclusions, our studies on high quality datasets of orthologous and paralogous proteins have confirmed that exon shuffling played a major role in the evolution of multidomain proteins of Metazoa.

## 9.1 Evolutionary Significance of Multidomain Proteins

**Gene duplication** and speciation are the two most important pathways that may lead to independently evolving replicas of a given protein-coding gene. The evolutionary significance of gene duplication is that it gives rise to duplicates of a

L. Patthy (✉)
Research Centre for Natural Sciences, Hungarian Academy of Sciences,
Institute of Enzymology, Budapest, Hungary
e-mail: patthy@enzim.hu

P. Pontarotti (ed.), *Evolutionary Biology: Exobiology and Evolutionary Mechanisms*,
DOI: 10.1007/978-3-642-38212-3_9, 

gene that may acquire divergent mutations and the duplicates eventually may emerge as new genes with novel or significantly modified functions (Ohno 1972). Similarly, speciation may open the way for adaptive changes in the functions of the diverging orthologs.

Many types of mutations can endow protein-coding genes with advantageous novel or modified functions. Most frequently, the new functions appear stepwise as a result of several point mutations: continuous modification of the original function may eventually appear as a novel function. New functions may also arise as a result of more dramatic changes, e.g., when the diverging genes experience major changes, such as internal gene duplication, fusion with other genes, insertion of pieces of other genes, etc. Unlike point mutations, such mutations frequently lead to a major and immediate change in function. It is now well established that formation of large multidomain proteins by assembly from distinct domains has played an important role in major evolutionary innovations and contributed significantly to the evolution of increased organismal complexity (Patthy 1985, 1999a, b, 2003; Tordai et al. 2005).

There are several reasons why assembly of multidomain proteins from pre-existing domains may have major evolutionary significance. First, this is the fastest way to create novel, structurally viable (properly folded), complex proteins. Second, acquisition of a new domain with an established molecular function (e.g., binding specificity, catalytic activity) can bring about an immediate change in the function of the recipient protein. Third, in multidomain proteins a large number of molecular functions (different binding activities, catalytic activities) may coexist, making such proteins critical components of regulatory or structural networks where multiple interactions (e.g., protein–protein, protein-DNA, protein-RNA, protein-carbohydrate, protein–ligand interactions) are essential. Indeed, recent analyses have confirmed that domain shuffling had a significant influence on the topology of protein–protein interaction networks of Metazoa (Cancherini et al. 2010; de Souza 2012).

As a corollary of their involvement in multiple interactions, formation of novel multidomain proteins contributes significantly to the evolution of increased organismal complexity since the latter reflects the complexity of interactions among genes, proteins, cells, tissues, and organs (Patthy 2003; Tordai et al. 2005). Comparative analyses of the genomes of the fly, worm, and yeast have shown that the multidomain proteins of the fly and worm are far more complex than those of yeast, suggesting that the organismal complexity of metazoans "does not principally depend on the generation of new genes, but on new combinations of protein domains or novel interactions" (Rubin et al. 2000; Jasny 2000). This conclusion has been confirmed and extended by comparative genomic studies of the human genome: the frequency of multidomain proteins was found to increase concomitant with the evolution more complex eukaryotes (Li et al. 2001; Lander et al. 2001). It has been concluded that human genes are more complex than worm or fly genes primarily because "vertebrates appear to have arranged pre-existing components into a richer collection of domain architectures" (Lander et al. 2001). Similarly, comparison of the human genome with the fly and worm genomes has led Venter

et al. (2001) to conclude that one of the most important differences between these genomes is that the human genome contains greater numbers of multidomain proteins with multiple functions and domain architectures (DAs). (Note that in this chapter domain architecture is defined as the linear sequence of constituent domains).

Thus, the most important conclusion from the comparison of genomes of different eukaryotes is that the increase in organismal complexity from yeast to vertebrates is paralleled by the generation of novel multidomain proteins with increasingly more complex domain architectures (DAs).

## 9.2 Molecular Mechanisms Leading to the Formation of Multidomain Proteins

**Unequal crossing over**. In both prokaryotes and eukaryotes, one of the major mechanisms responsible for creating multidomain proteins is unequal crossing over that can lead to tandem duplication of domains. Multidomain proteins formed by this mechanism contain multiple copies of a single type of structural domain.

The majority of multidomain proteins, however, contain different types of domains (i.e., domains that are not homologous to each other) indicating that different parts have different evolutionary origins. Such chimeric proteins are frequently referred to as mosaic proteins.

**Gene fusion**. It is generally accepted that fusion of genes is an important mechanism for the creation of chimeric multidomain proteins in both prokaryotes and eukaryotes. In the simplest case, tandem genes (of identical orientation) are fused by deleting the intergenic region, including the termination codon of the upstream gene and the initiation codon of the downstream gene. In Metazoa, the most plausible pathway for fusion of tandem genes is through co-transcription and alternative splicing of neighboring genes, followed by fixation of genomic changes that favor fusion over separate transcription of the constituent genes (Magrangeas et al. 1998). The significance of this pathway is supported by the observation that transcripts frequently span two adjacent, tandem genes (Akiva et al. 2006; Parra et al. 2006).

It must be emphasized that the DA of the chimeric gene formed by fusion of tandem genes is defined by the relative position of the neighboring genes in the species where gene fusion occurs: the upstream gene provides the N-terminal domain(s), the downstream gene provides the C-terminal domain(s) of the resulting chimeric multidomain protein. An important point for our present discussion is that DA changes resulting from gene fusion are N-terminal from the perspective of the downstream gene and C-terminal from the perspective of the upstream gene, i.e., gene fusion is characterized by terminal changes of the ancestral DA.

It may be expected that fusion of the same gene pair may occur several times independently as long as synteny and orientation of the neighboring genes is preserved. It seems possible that the recent observation of Zmasek and Godzik

(2012) that about 25 % of all currently observed domain combinations have evolved multiple times is explained by recurrent fusion of genes in regions of conserved synteny. Although conserved synteny may severely limit the degree of freedom with which genes (and their domains) may be combined with other genes by gene fusion, genomic rearrangements (translocations) disrupting synteny may promote novel combinations of genes by this mechanism.

**Domain shuffling**. The third major mechanism leading to the formation of novel multidomain proteins with novel DAs is domain shuffling. The term 'domain shuffling' is intended to be restricted to events when a genomic region (encoding domains) is moved from one genomic location to a new genomic location, i.e., a piece of a protein-coding gene (encoding a domain) is inserted into a gene in a different genomic location. From the perspective of DA evolution, a major difference between gene fusion and domain shuffling is that the former may alter the DA of the ancestral form only at the termini whereas domain shuffling does not have this requirement: it may add domains both internally (e.g., between domains) or at the termini. As a corollary, the relative frequency of DA change in internal positions versus N-terminal and C-terminal positions may be used to assess the relative contribution of gene fusion and domain shuffling to DA evolution.

Since domain shuffling is well documented in prokaryotes it is clear that intronic recombination is not an absolute prerequisite of domain shuffling: domain shuffling may not be equated with exon shuffling (Patthy 1996). Some detailed analyses have shown that domain shuffling in prokaryotes may be facilitated by the presence of special recombinogenic DNA sequences in interdomain linker regions (de Chateau and Bjorck 1994, 1996). In a recent study Furuta and Kobayashi (2012) have shown that movement of DNA sequence recognition domains between non-orthologous proteins within a prokaryote genome may occur through homologous recombination of similar DNA sequences.

Although domain shuffling may occur in the absence of introns, it is clear that recombination in introns can significantly increase the rate of domain shuffling. Gilbert (1978) was the first to suggest that middle repetitious sequences present in introns may create hotspots for recombination to shuffle exonic sequences. Since, however, only spliceosomal introns are suitable for exon shuffling and since these appeared at a relatively late stage of eukaryotic evolution, exon shuffling became significant only at relatively late stages of eukaryote evolution (Patthy 1987, 1991, 1999a, b).

In fact, the unquestionable cases of exon shuffling are apparently restricted to 'young' proteins unique to Metazoa. Analysis of the evolutionary distribution of multidomain proteins assembled from modules by intronic recombination suggests that exon shuffling became significant at the time of the appearance of the first multicellular animals and could in fact contribute to the explosive nature of metazoan radiation (Patthy 1996, 1999a, b). Analysis of established cases of exon shuffling has also shown that only a special group of exons (exon sets), the 'symmetrical' class 1-1, class 2-2, and class 0-0 modules (flanked by introns of identical phase) are really valuable for exon shuffling (Patthy 1987). Analysis of a large number of cases where the evolutionary history of the DA change involving

class 1-1 domains could be reliably reconstructed revealed that—as a rule—exons/exons-sets encoding class 1-1 domains were inserted into pre-existing phase 1 introns of the recipient gene (Patthy 1999a, b). The resulting DA changes may be classified as N-terminal, C-terminal, or internal DA changes, depending on the position of the intron where the class 1-1 domain was inserted.

Although these findings clearly support exon shuffling, very little is known about the molecular details of how an exon of a gene is inserted into another gene. The reason for this lack of information is that most known cases of exon insertion occurred hundreds of million years ago and this precludes the reconstruction of the sequences of the ancestral donor and acceptor genes. In many cases even the identity of the donor genes is unknown.

One of the most parsimonious explanations assumes that the same mechanism of intronic recombination that operates during tandem duplications and deletions of exons or fusion of exons of different genes is also responsible for the transfer of exons from one gene to another. The molecular details of these events are quite well understood since there is a wealth of information on contemporary or recent cases of domain deletion and domain duplication. For example, comparison of the structures of normal and abnormal LDL-receptor genes revealed that in many cases misalignment and recombination involving middle repetitious sequences (frequently Alu repeats) of introns flanking the exons was responsible for the deletion or duplication of domains or entire sets of domains. Several cases are known where the introns involved in recombination belong to different genes: intronic recombination may lead to the formation of chimeric protein-coding genes in which parts of different genes are fused (Yang et al. 1995; Jones et al. 2000). Alu-mediated intronic recombination can mediate even the fusion of genes located on different chromosomes. For example, Babcock et al. (2003) were able to trace interchromosomal Alu-mediated fusion between the IGSF3 gene located on chromosome 1p13.1 and the GGT gene located on chromosome 22q11.2.

Based on these data we have suggested earlier that middle repetitious sequences (such as Alu repeats) present in introns may also facilitate domain transposition. According to this hypothesis intra-chromosomal recombination between introns flanking a domain may lead to the formation of an extrachromosomal circular DNA (eccDNA) and the exon of this eccDNA (flanked by introns) is inserted into an intron of another gene (Patthy 1999a, b).

The arguments in favor of a role of this mechanism are indirect. It is well established that non-allelic homologous recombination in Alu repeats predispose Alu-rich regions to translocation and transposition (Zelnick et al. 1987; Hill et al. 2000; Bailey et al. 2003; Antonell et al. 2005). Transposition of members of the VK family of immunoglobulin light-chain genes was shown to be facilitated by the Alu repeats surrounding them: the VK orphons (genes that are located outside the cluster of the VK gene family) probably arose by alignment of nonadjacent members of the VK gene family with subsequent excision of the resulting looped-out region, which then could be inserted into other chromosomal locations. In general, sequence homology between duplicated DNA segments increases the

chance for misalignment during meiosis and may lead to transposition of genes or genomic segments to other chromosomes (Ji et al. 2000; Chai et al. 2003).

By analogy with transposition of whole genes, we hypothesize that recombination of middle repetitious sequences present in introns that flank a domain may facilitate domain transposition. We further assume that, just as misalignment of different duplicons in a gene cluster can lead to the creation of orphons by their transposition to other chromosomal locations, misalignment of different duplicated domains of a multidomain protein can lead to the transposition of a domain (Patthy 1999a, b). It is well known that multidomain proteins, containing tandem-repeated domains frequently undergo misalignment, leading to the contraction and expansion of such genes by domain duplications. During this process looping out and excision of exons (encoding domains) may occur and the exons present on such eccDNAs may be inserted in introns of other genes in distant chromosomal locations.

The plausibility of our eccDNA hypothesis of exon shuffling is supported by the fact that, although the eukaryotic genome is organized in linear chromosomes, a minor fraction of the DNA also exists as a population of small circular molecules. Extrachromosomal circular DNA is ubiquitous in eukaryotic genomes and has been detected in every organism tested, including human tissues, yeast, plants, *Drosophila*, *Xenopus*, mice, hamster, monkeys (Cohen et al. 2010). It is now well established that—although all types of chromosomal DNA are present in eccDNA—chromosomal tandem repeats are over-represented in the population of eccDNA. These include non-coding satellite repeats and tandemly organized coding genes, such as ribosomal DNA or histone genes (Cohen and Mechali 2002; Cohen et al. 2003, 2010). Cohen et al. (2001) have also shown that the formation of eccDNA is independent of chromosomal DNA replication, suggesting that the circles are excised from the chromosomal substrate through intra-chromosomal homologous recombination and, hence, may generate a deletion in the corresponding region of the chromosome. The junctions found in many eccDNA species represent short homologous sequences (Autiero et al. 2002) pointing to the involvement of nonhomologous end joining in the process.

In addition to an eccDNA-mediated pathway, there are other possibilities for transposing exons of one gene into another gene. One mode of exon shuffling is through retroposition, e.g., by integration of reverse transcripts encoding exons of the donor gene into an intron of the recipient gene (Long et al. 2003). Hickey et al. (1989) have suggested that insertion of exons might occur by the same mechanism as the insertion of spliceosomal introns. According to this hypothesis exon shuffling may be a consequence of the occasional inclusion of exon sequences in the retrotransposition cycle of introns. Exon skipping during alternative splicing may yield exons with flanking introns and if such a composite intron–exon-intron structure is reinserted elsewhere in the genome by the same mechanism that governs mobility of single introns (e.g., through reverse splicing, reverse transcription) we would have exon shuffling.

A fourth mechanism capable of exon shuffling—retrotransposon-mediated transduction of exons—has received significant experimental support. During retrotransposition L1 often associates 3′ flanking DNA as a read-through transcript

and may carry such non-L1 sequences to new genomic locations (Rozmahel et al. 1997; Eickbush 1999; Moran et al. 1999; Ejima and Yang 2003; Nisole et al. 2004; Sayah et al. 2004). In a systematic study Xing et al. (2006) have examined all of the full-length SVA elements (members of the youngest retrotransposon family in primates) in the human genome to assess the frequency and impact of SVA-mediated 3′ sequence transduction. They have shown that ≈53 kb of genomic sequences have been duplicated by 143 different SVA-mediated transduction events, suggesting that retrotransposon-mediated sequence transduction may represent an important mechanism for shuffling of exons.

There are major differences between the 'reverse splicing', 'retroposition', 'retrotransposon-mediated transduction', and 'eccDNA' models of exon shuffling. In the latter three models the exon is inserted into a pre-existing intron, whereas the reverse splicing model does not need a recipient intron for exon insertion. In the case of the transduction model a transposon is introduced into the intron of the recipient gene (together with the exon that is transduced), whereas in the retrotransposon and eccDNA models the exon/exon-set is inserted into an intron of the recipient gene without the concomitant insertion of a retrotransposon.

As mentioned above, from an overview of the vast majority of cases of exon shuffling of Metazoan evolution it is clear that exon insertion occurred into a pre-existing intron with the appropriate phase: this observation argues against a significant role of the 'reverse splicing' model in exon shuffling. As for the other models: since the exon shuffling events that contributed significantly to Metazoan evolution occurred hundreds of million years ago prevents the precise reconstruction of the molecular details of the transposition therefore we cannot decide whether they occurred by retroposition, retrotransposon-mediated transduction, or through eccDNA intermediates.

In summary, a survey of the various genetic mechanisms that may change the DA of proteins suggests that unequal crossing over, gene fusion, and domain shuffling are the key 'creative' mechanisms that may increase the complexity of the DA of multidomain proteins.

Some genetic changes, however, may lead to a decrease in the complexity of the DA of multidomain proteins. For example, unequal crossing over that can lead to tandem duplication of domains may also lead to deletion of tandem duplicated domains and gene fission may results in the loss of terminal domains of multidomain proteins. Interestingly, the same genomic features that are essential for exon shuffling (introns of identical phase at the boundaries of the domain that is shuffled) also facilitate the loss of domains acquired by exon shuffling through fixation of exon skipping (Patthy 2008).

## 9.3 Importance of Genome-scale Analyses of Multidomain Proteins

Genome-scale studies on key eukaryotic model organisms have convincingly shown that the increase in organismal complexity from yeast to vertebrates is

paralleled by the generation of novel multidomain proteins with increasingly more complex domain architectures (Rubin et al. 2000; Jasny 2000; Li et al. 2001; Lander et al. 2001; Venter et al. 2001).

Quantitative genome-scale analyses have also provided strong evidence that there is an intimate connection between the propensity of protein domains to form multidomain architectures and organismal complexity (Wuchty 2001; Koonin et al. 2002) and that the number of domains, the number of domain combinations, and the size of the largest connected component of domain-combination networks of each organism increase with the complexity of the organisms (Ye and Godzik 2004). Tordai et al. (2005) have shown that the proportion of multidomain proteins increases in the order Archaea < Bacteria < Fungi < Plants < Metazoa, and that the multidomain proteins of Metazoa tend to be larger than those in Archaea. Quantitative analyses of the abundance of intracellular (cytoplasmic and nuclear), extracellular, and transmembrane multidomain proteins in various groups of eukaryotes have revealed that Metazoa are unique among Eukaryotes in as much as the world of extracellular and transmembrane multidomain proteins has significantly expanded in this lineage (Tordai et al. 2005).

However, despite the rapid progress with genome projects, several important issues of DA evolution of multidomain proteins remain unanswered.

First, although in many cases there is strong evidence for a cause and effect relationship between the creation of novel multidomain protein architectures and the appearance of novel biological functions (Patthy 2003), in most cases it is not clear how a novel multidomain protein contributed to the biology of the lineage in which it appeared. More reliable and more detailed reconstruction of the evolutionary history of multidomain proteins is essential if we wish to correlate the formation of lineage-specific multidomain architectures with the appearance of lineage-specific biological traits.

Second, although it is clear that both gene fusion and domain shuffling are capable of creating novel chimeric domain architectures, very little quantitative information is available about their rates in different evolutionary lineages and how their rates (and their relative contributions) may correlate with lineage-specific differences of genomic features that are likely to account for such differences.

Third, since changes in domain architecture are generally assumed to be relatively rare evolutionary events they are sometimes used as evolutionary markers to resolve outstanding questions of eukaryote evolution. The fact that similar DAs may form several times independently Zmasek and Godzik (2012), however, cautions that detailed analysis of the evolutionary history of multidomain proteins is essential to decide whether two similar DAs arose by convergent or divergent evolution.

## 9.4 Challenges in Genome-scale Evolutionary Analyses of Protein-Coding Genes

Obviously, a prerequisite of genome-scale analyses of the evolution multidomain proteins is that (1) the protein sequences compared are valid, correct, and complete

and that (2) the evolutionary relationship of multidomain proteins compared is correctly defined. Although these requirements seem to be trivial, a survey of recent papers describing analyses of the evolutionary history of the DA of proteins suggests that problems with each of these points may have had a major impact on the validity of the conclusions of genome scale analyses.

As for the first requirement: a general problem of studies on DA evolution is that true change of DA (at the genome level) may be confused with change of DA only at the transcript level, due to alternative splicing. This type of problem is sometimes encountered even in the case of the high quality, manually curated Swiss-Prot section of UniProtKB; different isoforms (with different DA) are presented for orthologous genes with similar genomic structure (Nagy et al. 2011a).

Although the Swiss-Prot section of UniProtKB contains only a few non-valid, incomplete, or erroneous sequences, the TrEMBL section of UniProtKB is heavily contaminated with N-terminally or C-terminally truncated sequences as well as chimeric sequences produced by chromosomal translocations (Nagy et al. 2008). In view of this fact, DA data obtained by analyses of UniProtKB datasets containing both the Swiss-Prot and TrEMBL sections of UniProtKB may be biased in favor of differences at the N- and C-termini of proteins. It is noteworthy in this respect that—based on analyses of whole UniProtKB (Swiss-Prot plusTrEMBL) sets of proteins—Weiner et al. (2006) concluded that domain losses and duplications were more frequent at the ends of proteins. This finding led the authors to conclude that the genetic mechanism leading to DA changes acts predominantly on sequence termini and that modular evolution of proteins is dominated by two major types of events: fusion, on the one hand, and deletion and fission on the other. As discussed in Sect. 9.5, this conclusion is unjustified since it simply reflects the abundance of incomplete and aberrant sequences in the TrEMBL section of UniProtKB (Nagy et al. 2011a; Nagy and Patthy 2011).

Most genome-scale analyses of DA changes rely on databases (e.g., EnsEMBL, RefSeq) in which the majority of protein sequences analyzed is predicted: the accuracy of the predicted sequences depends on the type of genome and the performance of protocols used for the identification of protein-coding genes in genomic sequences. Protein-coding genes encoded by intronless and intron-poor genomes are usually predicted with great specificity and sensitivity. However, correct prediction of the genomic structure of the protein-coding genes of higher eukaryotes with intron-rich genomes is still a very difficult task.

Recent analyses have shown that the exact genomic structure of protein-coding genes of higher eukaryotes is correctly predicted for only about 50–60 % of the genes (Guigo et al. 2006; Nagy et al. 2008; Harrow et al. 2009). In harmony with this conclusion, recently Zhang et al. (2012) have shown that approximately 50 % of the rhesus gene models are missing, incomplete, or incorrect. Similarly, Tu et al. (2012) have shown that in the case of sea urchin S*trongylocentrotus purpuratus* more than half the computational gene model predictions were imperfect, containing errors such as missing exons, prediction of nonexistent exons, erroneous intron/exon boundaries, fusion of adjacent genes, and prediction of multiple genes from single genes. Since these problems are likely to apply to all Metazoan

genomes annotated with automated gene model pipelines, such a high rate of misprediction is expected to have a major impact on conclusions drawn from all types of genome-scale evolutionary analyses of Metazoan proteins.

Indeed, an increasing number of studies show that there is a significant danger that sequence errors may be confused with evolutionary changes of protein sequences. For example, using datasets of several vertebrate genomes Prosdocimi et al. (2012) have shown that the majority of the detected "evolutionary events" (57 %) are in fact artifacts due to erroneous sequences and that these artifacts mask the true functional significance of the evolutionary events.

Obviously, a high rate of misprediction may have significant influence on the results of DA analyses of Metazoan proteins. Although many authors realize that some of the DA differences may result from errors of gene prediction, they have performed their analyses with the implicit assumption that (1) the 'signals' of DA difference resulting from sequence errors are much weaker than those originating from true evolutionary DA changes and that (2) these sequence errors occur at random.

Our recent analyses have shown that neither of these assumptions is valid (Nagy et al. 2011a). We have estimated the contribution of errors of gene prediction to differences in DA of orthologous and paralogous sequences of Metazoa by comparing data obtained on the high quality, manually curated Swiss-Prot database with those obtained on databases containing predicted sequences (e.g., RefSeq, EnsEMBL, and NCBI's GNOMON predicted protein sequences). Our analyses have confirmed that (especially in the case of EnsEMBL and GNOMON predicted sequences) DA differences due to errors in gene prediction may significantly exceed the rate of true DA change. We have also demonstrated that contamination of RefSeq, EnsEMBL, and NCBI's GNOMON databases with mispredicted sequences introduces a bias in DA differences in as much as it increases the proportion of terminal over internal DA differences. The latter observation appears to be an inevitable consequence of the fact that computational gene prediction introduces a strong positional bias in the distribution of errors in as much as the initial and terminal exons of genes are predicted with significantly lower accuracy than internal exons (Bernal et al. 2007). In terms of DA, this means that DA differences due to misprediction are more likely to be observed at the N-terminal end and the C-terminal end than internally.

Our findings thus caution that earlier genome-scale studies based on comparison of predicted (frequently mispredicted) protein sequences may have led to some erroneous conclusions about the evolution of novel domain architectures of multidomain proteins. This issue is discussed in more detail in Sect. 9.5.

As pointed out above, genome-scale analyses of the evolutionary history of multidomain proteins also require that their evolutionary relationships be correctly defined. Establishing the evolutionary relationship of multidoman proteins, however, is much less trivial than in the case of single-domain proteins. Since in the case of chimeric proteins produced by gene fusion or domain shuffling not all parts of two homologous multidomain proteins have the same evolutionary history the exact evolutionary relationship of homologous multidomain proteins may be

defined only through the detailed analysis of the evolutionary histories of their constituent domains (Patthy 1985). Such analyses, however, are usually not performed on a genome-scale.

As a consequence of these problems it is generally accepted that the procedures used for orthology or paralogy group construction and construction of sequence-based gene trees are more likely to misassign multidomain proteins than single domain proteins (Li et al. 2003; Ruan et al. 2008; Chen et al. 2007). Despite these problems, some authors analyzing DA changes have relied on trees determined for entire multidomain proteins (Buljan and Bateman 2009; Buljan et al. 2010) but most studies have circumvented the problem of sequence-based phylogeny of multidomain proteins by using phylogenies based on similarities of domain architectures (Przytycka et al. 2006; Forslund et al. 2008; Koonin et al. 2000; Fong et al. 2007; Moore et al. 2008). The problem with the latter approach, however, is that it may distort true evolutionary relationships: distantly related proteins with more similar DAs may appear to be closely related; closely related proteins with less similar architectures may appear to be distantly related and these procedures tend to underestimate the number of DA changes. Another conceptual problem with such approaches is that they neglect the fact that convergent evolution of the same DA is quite significant (Zhang et al. 2010; Zmasek and Godzik 2012).

In view of the complexity of the evolutionary histories of chimeric multidomain proteins produced by domain shuffling, the usual terms for homology (orthology, paralogy) are not suitable to describe their evolutionary relationships. Realizing this problem, some authors suggested that the use of the concept of orthology is applicable only at the level of domains rather than at the level of proteins, except for proteins with identical domain architectures (Koonin et al. 2000; Ponting and Russell 2002).

The source of the problem, however, is not that the concept of orthology or paralogy is not applicable for multidomain proteins. The problem is that the three major mutually exclusive subtypes of homology (orthology, paralogy, pseudoparalogy) do not account for all types of relationships that may hold for two homologous multidomain proteins. There exists an additional category of homologous multidomain proteins that are neither orthologous (they do not have the same common ancestor in the last common ancestor of the two species), nor paralogous (they do not derive from the same gene that was duplicated within a genome), nor pseudoparalogous (neither of them were acquired by inter-species horizontal gene transfer): their homology is the result of independent acquisition of the same domain-type.

Several attempts have been made to define and distinguish homology of multidomain proteins due to shuffling of domains from other types of homology. Some authors have used the criterion that two sequences are "homologous only if they are encoded by genes that share an ancestral locus" (Song et al. 2008), thus excluding domain shuffling based homology from the world of homologies. Fitch (2000) recommended the use of the term 'partial homology' for cases where homology of two proteins does not hold for the entire length of both proteins. The terms 'partial homology', or 'local homology' (as opposed to 'global homology')

(Krishnamurthy et al. 2007), however, do not solve the problem of homology based on independent acquisition of the same domain type since they are also valid for cases where two sequences are partially homologous but at the same time they are orthologous or paralogous, e.g., because one of them lost or gained a domain. We have proposed a new term for homologs that are neither orthologs nor paralogs/pseudoparalogs of each other yet they are related through the acquisition of homologous domains. Since the basis of their homology is the import of homologous mobile domain(s), we suggested the term epaktolog from the ancient Greek *επακτοζ*, 'imported' (Nagy et al. 2011b).

We have shown that standard procedures used for orthology or paralogy group construction frequently confuse epaktologs and paralogs (Nagy et al. 2011b; Nagy and Patthy 2011). The significance of this problem may be illustrated by the fact that, in TreeFam (Ruan et al. 2008) several trees for orthologous/paralagous multidomain proteins are contaminated with epaktologous multidomain proteins (Nagy and Patthy 2011).

Failure to distinguish paralogs from epaktologs may lead to serious errors in the interpretation of DA differences: if we compare the DA of two paralogs (separated by a single duplication event) we are likely to reconstruct the actual events (gain or loss of domains) that have occurred since the duplication of the ancestral gene, whereas if we compare two epaktologs (as if they were paralogs) we will be misled as to the evolutionary history of DA changes (Nagy et al. 2011b). Epaktologs are most likely to be mistaken for paralogs in the case of epaktologous proteins in which the independently acquired domain was subsequently duplicated (see

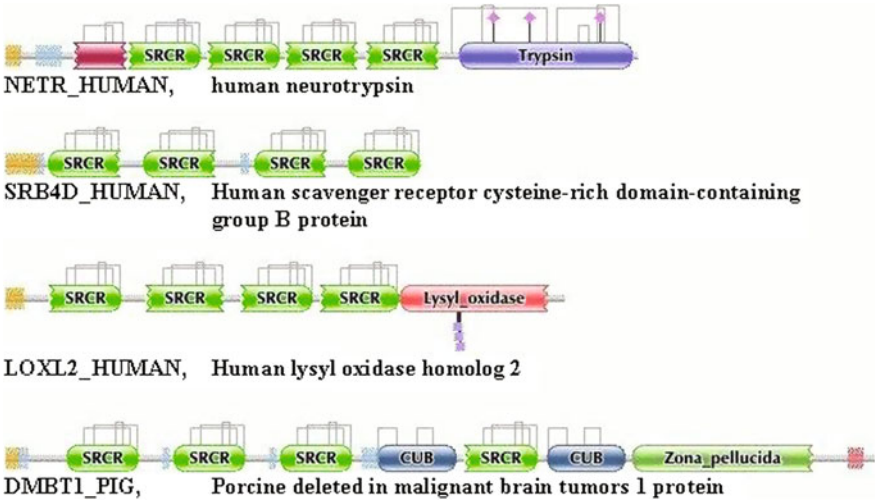



**Fig. 9.1** Domain architectures of some epaktologous multidomain proteins. Note that human neurotrypsin, human scavenger receptor cysteine-rich domain-containing group B protein, human lysyl oxidase homolog 2 and porcine deleted in malignant brain tumors 1 protein are related to each other only because their unrelated ancestors acquired the same domain type, an SRCR domain

Fig. 9.1). Thanks to the presence of tandem duplicated domains the sequence similarity of such epaktologs may be so significant that most techniques used for protein family construction will include the epaktologs in the same protein family (although they are neither orthologs nor paralogs of each other).

We have shown that misidentification of epaktologs as paralogs, not only falsifies the true evolutionary history of multidomain proteins but also has a major impact on conclusions about DA evolution: if the DA of epaktologs is compared their DA is more likely to differ at terminal positions relative to the homologous run of tandem duplicated domains (Nagy et al. 2011b). As a corollary, contamination of protein families with epaktologs not only increases the apparent rate of DA change but also introduces a bias in DA differences in as much as it increases the proportion of terminal over internal DA differences (Nagy et al. 2011b). Our findings thus cautioned that earlier studies based on analysis of datasets of protein families that were contaminated with epaktologs might have led to some erroneous conclusions about the evolution of novel domain architectures of multidomain proteins. This issue is discussed in Sect. 9.5.

## 9.5 Role of Exon Shuffling in Domain Architecture Evolution of Metazoan proteins

The fact that the multidomain proteins crucial for multicellularity of Metazoa and the overwhelming majority of vertebrate-specific novel multidomain proteins were constructed by exon shuffling underlines the extreme importance of this evolutionary mechanism (Patthy 1999a, b, 2003). Analyses of the evolutionary distribution of proteins created by exon shuffling suggested that the rate of creation of novel multidomain architectures by exon shuffling is not constant; certain evolutionary lineages/periods are characterized by bursts of creativity, whereas in other lineages/periods there is little creativity in this respect. The fact that all Eumetazoa share the basic complement of multidomain proteins essential for Metazoan type multicellularity and that these proteins were created by exon shuffling indicate that there was a burst of creation of novel multidomain architectures by exon shuffling in early Metazoa (Patthy 1999a, b, 2003). Similarly, the fact that there is a large collection of multidomain proteins unique to and indispensable for vertebrate biology and that these proteins were created by exon shuffling suggest that there was a burst of creation of novel architectures by exon shuffling in a relatively brief period of chordate evolution (Patthy 1985, 1999a, b, 2003; Jiang and Doolittle 2003).

Consistent with this notion, quantitative analyses by Ekman et al. (2007) have shown that the rate of creation of novel multidomain architectures varies, the rate is lowest (0.1–0.2 per million years) in the lineages of Archea, Bacteria, and yeast (i.e., where exon shuffling plays no role), it is somewhat higher (0.3 per million years) in Plant and Invertebrate (worm, fly) lineages but it is significantly higher in the early Metazoan lineage (1.1 per million years) or the vertebrate lineage (0.9 per million years). The most plausible (and generally accepted) explanation for these

observations is that although all three types of genetic events leading to changes in the domain organization of proteins (internal gene duplication, gene fusion, insertion of gene segments into other genes) can occur in both prokaryotes and eukaryotes, the probability of domain shuffling may be significantly increased through intronic recombination in intron-rich genomes (Patthy 1999a, b). Thus, in early Metazoa and Chordates the high rate of creation of novel multidomain proteins is explained by the fact that they had intron-rich genomes, whereas some lineages—such as the worm and fly lineages—display lower rates since they lost the majority of ancestral metazoan introns (Patthy 1999a, b; Bányai and Patthy 2004).

Recently, however, several studies have questioned the role of exon shuffling in Metazoan evolution. In these genome-scale studies on completely sequenced Archean, Bacterial, and Eukaryotic genomes (including Metazoan genomes) the authors have noted a strong bias in favor of terminal over internal domain architecture changes leading them to conclude that this bias is generally valid for all groups of organisms and that these results support the view that evolution of multidomain proteins (in both prokaryotes and eukaryotes) is dominated by gene fusion, rather than exon shuffling (Bornberg-Bauer et al. 2010; Buljan and Bateman 2009; Buljan et al. 2010; Ekman et al. 2007; Moore et al. 2008; Weiner et al. 2006). In a recent commentary Marsh and Teichmann (2010) have concluded that "although recombination between introns has been speculated to be one of the main mechanisms behind the diverse domain rearrangements observed in complex eukaryotes, it seems to have made a fairly limited contribution to the domain gain events".

Since there are major differences in the organization of genomes/genes of prokaryotes and higher eukaryotes such as Metazoa one would expect that these differences have some impact on the mode and tempo of domain architecture evolution. Indeed, there is a general consensus that the rate of formation of new domain architecture is significantly higher in Metazoa than in prokaryotes or other eukaryotes (e.g., Ekman et al. 2007). It was therefore surprising that this increase in the rate of domain architecture evolution (that is generally attributed to an increased role of exon shuffling in the Metazoan lineage) is not reflected in a shift in favor of internal domain architecture changes.

As outlined in Sect. 9.4, our recent studies have provided some explanations for this contradiction. First, we have shown that a major problem with the aforementioned genome–scale -studies is that the authors ignored the fact that in the case of intron-rich genomes (e.g., those of Metazoa) the reliability of gene prediction is very low (Nagy et al. 2008; Harrow et al. 2009) and that large multidomain proteins (likely to be encoded by a large number of exons) are particularly affected by gene prediction errors (Nagy and Patthy 2011a). A serious consequence is that when low quality datasets are compared mispredicted multidomain protein sequences might appear as novel multidomain architectures resulting in significant overprediction of novelties (Nagy et al. 2011a). Furthermore, we also demonstrated that mispredictions are most likely to appear as terminal domain architecture changes, thus the relative proportion of terminal over internal changes

is falsified, introducing a strong bias in favor of terminal over internal of DA change (Nagy et al. 2011a).

Second, in some studies the authors failed to separate epaktologs and paralogs, a methodological error that also increases the proportion of terminal over internal DA differences (Nagy et al. 2011b).

The issue of whether terminal or internal domain architecture changes occur with greater probability has very important implications for the evolution of multidomain proteins since gene fusion can add domains only at terminal positions, whereas exon shuffling is capable of inserting domains both at internal and terminal positions. As a corollary, in the aforementioned genome-scale studies the overestimation of terminal domain architecture changes necessarily led to the conclusion that in Metazoan evolution gene fusion played a dominant role, whereas the contribution of exon shuffling was negligible.

In contrast with this conclusion our studies on high quality orthologous and paralogous datasets (excluding mispredicted sequences and epaktologous sequences) confirmed that exon shuffling had a major role in the evolution of multidomain proteins of Metazoa (Nagy and Patthy 2011). A recent study using genome information from a variety of eukaryotic species has also confirmed our conclusion that exon shuffling was one of the major evolutionary forces shaping both the genomes and the proteomes of Metazoa (França et al. 2012).

**Acknowledgments** This work was supported by the grant TECH_09_A1-FixPred9 from the National Office for Research and Technology of Hungary.

## References

Akiva P, Toporik A, Edelheit S, Peretz Y, Diber A, Shemesh R, Novik A, Sorek R (2006) Transcription-mediated gene fusion in the human genome. Genome Res 16:30–36

Antonell A, de Luis O, Domingo-Roura X, Perez-Jurado LA (2005) Evolutionary mechanisms shaping the genomic structure of the Williams-Beuren syndrome chromosomal region at human 7q11 23. Genome Res 15:1179–1188

Autiero M, Camarca A, Ciullo M, Debily MA, El Marhomy S, Pasquinelli R, Capasso I, D'Aiuto G, Anzisi AM, Piatier-Tonneau D, Guardiola J (2002) Intragenic amplification and formation of extrachromosomal small circular DNA molecules from the PIP gene on chromosome 7 in primary breast carcinomas. Int J Cancer 99:370–377

Babcock M, Pavlicek A, Spiteri E, Kashork CD, Ioshikhes I, Shaffer LG, Jurka J, Morrow BE (2003) Shuffling of genes within low-copy repeats on 22q11 (LCR22) by Alu-mediated recombination events during evolution. Genome Res 13:2519–2532

Bailey JA, Liu G, Eichler EE (2003) An Alu transposition model for the origin and expansion of human segmental duplications. Am J Hum Genet 73:823–834

Bányai L, Patthy L (2004) Evidence that human genes of modular proteins have retained significantly more ancestral introns than their fly or worm orthologues. FEBS Lett 565:127–132

Bernal A, Crammer K, Hatzigeorgiou A, Pereira F (2007) Global discriminative learning for higher-accuracy computational gene prediction. PLoS Comput Biol 3:e54

Bornberg-Bauer E, Huylmans AK, Sikosek T (2010) How do new proteins arise? Curr Opin Struct Biol 20:1–7

Buljan M, Bateman A (2009) The evolution of protein domain families. Biochem Soc Trans 37(Pt 4):751–755
Buljan M, Frankish A, Bateman A (2010) Quantifying the mechanisms of domain gain in animal proteins. Genome Biol 11:R74
Cancherini DV, França GS, de Souza SJ (2010) The role of exon shuffling in shaping protein-protein interaction networks. BMC Genomics 11(Suppl 5):S11
Chai JH, Locke DP, Greally JM, Knoll JH, Ohta T, Dunai J, Yavor A, Eichler EE, Nicholls RD (2003) Identification of four highly conserved genes between breakpoint hotspots BP1 and BP2 of the Prader–Willi/Angelman syndromes deletion region that have undergone evolutionary transposition mediated by flanking duplicons. Am J Hum Genet 73:898–925
Chen F, Mackey AJ, Vermunt JK, Roos DS (2007) Assessing performance of orthology detection strategies applied to eukaryotic genomes. PLoS One 2(4):e383
Cohen S, Mechali M (2001) A novel cell-free system reveals a mechanism of circular DNA formation from tandem repeats. Nucleic Acids Res 29:2542–2548
Cohen S, Méchali M (2002) Formation of extrachromosomal circles from telomeric DNA in *Xenopus laevis*. EMBO Rep 3:1168–1174
Cohen S, Yacobi K, Segal D (2003) Extrachromosomal circular DNA of tandemly repeated genomic sequences in *Drosophila*. Genome Res 13:1133–1145
Cohen S, Agmon N, Sobol O, Segal D (2010) Extrachromosomal circles of satellite repeats and 5S ribosomal DNA in human cells. Mob DNA 1:11
de Chateau M, Bjorck L (1994) Protein PAB, a mosaic albumin-binding bacterial protein representing the first contemporary example of module shuffling. J Biol Chem 269:12147–12151
de Chateau M, Bjorck L (1996) Identification of interdomain sequences promoting the intronless evolution of a bacterial protein family. Proc Natl Acad Sci USA 93:8490–8495
de Souza SJ (2012) Domain shuffling and the increasing complexity of biological networks. BioEssays 34:655–657
Eickbush T (1999) Exon shuffling in retrospect. Science 283:1465–1467
Ejima Y, Yang L (2003) Transmobilization of genomic DNA as a mechanism for retrotransposon-mediated exon shuffling. Hum Mol Genet 12:1321–1328
Ekman D, Björklund AK, Elofsson A (2007) Quantification of the elevated rate of domain rearrangements in metazoa. J Mol Biol 372:1337–1348
Fitch WM (2000) Homology: a personal view on some of the problems. Trends Genet 16:227–231
Fong JH, Geer LY, Panchenko AR, Bryant SH (2007) Modeling the evolution of protein domain architectures using maximum parsimony. J Mol Biol 366:307–315
Forslund K, Henricson A, Hollich V, Sonnhammer EL (2008) Domain tree-based analysis of protein architecture evolution. Mol Biol Evol 25:254–264
França GS, Cancherini DV, de Souza SJ (2012) Evolutionary history of exon shuffling. Genetica 140:249–257
Furuta Y, Kobayashi I (2012) Movement of DNA sequence recognition domains between non-orthologous proteins. Nucleic Acids Res 40:9218–9232
Gilbert W (1978) Why genes in pieces? Nature 271:501
Guigó R, Flicek P, Abril JF, Reymond A, Lagarde J, Denoeud F, Antonarakis S, Ashburner M, Bajic VB, Birney E, Castelo R, Eyras E, Ucla C, Gingeras TR, Harrow J, Hubbard T, Lewis SE, Reese MG (2006) EGASP: the human ENCODE genome annotation assessment project. Genome Biol 7(Suppl 1):S21–S31
Harrow J, Nagy A, Reymond A, Alioto T, Patthy L, Antonarakis SE, Guigó R (2009) Identifying protein-coding genes in genomic sequences. Genome Biol 10:201
Hickey DA, Benkel BF, Abukashawa SM (1989) A general model for the evolution of nuclear pre-mRNA introns. J Theor Biol 137:41–53
Hill AS, Foot NJ, Chaplin TL, Young BD (2000) The most frequent constitutional translocation in humans, the t (11;22)(q23;q11) is due to a highly specific Alu-mediated recombination. Hum Mol Genet 9:1525–1532

Jasny BR (2000) The universe of Drosophila genes. Science 287:2181
Ji Y, Eichler EE, Schwartz S, Nicholls RD (2000) Structure of chromosomal duplicons and their role in mediating human genomic disorders. Genome Res 10:597–610
Jiang Y, Doolittle RF (2003) The evolution of vertebrate blood coagulation as viewed from a comparison of puffer fish and sea squirt genomes. Proc Natl Acad Sci USA 100:7527–7532
Jones JM, Huang JD, Mermall V, Hamilton BA, Mooseker MS, Escayg A, Copeland NG, Jenkins NA, Meisler MH (2000) The mouse neurological mutant flailer expresses a novel hybrid gene derived by exon shuffling between Gnb5 and Myo5a. Hum Mol Genet 9:821–828
Koonin EV, Aravind L, Kondrashov AS (2000) The impact of comparative genomics on our understanding of evolution. Cell 101:573–576
Koonin EV, Wolf YI, Karev GP (2002) The structure of the protein universe and genome evolution. Nature 420:218–223
Krishnamurthy N, Brown D, Sjölander K (2007) FlowerPower: clustering proteins into domain architecture classes for phylogenomic inference of protein function. BMC Evol Biol 7 (Suppl 1):S12
Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC et al (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921
Li WH, Gu Z, Wang H, Nekrutenko A (2001) Evolutionary analyses of the human genome. Nature 409:847–849
Li L, Stoeckert CJ Jr, Roos DS (2003) OrthoMCL: Identification of ortholog groups for eukaryotic genomes. Genome Res 13:2178–2189
Long M, Deutsch M, Wang W, Betrán E, Brunet FG, Zhang J (2003) Origin of new genes: evidence from experimental and computational analyses. Genetica 118:171–182
Magrangeas F, Pitiot G, Dubois S, Bragado-Nilsson E, Cherel M, Jobert S, Lebeau B, Boisteau O, Lethe B, Mallet J, Jacques Y, Minvielle S (1998) Cotranscription and intergenic splicing of human galactose-1-phosphate uridylyltransferase and interleukin-11 receptor alpha-chain genes generate a fusion mRNA in normal cells. Implication for the production of multidomain proteins during evolution. J Biol Chem 273:16005–16010
Marsh JA, Teichmann SA (2010) How do proteins gain new domains? Genome Biol 11:126
Moore AD, Björklund AK, Ekman D, Bornberg-Bauer E, Elofsson A (2008) Arrangements in the modular evolution of proteins. Trends Biochem Sci 33:444–451
Moran JV, DeBerardinis RJ, Kazazian HH Jr (1999) Exon shuffling by L1 retrotransposition. Science 283:1530–1534
Nagy A, Patthy L (2011) Reassessing domain architecture evolution of metazoan proteins: the contribution of different evolutionary mechanisms. Genes 2:578–598
Nagy A, Hegyi H, Farkas K, Tordai H, Kozma E, Banyai L, Patthy L (2008) Identification and correction of abnormal, incomplete and mispredicted proteins in public databases. BMC Bioinformatics 9:353
Nagy A, Szláma G, Szarka E, Trexler M, Bányai L, Patthy L (2011a) Reassessing domain architecture evolution of metazoan proteins: major impact of gene prediction errors. Genes 2:449–501
Nagy A, Bányai L, Patthy L (2011b) Reassessing domain architecture evolution of metazoan proteins: major impact of errors caused by confusing paralogs and epaktologs. Genes 2:516–561
Nisole S, Lynch C, Stoye JP, Yap MW (2004) A Trim5-cyclophilin A fusion protein found in owl monkey kidney cells can restrict HIV-1. Proc Natl Acad Sci USA 100(101):13324–13328
Ohno S (1972) Evolution by gene duplication. Springer, New York
Parra G, Reymond A, Dabbouseh N, Dermitzakis ET, Castelo R, Thomson TM, Antonarakis SE, Guigo R (2006) Tandem chimerism as a means to increase protein complexity in the human genome. Genome Res 16:37–44
Patthy L (1985) Evolution of the proteases of blood coagulation and fibrinolysis by assembly from modules. Cell 41:657–663
Patthy L (1987) Intron-dependent evolution: preferred types of exons and introns. FEBS Lett 214:1–7

Patthy L (1991) Modular exchange principles in proteins. Curr Opin Struct Biol 1:351–361
Patthy L (1996) Exon shuffling and other ways of module exchange. Matrix Biol 15:301–310
Patthy L (1999a) Genome evolution and the evolution of exon-shuffling—a review. Gene 238:103–114
Patthy L (1999b) Protein evolution. Blackwell Publishing Ltd, Oxford
Patthy L (2003) Modular assembly of genes and the evolution of new functions. Genetica 118:217–231
Patthy L (2008) Alternative splicing: evolution. In encyclopedia of life sciences (ELS). Wiley, Chichester
Ponting CP, Russell RR (2002) The natural history of protein domains. Annu Rev Biophys Biomol Struct 31:45–71
Prosdocimi F, Linard B, Pontarotti P, Poch O, Thompson JD (2012) Controversies in modern evolutionary biology: the imperative for error detection and quality control. BMC Genomics 13:5
Przytycka T, Davis G, Song N, Durand DJ (2006) Graph theoretical insights into evolution of multidomain proteins. Comput Biol 13:351–363
Rozmahel R, Heng HH, Duncan AM, Shi XM, Rommens JM, Tsui LC (1997) Amplification of CFTR exon 9 sequences to multiple locations in the human genome. Genomics 45:554–561
Ruan J, Li H, Chen Z, Coghlan A, Coin LJ, Guo Y, Hériché JK, Hu Y, Kristiansen K, Li R, Liu T, Moses A, Qin J, Vang S, Vilella AJ, Ureta-Vidal A, Bolund L, Wang J, Durbin R (2008) TreeFam: 2008 update. Nucleic Acids Res 36:D735–D740
Rubin GM, Yandell MD, Wortman JR, Gabor Miklos GL, Nelson CR et al (2000) Comparative genomics of the eukaryotes. Science 24(287):2204–2215
Sayah DM, Sokolskaja E, Berthoux L, Luban J (2004) Cyclophilin A retrotransposition into TRIM5 explains owl monkey resistance to HIV-1. Nature 430:569–573
Song N, Joseph JM, Davis GB, Durand D (2008) Sequence similarity network reveals common ancestry of multidomain proteins. PLoS Comput Biol 4(5):e1000063
Tordai H, Nagy A, Farkas K, Banyai L, Patthy L (2005) Modules, multidomain proteins and organismic complexity. FEBS J 272:5064–5078
Tu Q, Cameron RA, Worley KC, Gibbs RA, Davidson EH (2012) Gene structure in the sea urchin Strongylocentrotus purpuratus based on transcriptome analysis. Genome Res 22:2079–2087
Venter JC, Adams MD, Myers EW, Li PW, Mural RJ et al (2001) The sequence of the human genome. Science 291:1304–1351
Weiner J 3rd, Beaussart F, Bornberg-Bauer E (2006) Domain deletions and substitutions in the modular protein evolution. FEBS J 273:2037–2047
Wuchty S (2001) Scale-free behavior in protein domain networks. Mol Biol Evol 18:1694–1702
Xing J, Wang H, Belancio VP, Cordaux R, Deininger PL, Batzer MA (2006) Emergence of primate genes by retrotransposon-mediated sequence transduction. Proc Natl Acad Sci USA 103:17608–17613
Yang J, Brown MS, Ho YK, Goldstein JL (1995) Three different rearrangements in a single intron truncate sterol regulatory element binding protein-2 and produce sterol-resistant phenotype in three cell lines: role of introns in protein evolution. J Biol Chem 270:12152–12161
Ye Y, Godzik A (2004) Comparative analysis of protein domain organization. Genome Res 14:343–353
Zelnick CR, Burks DJ, Duncan CH (1987) A composite transposon 3′ to the cow fetal globin gene binds a sequence specific factor. Nucleic Acids Res 15:10437–10453
Zhang Q, Zmasek CM, Godzik A (2010) Domain architecture evolution of pattern-recognition receptors. Immunogenetics 62:263–272
Zmasek CM, Godzik A (2012) This déjà vu feeling-analysis of multidomain protein evolution in eukaryotic genomes. PLoS Comput Biol 8(11):e1002701

# Chapter 10
# Of Trees and Bushes: Phylogenetic Networks as Tools to Detect, Visualize and Model Reticulate Evolution

**Antonio Hernandez-Lopez**

**Abstract** Phylogenetic trees have been the main tools for representing evolutionary relationships among biological entities at the level of species and above, and their use has greatly facilitated the discussion and testing of phylogenetic and evolutionary hypotheses. However, they are not well suited to model nor represent well known non-vertical evolutionary events such as recombination, horizontal gene transfer, and polyploid and hybrid speciation. Recent advances in phylogenetic network estimation and implementation to genomic data sets, of which I present a brief summary, can better account for reticulated and complex evolutionary histories.

> *Molecular phylogeneticists will have failed to find the true tree, not because their methods are inadequate or because they have chosen the wrong genes, but because the history of life cannot properly be represented as a tree*
>
> Ford Doolittle

## 10.1 Trees or Networks?

Phylogenetic trees have been the main tools for representing evolutionary relationships among biological entities at the level of species and above, and their main goal is to uncover the evolutionary relationships between different species or taxa, and to ultimately understand the evolution of life on Earth. The use of bifurcating trees as explicit evolutionary models goes back to Charles Darwin,

A. Hernandez-Lopez (✉)
LATP UMR—CNRS 7353, Evolution Biologique et Modélisation, Aix-Marseille Université, Case 19, 3 Place Victor Hugo 13331, Marseille CEDEX 3, France
e-mail: antonio.hernandezlopez@univ-provence.fr

P. Pontarotti (ed.), *Evolutionary Biology: Exobiology and Evolutionary Mechanisms*,
DOI: 10.1007/978-3-642-38212-3_10, 

who famously used a tree to illustrate the notion of gradual species evolution in his seminal book The Origin of Species. Ever since, biologists have come to embrace reconstruction of phylogenetic trees as a major research goal in itself (Felsenstein 2001; Huelsenbeck et al. 2000; Huelsenbeck et al. 1997), and indeed their use has greatly facilitated the discussion and testing of phylogenetic and evolutionary hypotheses, as they are well suited to represent evolutionary histories in which the main events are speciation (at the internal nodes of the tree) and descent with modification (along the edges of the tree).

However, well-known events such as recombination, horizontal gene transfer, and polyploid and hybrid speciation cannot be modeled by bifurcating trees. Meiotic recombination occurs in every generation at the level of individual chromosomes; sexual recombination commonly acts at the population level and recombines the evolutionary histories of genomes. Hybrid speciation is very common in large groups of organisms: plants, fungi, fish, frogs, birds, mammals, and many lineages of invertebrates, and horizontal gene transfer is ubiquitous in bacteria. Although the mixing of different evolutionary histories (reticulation) has long been widely appreciated and acknowledged at population level and below, there has been comparatively little work on mathematical modeling and computational methods for studying and estimating reticulate evolution at the species level or above. In this survey, I will address the biological aspects of reticulate evolution, and briefly present existing mathematical and computational methods to infer phylogenetic networks with an emphasis on prokaryotes.

### *10.1.1 Biological Causes of Reticulate Evolution: Sexual Eukaryotes*

The term reticulation refers to the lack of independence between separate evolutionary groups, and occurs when two or more lineages are combined at some level of biological organization: chromosomes, genomes, or species. Reticulation within lineages of sexually reproducing organisms occurs during each round of sexual reproduction, through meiotic recombination. Each parent contributes half of its original nuclear genome—one sister chromatid from each chromosome—with each of these chromosomes having themselves undergone meiotic recombination during the process of producing the haploid gametes (Fig. 10.1a). Because different parts of each parent's contribution to the genome of the next generation may have a different evolutionary history from that of the other parent, sexual (meiotic) recombination is a form of population-level reticulation (Fig. 10.1b). Organellar genomes (mitochondria and chloroplasts) are usually inherited uniparentally so they do not usually undergo any sort of sexual recombination. But reticulation can also occur among lineages through a special kind of sexual recombination involving different species, a phenomenon called hybridization.

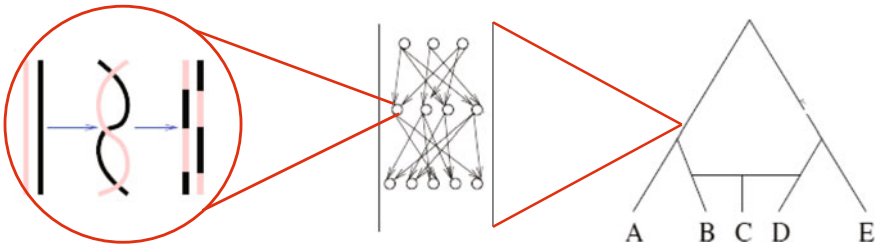



**Fig. 10.1** Reticulation at various levels: **a** individual (chromosomal) level, **b** population level, and **c** species level

#### 10.1.1.1 Hybridization and Polyploidization

The origin of a new species resulting from this processes is called hybrid speciation, and can happen in two different ways: in diploid or homoploid hybrid speciation the new species have the same number of chromosomes as its parent species, while in polyploid hybridization (autopolyploidization, allopolyploidization) it has the sum of the number of the parental species. Diploid hybrid speciation is a normal sexual recombination event where gametes have a haploid complement of the chromosomes from each parent, but the gametes that form the zygote come from different species. In nearly all cases, both parental species must have the same number of chromosomes. In this case, successful backcrossing to the parents is possible, so it is thought that the hybrids have to be isolated by undergoing selection eventually resulting in reproductive isolation from the parental species (Rieseberg and Carney 1998). Even if hybridization does not always results in speciation, it can incorporate large DNA segments into the genomes of the hybridizing species in a phenomenon called introgression or admixture (Fig. 10.1c). For example, modern humans got up to 4 % of genome sequence through hybridization and introgression with the closely related Neanderthals (Green et al. 2010).

Allopolyploidization is hybrid speciation between two species resulting in offspring with the complete diploid chromosome complement of both its parents. Each parent need not have the same number of chromosomes. Allopolyploidization results in instantaneous speciation because any backcrossing to the diploid parents would produce unviable or sterile polyploid offspring. It is very common in plants, and some of the best examples are known in the *Brassica* genus, with the genomes of three ancestral species combined to create three of the common contemporary vegetables and oilseed crop species (Lysak et al. 2007). Autopolyploidization can be considered as a specialized form of normal (bifurcating) speciation, rather than as a true hybrid speciation, as only a single parental species is involved in its production. Autopolyploids can arise from a spontaneous, naturally occurring genome duplication, like in the potato (Huang et al. 2011).

#### 10.1.1.2 Horizontal Gene Transfer: "Asexual" Prokaryotes

Even if the term can be applied generally to introgression through hybridization, it is commonly used to refer to asexual DNA transfers. In horizontal (or lateral) gene transfer (HGT for short), genetic material is transferred from one lineage to another. Even if bacteria and archaea reproduce only clonally, they can exchange DNA horizontally through different processes, the most studied being transformation, conjugation, transduction, and gene transfer agents. Transformation involves the uptake of naked DNA from the environment, and it is well known for the transformation of harmless strains to virulent (Thomas and Nielsen 2005). Conjugation is the transfer of genetic material, mediated by plasmids, by a bridge-like connection between two cells (Wozniak and Waldor 2010). Transduction is DNA acquisition through viral vectors, and gene transfer agents (GTA) are phage-like DNA-vehicles that are produced by a donor cell and released to the environment (Lang and Beatty 2007). Another recently discovered mechanism are nanotubes, tubular protrusions composed of membrane components that can bridge between neighboring cells, and conduct the transfer of DNA and proteins (Dubey and Ben-Yehuda 2011). Furthermore, massive genetic exchange within an Archaean lineage (*Haloferax* spp.) that resembles sex in eukaryotes—cells fuse and huge segments of DNA are recombined—has been recently described, challenging the current understanding of barriers to genetic exchange between prokaryotic species, as there is surprisingly little "sexual" isolation between the species involved (Naor et al. 2012).

Given their clonal mode of reproduction, HGT in prokaryotes takes place at all taxonomic levels: from individuals of the same population, up to inter-kingdom and domain transfers (Lawrence et al. 2011). Regardless of the process, homologous recombination is perhaps the most important mechanism for integrating donor DNA into a recipient genome after a transfer (Majewski et al. 2000; Vulic et al. 1997). Homologous recombination requires incoming DNA to be highly similar to the recipient DNA, and so is usually more common between closely related bacteria, constituting a barrier to HGT (Didelot et al. 2011; Fraser et al. 2005; Raymond et al. 2010; Thomas and Nielson 2005; Zawadzki et al. 1995). However, the insertion of a foreign DNA segment into an arbitrary (non homologous) position in the recipient genome is also possible, if less frequent, through transposable elements, allowing HGT between distantly related bacteria and inter-kingdom transfers (Arias et al. 2012; Ochman et al. 2000).

#### 10.1.1.3 Gene Duplication, Loss, and Incomplete Lineage Sorting

These three phenomena are not truly reticulate in nature, in the sense that no horizontal inheritance is involved, but they do produce reticulate-like patterns (i.e., incongruent gene—species trees), and it can be very difficult to distinguish them

from true horizontal inheritance. Gene duplication is a major mechanism that generates new genetic material, novel functions or structures, molecular diversity, and eventually adaptation. It can be defined as any duplication of a region of DNA that contains a gene; it may occur as an error in homologous recombination (both meiotic and asexual), a retrotransposition event, or duplication of an entire chromosome or even genome, as in autopolyploidization (Zhang 2003). The second copy of the gene is often free from selective pressure, so it accumulates mutations faster than a functional single-copy gene, and can have different evolutionary outcomes: neo-functionalization (gene conversion), sub-functionalization, or loss of function (pseudogenization) (Conant and Wolfe 2008). The two copies resulting from a duplication event are called paralogs, as opposed to orthologs, which are duplicated as a result of a speciation event. Lineage-specific gene losses of well-established genes can be due to deletion events or to pseudogenizations. Failure of correctly identifying orthologs and paralogs produce apparently reticulated evolution patterns. Deletions can be caused by errors in chromosomal crossover during meiosis, or during bacterial recombination and DNA repair. In bacteria, gene fragmentation and silencing is also common, particularly in endosymbiont or parasitic bacteria (Amiri et al. 2003). After a certain time it is not possible to differentiate between the two cases (pseudogenization and deletions) due to the accumulation of mutations in pseudogenes. Incomplete lineage sorting (ILS) refers to the retention and incomplete random coalescence of alleles at many loci independently, due to short intervals between speciation events. This is more likely if the effective population sizes are large relative to the divergence time after speciation. In this case, genetic drift is unlikely to have time to bring genes to fixation before subsequent divergences (Pamilo and Nei 1988).

## 10.2 Detecting Reticulate Evolution

From here on, any DNA incorporation event that involves two distinct species, where the donor's genetic material is inserted into the host (recipient) genome, will be referred to as HGT, regardless of the underlying biological mechanisms involved. Similarly, I will refer to reticulation as the result of HGT. The main challenge in detecting reticulation derived from HGT is that patterns in the data that are suggestive of reticulation may also be present due to other factors (ILS, gene duplication/loss, homoplasy, sequence saturation), making it very difficult to distinguish between the different causes (more than one can cause reticulation); see (Rokas et al. 2003) for further discussion of the topic. Computational methods developed to infer HGT events are based on two major principles: atypical sequence composition left by HGT candidates in the host genome, and the evolutionary history (phylogenetic signal) of the host with respect to the HGT candidate (Fig. 10.2).

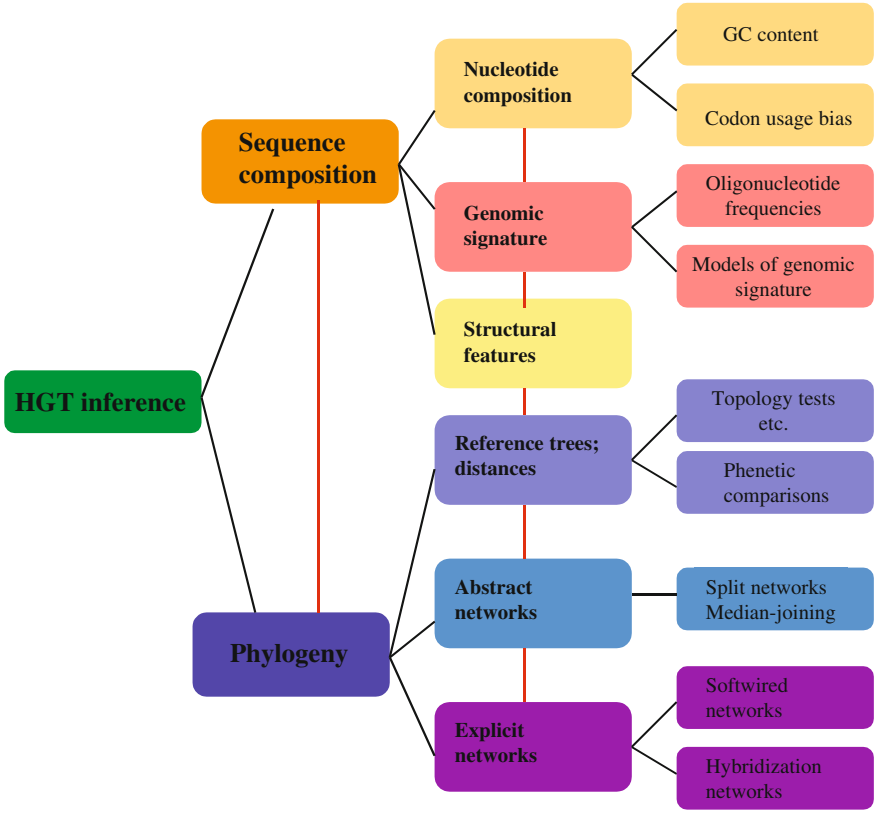



**Fig. 10.2** Main methods used to detect horizontal gene transfer (*HGT*). Sequence composition-based methods compare *HGT* candidate's nucleotide composition, genomic signature, or structural features to the host's genomic average. Phylogeny-based methods search for incongruence in the evolutionary history of the host and the *HGT* candidate. Phylogenetic networks methods are further subdivided into abstract methods that help visualizing incongruencies between gene trees, and explicit methods that represent precise reticulation events. These methods can, and often are, used simultaneously to detect reticulation in a given data set (*red lines*)

### *10.2.1 Sequence Composition*

Given the unequivocal evidence of the importance of HGT on prokaryotic evolution, computational methods used in genome-wide HGT detection were mainly developed, and are more often applied to prokaryotes than to eukaryotes (Boto 2010). Nevertheless, apart from the computational obstacles—not negligible when analyzing the much larger Eukaryotic genomes—most computational methods can be adapted to detect HGT in Eukaryotes (Boschetti et al. 2012; Mallet et al. 2010). In the context of these methods, if a fragment of the genome strongly deviates

from the genomic average, it is considered as a possible horizontal transfer. On the basis of such approaches, a database of genes acquired by horizontal transfer in prokaryotes has been established at www.fut.es/debb/HGT/.

#### 10.2.1.1 Nucleotide Composition: GC Content and Codon Usage Bias

Bacteria display very large GC content variations (Nakabachi et al. 2006; Zhandong et al. 2008), and even within a closely related group of α-Proteobacteria, values range from about 30 % to about 65 % (Bentley and Parkhill 2004). Such differences have been exploited in detecting HGT events: a strikingly different GC content of a genome segment is an indication of its foreign origin (Daubin et al. 2003). The different evolutionary histories that determine diverse GC contents also account for different synonymous codon preferences, known as codon usage bias (Grantham et al. 1981; Supek et al. 2010). In one of the first papers on HGT detection (Médigue et al. 1991), multivariate analysis of codon usage was applied to identify transferred genes, with two features characterizing HGT candidate genes: first, they had a strong codon usage bias, and second, the preferred codons of HGT candidate genes were distinct from those preferred in the host's genome.

#### 10.2.1.2 Genomic Signature

GC content and codon usage bias rely on mononucleotide frequency. But the use of larger sequence fragments (oligonucleotides) allows more discriminatory power. In fact, oligonucleotide frequencies vary less along a genome than between genomes, and constitute a genomic signature: a deviation from the genomic signature on the host makes the fragment an HGT candidate (Campbell et al. 1999). A more complex way of capturing the genomic signature, is the use of models (Markov, Bayesian) to identify clusters of genes that show atypical sequence composition compared to typical ones, and are thus likely to be recently integrated foreign elements (Cortez et al. 2009; Nakamura et al. 2004).

#### 10.2.1.3 Structural Features

The structural features of a DNA molecule include interaction energies between neighboring base pairs, the twist that makes two bases of a pair non-coplanar, or DNA deformability induced by the proteins shaping the chromatin. These structural features can be encoded in a numerical sequence for correlation and frequency analysis. The autocorrelation analysis of such a numerical sequence shows characteristic periodicities in complete genomes (Herzel et al. 1999). In fact, upon detecting Archaea-like regions in thermophilic bacteria (Nelson et al. 1999), periodicity spectra of these regions were compared to those of the homologous regions in the Archaea (Worning et al. 2000), revealing similarities in the periodicity and strongly supporting a case of massive HGT between two kingdoms.

There are a number of limitations for sequence composition methods. First, the recipient's average signature must be clearly recognizable: failure to account for intra-genomic variability results in overestimations (i.e., identifying native segments as possible HGT events; (Guindon and Perriere 2001). Similarly, the transferred segments need to exhibit the donor's signature. These conditions are met only in recent transfers, since transferred segments are subject to the same mutational processes as the rest of the host genome, so their distinct signatures eventually disappear (Lawrence and Ochman 1997). Thus, ancient events are likely to be missed by sequence composition methods. This is also the case for inserted segments previously adapted to the recipient's genome, as in prophage insertions (Vernikos et al. 2007). A remarkable example of the reduced power of sequence composition methods comes from *Bdellovibrio bacteriovorus*, a predatory $\delta$-Proteobacterium. An initial analysis based on the homogeneous GC content, found that its genome was resistant to HGT (Rendulic et al. 2004). However, subsequent phylogenetic analysis identified a number of ancient LGT events in the genome (Gophna et al. 2006).

### *10.2.2 Phylogeny*

If transferred elements have different evolutionary histories, the use of phylogenetic trees could in principle identify discrepancies due to transfer events and the time of their occurrence. However, tree-based methods will be discussed briefly, as the focus of the survey is on network-based approaches. Phylogenetic methods detect inconsistencies in gene and genome evolutionary history by reconstructing gene trees and reconciling them with a reference tree. The common sense conclusion that many genes inherited through lineal descent would override the confusing signal generated by a few genes acquired through horizontal transfer is too simplistic, if not wrong (Brown et al. 2001; Teichmann and Mitchison 1999). And the main limitations of tree-based approaches stems from the gene trees versus species tree old problem: discrepancies between the trees derived from different genes do not necessarily indicate reticulate evolution, but may simply testify to the incongruent evolution of two or more genes, all within a valid, tree-shaped evolution of the species. Furthermore, the selection of the reference tree reconstruction method is still controversial due to the statistical bias involved in creating it (Brown et al. 1994; Hilario and Gogarten 1993; Nesb et al. 2001). Even if there is no doubt in the reference tree, the conflicting phylogenies can be the result of an unrecognized paralogy or a gene loss, as well as an HGT event. Moreover, inferred HGT scenarios are not necessarily unique. The availability of genomes included in the analysis limits the application for two more reasons: first, the computational complexity of reconstructing a gene tree or a species tree is still a challenge. Second, long branch attraction will confound the interpretation when fast evolving orthologous groups are considered.

#### 10.2.2.1 Phenetic Comparisons

This methods estimate sequence similarity as a measure of evolutionary distance—the number of per-site substitutions since the genes diverged from their common ancestor. If a gene has unexpected evolutionary distance from the reference, i.e., its gene family or the genomic average, an HGT event is inferred. In sequence similarity correlations, the evolutionary distances of a group of ortholog genes are expected to be proportional to the evolutionary distances of the respective genomes (Lindell and Penny 2003). If a group of orthologs contains xenologs—inserted by an HGT event—the proportionality of evolutionary relationships will hold for orthologs, but not for xenologs (Novichkov et al. 2004). Phyletic profiles refer to the presence/absence patterns of ortholog group members in reference genomes, and HGT is inferred if unusual genes are found on a phyletic profile, indicated by the presence of top-scoring BLAST hits in an unrelated species (Boschetti et al. 2012; Nelson et al. 1999). Phenetic distances can also be represented as networks, most commonly as networks of gene sharing, to better visualize shared genes, and possible HGT events (Dagan et al. 2008; Popa and Dagan 2011). Often these methods are used as a first step in HGT detection, followed by sequence composition and phylogenetic analyses.

## 10.3 Phylogenetic Networks

In the literature, networks are called in many different ways, with some named by the algorithms that compute them or by mathematical properties that define them, such as "neighbor-nets" or "median networks." Others are named by topological constraints that are imposed on them for computational reasons, such as "galled trees," "galled networks," or "level-k networks." Yet others are named by the types of evolutionary events that they model, such as "hybridization networks," "recombination networks," or "duplication-loss-transfer (DLT) networks." But as with phylogenetic trees, the first major distinction is between unrooted and rooted phylogenetic networks (Fig. 10.3): (1) unrooted, known as abstract networks, are used as a tool to classify and better visualize incompatibilities in datasets; (2) rooted, explicit networks, are used to model evolution by reconstructing the evolutionary history of putative reticulate events (Huson et al. 2010). Phylogenetic networks can be constructed from a wide range of data, including multiple sequence alignments, distance matrices, and sets of trees.

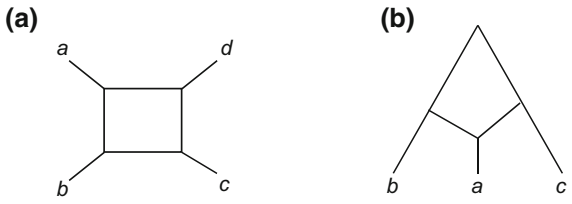



**Fig. 10.3** An unrooted phylogenetic network **a** for taxa (*a, b, c*) *d*, and a rooted phylogenetic network; **b** for taxa (*a*, *b*, *c*) in which the top node *d* is the root

### *10.3.1 Abstract (Unrooted) Networks*

There are many types of unrooted phylogenetic networks in use, such as splits networks, median and quasi-median networks, haplotype networks, and reticulograms. I will only describe the networks suited for above-species levels (except for splits networks from distance data, which can in principle be used for both population and above species levels).

#### 10.3.1.1 Split Networks

In an unrooted phylogenetic tree, a node defines a bipartition into two different and disjoint subsets, known as splits. Splits define the topology of a tree, and are the basis for comparing different topologies or to compute consensus trees. Any set of splits that are compatible between two trees can be summarized as a phylogenetic consensus tree. But incompatible sets of splits are not considered in tree reconciliation, and they result in unresolved, comb-like trees. A graphical representation of incompatible splits constitutes a splits network, in which every split is represented by an array of parallel edges (branches).

#### 10.3.1.2 Split Networks from Distances

There are two commonly used methods to compute split networks from distance data. Split decomposition takes a distance matrix as input and produces a set of weighted splits that is "weakly compatible", a property that ensures the corresponding split network is not too complicated (Bandelt and Dress 1992). In practical terms, split decomposition is a very conservative method, in the sense that a split will only be present in the output if there is global support for it in the given data set. For large or very divergent data sets, the method tends to exhibit very low resolution (Huson and Scornavacca 2011). Neighbor-Net takes a distance matrix as input and produces a set of weighted splits that is circular and can be represented as an outer-planar split network, a graph in which no two edges intersect and all labeled nodes lie on the outside of the graph (Bryant and Moulton 2004). Neighbor-Net is more popular than split decomposition because it is less conservative and so does not lose resolution on larger data sets. Moreover, the fact that the output of the method can always be represented by an outer-planar split network and is thus easy to visualize adds to its attraction. Both methods have the attractive property that they produce a set of splits corresponding to a correct tree when given a tree-like matrix, making it easier to visualize tree-like and network-like evolution in the data. In Fig. 10.4, a Neighbor-Net distance analysis of 1,174 concatenated orthologous genes alignment shows conflicting phylogenetic signals among 17 enteric bacteria species that may reflect putative recombination (Retchless and Lawrence 2010).

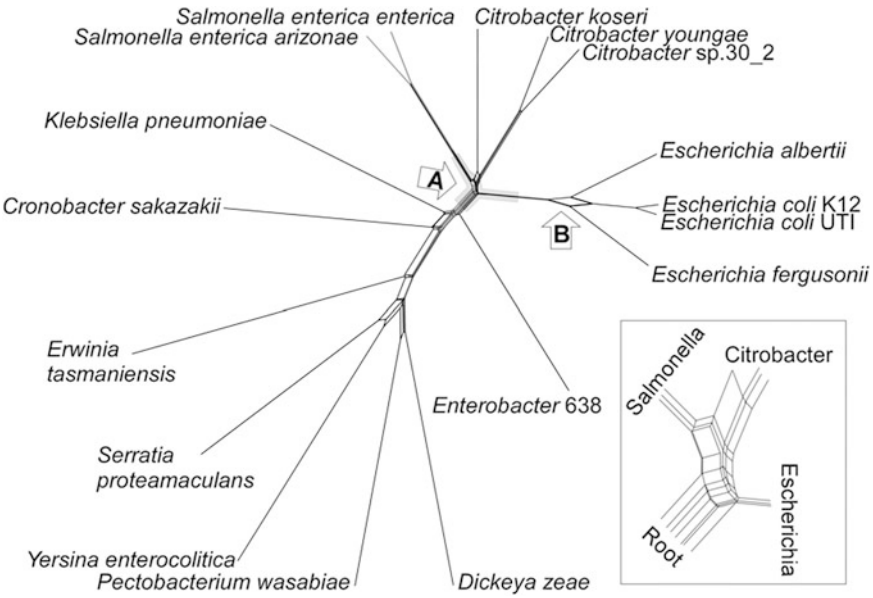



**Fig. 10.4** Split philogenetic network of enteric bacteria (Retchless and Lawrence 2010). The inset reveals the conflicts, probably due to HGT, in the divergence of *Escherichia*, *Salmonella*, and *Citrobacter* (*arrow A*). *Arrow B* also highlights potential HGT events at the divergence of the different *E. coli* starins

### 10.3.1.3 Splits Networks from Trees

Split networks can also be used to visualize conflicting signals present in a collection of unrooted phylogenetic trees, i.e., different gene trees; trees for the same gene computed using different methods or obtained in a Bayesian analysis. The set of majority-consensus splits is defined as the set of all splits that are present in more than 50 % of the input trees. By lowering the threshold to a lower proportion, one obtains a set of splits that will not necessarily be compatible. The split network associated with the set of splits is called a "consensus split network" and can be used to visualize conflicting signals in a set of trees (Holland and Moulton 2003). For sets of trees with missing taxa, the set of taxa that occurs in each tree will often differ between trees. To address this, methods have been developed to compute a "super split network" for a given set of unrooted phylogenetic trees on overlapping but nonidentical taxon sets using the "Z-closure" algorithm (Huson et al. 2004); Whitfield, 2008 #310}. A split network computed from a forest of nearly 102 universal trees for a set of 100 prokaryote species (Fig. 10.5) reveals high levels of topological inconsistency, probably due to horizontal gene transfer (Puigbò et al. 2009).

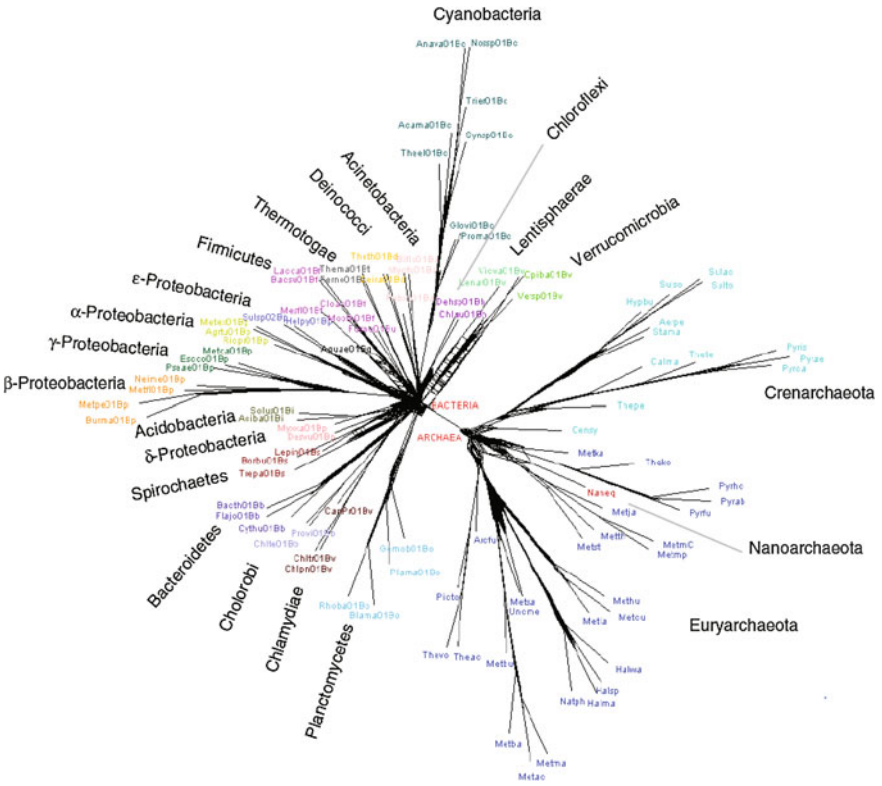



**Fig. 10.5** A consensus split network built from 102 nearly universal trees for 100 prokaryote species shows that the incongruence among the trees is mainly concentrated at the deepest levels (except for the clean archaeal-bacterial split), with a much greater congruence at shallow phylogenetic depths. This suggests that HGT took place mainly during the early divergence of the groups (Puigbò et al. 2009)

### 10.3.1.4 Splits Networks from Sequences

A first approach to build a split network from a sequence alignment is to compute a set of splits that represents the alignment using the "parsimony-splits" method (Bandelt and Dress 1993). This method produces a set of weakly compatible splits using a simple modification of the split decomposition algorithm. The parsimony-splits method has not been used much, probably because the resulting set of splits is usually very similar to the one obtained by the more widely known split decomposition method. Another way of computing an alignment-based split network is to first simplify the alignment to obtain a matrix containing only columns for variable sites. Then, every column defines a different split of the taxon set, and a split network can represent the set of splits obtained in this way. If each edge of the network is labeled by the columns in the alignment that correspond to the split represented by that edge, then the resulting split network is called a median

network (Bandelt et al. 1995). This construction is suitable for data sets that have very few differences in them, and are thus mainly used in phylogeography and population studies.

### *10.3.2 Explicit (Rooted) Networks*

An explicit network is a phylogenetic network where all reticulations can be interpreted as precise biological events, and which contains tree-like parts (bifurcating or tree nodes) and reticulated parts known as "blobs" (nodes with more than two parent edges or reticulate nodes). Any edge leading to a reticulate node is called a reticulate edge and all others are called tree edges. Every edge (branch) of a phylogentic tree represents a subset of the leaves, and these are known as clusters. If all the clusters present in a given data set can be fitted without conflict in a graph, this graph will be a phylogenetic tree. When a phylogenetic network contains all the possible clusters for a given data set in the form of rooted trees and they are connected to each other by reticulate nodes and edges, this is called a soft-wired network. On the other hand, hard-wired networks are those in which the clusters are represented by networks, not by trees. Most of the methods available to date are based on soft-wired clusters (Huson et al. 2010). Perhaps the most important feature of a rooted phylogenetic network is the set of "clusters" that the network represents, as clusters suggest putative monophyletic groups and thus provide hypotheses about the evolutionary relatedness of the taxa under consideration, so usually rooted phylogenetic networks are interpreted as representations of sets of clusters (Huson and Scornavacca 2011).

#### 10.3.2.1 Softwired Networks

In general, rooted phylogenetic networks interpreted in the softwired sense are computationally very hard to build (Kanj et al. 2008). The use of topologically constrained networks such as galled trees (Gusfield et al. 2003; Wang et al. 2000), galled networks (Huson et al. 2009), and level-k networks (Choy et al. 2005), avoid these computational problems and are commonly applied for large data sets. All these methods put constraints on how tangled the reticulations in a rooted phylogenetic network may be, and work in two steps: (1) solve cluster conflicts by deleting taxa and reconstructing the tree with the remaining taxa (maximum compatible clusters); and (2) attach conflicting taxa to the tree with the minimum number of reticulate edges (minimum attachment). One of the first implementations of these algorithms on a large dataset (Fig. 10.6) shows a galled network displaying potential HGT events for nine different gene trees on a set of 279 Archea and Bacteria species (Huson et al. 2009). Figure 10.7 is the galled tree representing all clusters contained in three rooted phylogenetic trees for 23, 210

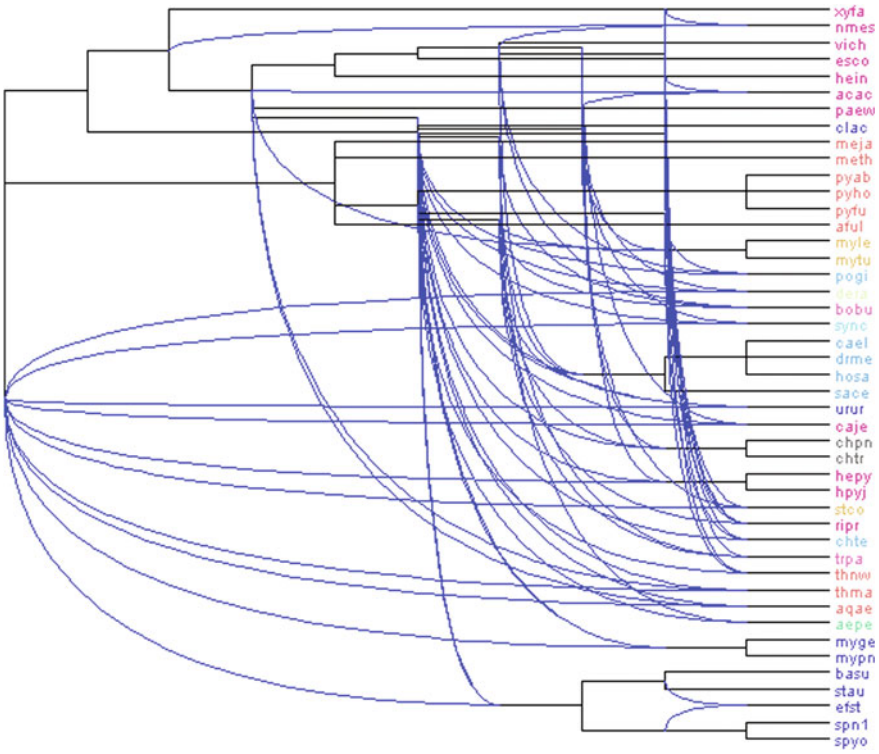



**Fig. 10.6** Galled phylogenetic network for 279 prokaryote species, computed from all clusters contained in more than one of nine different gene trees (Huson et al. 2009). Most of the reticulations (*blue lines*) are seen at deeper nodes, but in contrast to the split network in Fig. 10.5, the reticulations represent putative HGT events, not only phylogenetic conflict

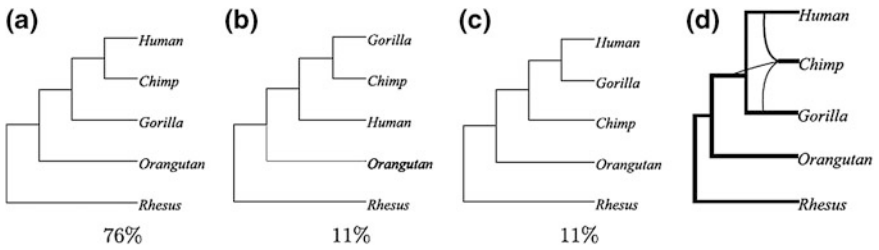



**Fig. 10.7** Reticulate evolution in the human lineage inferred from three rooted phylogenetic trees shown in (**a**), (**b**), and (**c**) supported by 76, 11, and 11 % of all genes studied in Ebersberger et al. (2007). The galled phylogenetic network (**d**) represents all clusters contained in the rooted phylogenetic trees. The line width of each edge is proportional to the number of trees that contain it. Adapted from Huson et al. (2010)

DNA sequence alignments (Ebersberger et al. 2007), showing reticulation in the human lineage most likely due to ancient hybridization/ILS.

#### 10.3.2.2 Hybridization and Recombination Networks

The algorithms used to reconstruct these type of networks, where the tree nodes correspond to speciation events and the reticulate nodes correspond to recombination events, follow the same parsimony principles as galled and other softwired networks, but are used specifically when there are suspected or known hybridization or recombination events (Huson and Scornavacca 2011). The network for grass species in Fig. 10.8 suggests that *P. glycerias* is a hybrid of the lineages leading to *P. melica* and *P. triticium*. In the case of the two other putative hybrid species, *P. lygeum* and *P. chusquea*, their evolution requires the postulation of additional lineages to resolve the fact that they appear to be hybrids of recent and ancient lineages. For recombination networks additional labeling is required, one for all nodes by sequences, and a labeling for all tree edges by the positions in the sequences at which mutations occur (Gusfield et al. 2003; Huson and Klopper 2005; Song and Hein 2005). These labels must be compatible in the sense that the sequences assigned to tree nodes of the network differ exactly by the indicated mutations, whereas the sequences assigned to reticulate nodes must be obtainable from the sequences assigned to the parent nodes by a crossover.

#### 10.3.2.3 DLT Networks

A fundamental and well-understood observation in phylogenetics is that different genes often have different evolutionary histories and thus give rise to different trees. These differences can be due to methodological errors or bad data, but once

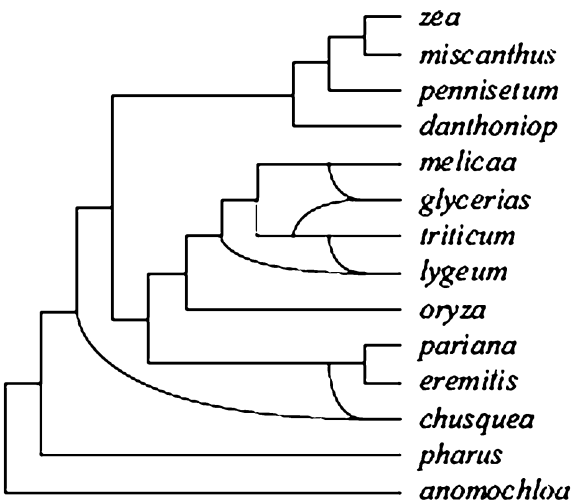



**Fig. 10.8** Hybridization network based on the *phyB* and *waxy* gene trees (Huson and Scornavacca 2011). The network displays a set of putative hybridization events that may explain the differences between two trees. Redrawn from Huson and Scornavacca (2011)

ruling those causes out, various evolutionary processes can cause the incongruence: gene duplications, losses, and HGT (DLT events). Assuming that the species phylogeny is correct, then any gene phylogeny will usually differ from the species tree, and it possible to reconcile all differences between the two trees by postulating an appropriate DLT scenario. Such a scenario provides a mapping of the gene tree onto the species tree that implies certain duplication, loss, and transfer events. Because of the presence of HGT events, this DTL scenario can be represented as a network, but the network can not differentiate between DLT events (Planet et al. 2003; Tofigh et al. 2010).

## 10.4 Limitations and Prospects of Phylogenetic Networks

Phylogenetic networks provide an alternative to phylogenetic trees and are more suitable for data sets where evolution involves significant amounts of reticulate events, such as hybridization, horizontal gene transfer, or recombination. Unrooted phylogenetic networks are now routinely used in phylogenetic analysis and phylogeography, particularly Neighbor-Net, consensus split networks, and median-joining, for data ranging from multiple locus analyses to whole genome alignments (Budroni et al. 2011; Coscollá et al. 2011; Puigbò et al. 2009; Retchless and Lawrence 2010). They have proven very useful in detecting and visualizing phylogenetic conflict between different markers, or for single markers when recombination is involved.

Although rooted phylogenetic networks are conceptually very appealing, the development of suitable methods for their computation remains a formidable challenge. At present, none of the existing methods for computing a rooted phylogenetic method is widely or routinely used as a tool to help understand the evolutionary history of a given set of taxa in terms of mutations, speciation, and specific types of reticulate events. Even if a number of algorithms have been proposed for computing rooted phylogenetic networks, some problems must be overcome. First, many of the algorithms have only proof-of-concept implementations that are too complicated to be used as tools in real studies. Second, the computational problems are often hard, and the algorithms have impractical running times. Third, the calculation of rooted phylogenetic networks must be more closely linked to detailed biological models of reticulate evolution so as to produce more plausible results. Furthermore, from a biological point of view it is important to not only compute one network, but all possible networks. Performance studies also need to be able to measure the error (distance) between the "true" network and the reconstructed one. Such a measure should be symmetric, and should be zero only when the two networks are identical. Such measures are commonplace for trees, but such measures between phylogenetic networks seem much harder to obtain. In a recent work, the first algorithm that calculates all maximum-acyclic-agreement forests for two rooted binary phylogenetic trees on the same set of taxa

was presented (Scornavacca et al. 2012), but additional work needs to be done in this regard.

Reticulate evolution, due to biological events such as horizontal gene transfer, recombination, and hybridization, make current approaches to phylogenetic analysis insufficient, and an accurate estimation of evolutionary histories will require the development of novel simulation tools, reconstruction methods, and mathematical theory for phylogenetic networks. The successful approaches in this area will have to combine population genetics and phylogenetics, and will lead to interesting questions in many technical areas, including statistical inference, molecular phylogenetics, and computer science.

## References

Amiri H, Davids W, Andersson SGE (2003) Birth and death of orphan genes in *Rickettsia*. Mol Biol Evol 20:1575–1587

Arias M, Danchin EGJ, Coutinho P, Henrissat B, Ball S (2012) Eukaryote to gut bacteria transfer of a glycoside hydrolase gene essential for starch breakdown in plants. Mob Genet Elem 2:81–87. doi:10.4161/mge.20375

Bandelt H-J, Dress AWM (1992) A canonical decomposition theory for metrics on a finite set. Adv Math 92:47–105

Bandelt HJ, Dress A (1993) A relational approach to split decomposition. In: Opitz O, Lausen B, Klar R (eds) Information and classification. Springer, Berlin, pp 123–131

Bandelt HJ, Forster P, Sykes BC, Richards MB (1995) Mitochondrial portraits of human population using median networks. Genetics 141:743–753

Bentley SD, Parkhill J (2004) Comparative genomic structure of prokaryotes. Annu Rev Genet 38:771–792. doi:710.1146/annurev.genet.1138.072902.094318

Boschetti C, Carr A, Crisp A, Eyres I, Wang-Koh Y, Lubzens E, Barraclough TG, Micklem G, Tunnacliffe A (2012) Biochemical diversification through foreign gene expression in Bdelloid rotifers. PLoS Genet 8:e1003035. doi:1003010.1001371/journal.pgen.1003035

Boto L (2010) Horizontal gene transfer in evolution: facts and challenges. Proc R Soc Lond B 277:819–827. doi:810.1098/rspb.2009.1679

Brown JR, Masuchi Y, Robb FT, Doolittle FW (1994) Evolutionary relationships of bacterial and archaeal glutamine synthetase genes. J Mol Evol 38:566–576

Brown JR, Douady CJ, Italia MJ, Marshall WE, Stanhope MJ (2001) Universal trees based on large combined protein sequence data sets. Nat Genet 28:281–285

Bryant D, Moulton V (2004) Neighbor-net: an agglomerative method for the construction of phylogenetic networks. Mol Biol Evol 21:255–265

Budroni S, Siena E, Dunning JCH, Seib KL, Serruto D, Nofroni C, Comanducci M, Riley DR, Daugherty SC, Angiuoli SV, Covacci A, Pizza M, Rappuoli R, Moxon ER, Tettelin H, Medini D (2011) *Neisseria meningitidis* is structured in clades associated with restriction modification systems that modulate homologous recombination. Proc Natl Acad Sci 108:4494–4499

Campbell A, Mrzek J, Karlin S (1999) Genome signature comparisons among prokaryote, plasmid, and mitochondrial DNA. Proc Natl Acad Sci 96:9184–9189

Choy C, Jansson J, Sadakane K, Sung WK (2005) Computing the maximum agreement of phylogenetic networks. Theor Comput Sci 335:93–107

Conant GC, Wolfe KH (2008) Turning a hobby into a job: how duplicated genes find new functions. Nat Rev Genet 12:938–950

Cortez D, Forterre P, Gribaldo S (2009) A hidden reservoir of integrative elements is the major source of recently acquired foreign genes and ORFans in archaeal and bacterial genomes. Genome Biol 10:R65. doi:10.1186/gb-2009-1110-1186-r1165

Coscollá M, Comas I, González-Candelas F (2011) Quantifying Nonvertical Inheritance in the evolution of *Legionella pneumophila*. Mol Biol Evol 28:985–1001

Dagan T, Artzy-Randrup Y, Martin W (2008) Modular networks and cumulative impact of lateral transfer in prokaryote genome evolution. Proc Natl Acad Sci 105:10039–10044

Daubin V, Lerat E, Perriere G (2003) The source of laterally transferred genes in bacterial genomes. *Genome Biology* 4(9):R57. doi:10.1186/gb-2003-1184-1189-r1157

Didelot X, Bowden R, Street T, Golubchik T, Spencer C, McVean G, Sangal V, Anjum MF, Achtman M, Falush D, Donnelly P (2011) Recombination and population structure in *Salmonella enterica*. PLoS Genetics 7:e1002191. doi:1002110.1001371/journal.pgen.1002191

Dubey GP, Ben-Yehuda S (2011) Intercellular nanotubes mediate bacterial communication. Cell 144:590–600

Ebersberger I et al (2007) Mapping human genetic ancestry. Moleculart Biol Evol 24:2266–2276

Felsenstein J (2001) The troubled growth of statistical phylogenetics. Syst Biol 50:465–467

Fraser C, Hanage WP, Spratt BG (2005) Neutral microepidemic evolution of bacterial pathogens. Proc Natl Acad Sci 102:1968–1973

Gophna U, Charlebois RL, Doolittle FW (2006) Ancient lateral gene transfer in the evolution of *Bdellovibrio bacteriovorus*. Trends Microbiol 14:64–69. doi:10.1016/j.tim.2005.1012.1008

Grantham R, Gautier C, Gouy M, Jacobzone M, Mercier R (1981) Codon catalog usage is a genome strategy modulated for gene expressivity. Nucleic Acids Res 9:43–74

Green RE, Krause J, Briggs AW et al (2010) A draft sequence of the neanderthal genome. Science 328:710–722

Guindon S, Perriere G (2001) Intragenomic base content variation is a potential source of biases when searching for horizontally transferred genes. Mol Biol Evol 18:1838–1840

Gusfield D, Eddhu S, Langley C (2003) Efficient reconstruction of phylogenetic networks with constrained recombinations. In: Proceedings of the IEEE computer science conference on bioinformatics (CSB), IEEE Computer Society, p 363

Herzel H, Weiss O, Trifonov EN (1999) 10–11 bp periodicities in complete genomes reflect protein structure and DNA folding. Bioinformatics 15:187–193

Hilario E, Gogarten JP (1993) Horizontal transfer of ATPase genes-the tree of life becomes a net of life. Bio Syst 31:111–119

Holland B, Moulton V (2003) Consensus networks: a method for visualizing incompatibilities in collections of trees. In: *WABI* 2003. *Algorithms in bioinformatics.* In: Benson G, Page R (eds) Proceedings of the third international workshop on algorithms in bioinformatics (*WABI*), Springer, Budapest, Hungary. Heidelberg (Germany), pp 165–176

Huang S, Buell CR, Visser RGF et al (2011) Genome sequence and analysis of the tuber crop potato. Nature 475:189–195

Huelsenbeck JP, Rannala B, Yang Z (1997) Statistical tests of host-parasite cospeciation. Evolution 51:410–419

Huelsenbeck JP, Rannala B, Larget B (2000) A Bayesian framework for the analysis of cospeciation. Evolution 54:353–364

Huson DH, Klopper TH (2005) Computing recombination networks from binary sequences. Bioinformatics 21:ii159–ii165

Huson DH, Scornavacca C (2011) A survey of combinatorial methods for phylogenetic networks. Genome Biol Evol 3:23–35

Huson DH, Dezulian T, Klopper T, Steel MA (2004) Phylogenetic supernetworks from partial trees. IEEE/ACM Trans Comput Biol Bioinform 1:151–158

Huson DH, Rupp R, Berry V, Gambette P, Paul C (2009) Computing galled networks from real data. Bioinformatics 25:i85–i93

Huson DH, Rupp R, Scornavacca C (2010) Phylogenetic networks. Concepts, algorithms and applications. Cambridge University Press, Cambridge

Kanj IA, Nakhleh L, Than C, Xia G (2008) Seeing the trees and their branches in the network is hard. Theor Comput Sci 401:153–164

Lang AS, Beatty JT (2007) Importance of widespread gene transfer agent genes in alpha-proteobacteria. Trends Microbiol 15:54–62

Lawrence JG, Ochman H (1997) Amelioration of bacterial genomes: rates of change and exchange. J Mol Evol 44:383–397

Lawrence DP, Kroken S, Pryor BM, Arnold AE (2011) Interkingdom gene transfer of a hybrid NPS/PKS from bacteria to filamentous ascomycota. PLoS One e28231. doi:10.1371/journal.pone.0028231

Lindell B, Penny D (2003) The modern molecular clock. Nat Rev Genet 4:216–224. doi:210.1038/nrg1020

Lysak MA, Cheung K, Kitschke M, Bures P (2007) Ancestral chromosomal blocks are triplicated in *Brassicaceae* species with varying chromosome number and genome size. Plant Physiol 145:402–410

Majewski J, Zawadzki P, Pickerill P, Cohan FM, Dowson CG (2000) Barriers to genetic exchange between bacterial species: streptococcus pneumoniae transformation. J Bacteriol 182:1016–1023

Mallet J, Ludovic V, Becq J, Deschavanne P (2010) Whole genome evaluation of horizontal transfers in the pathogenic fungus *Aspergillus fumigatus*. BMC Genomics 11:171. doi:110.1186/1471-2164-1111-1171

Médigue C, Rouxel T, Vigier P, Hénaut A, Danchin A (1991) Evidence of horizontal gene transfer in *Escherichia coli* speciation. J Mol Biol 222:851–856

Nakabachi A, Yamashita A, Toh H, Ishikawa H, Dunbar HE, Moran NA, Hattori M (2006) The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. Science 314:267. doi:210.1126/science.1134196

Nakamura Y, Itoh T, Matsuda H, Gojobori T (2004) Biased biological functions of horizontally transferred genes in prokaryotic genomes. Nat Genet 36:760–766. doi:710.1038/ng1381

Naor A, Lapierre P, Mevarech M, Papke RT, Gophna U (2012) Low species barriers in halophilic Archaea and the formation of recombinant hybrids. Curr Biol 22:1444–1448

Nelson KE, Clayton RA, Gill SR, Gwinn ML, Dodson RJ, Haft DH, Hickey EK et al (1999) Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. Nature 399:323–329. doi:310.1038/20601

Nesb CL, Boucher Y, Doolittle FW (2001) Defining the core of nontransferable prokaryotic genes: the euryarchaeal core. J Mol Evol 53:340–350. doi:310.1007/s002390010224

Novichkov P, Omelchenko M, Gelfand M, Mironov A, Wolf Y, Koonin E (2004) Genome-wide molecular clock and horizontal gene transfer in bacterial evolution. J Bacteriol 186:6575. doi:6510.1128/JB.6186.6519.6575-6585.2004

Ochman H, Lawrence JG, Groisman EA (2000) Lateral gene transfer and the nature of bacterial innovation. Nature 405:299–305

Pamilo P, Nei M (1988) Relationships between gene trees and species trees. Mol Biol Evol 5:568–583

Planet PJ, Kachlany SC, Fine DH, DeSalle R, Figurski DH (2003) The widespread colonization island of *Actinobacillus actinomycetemcomitans*. Nat Genet 34:193–198

Popa O, Dagan T (2011) Trends and barriers to lateral gene transfer in prokaryotes. Curr Opin Microbiol 14:1–9. doi:10.1016/j.mib.2011.1007.1027

Puigbò P, Wolf YI, Koonin EV (2009) Search for a 'Tree of life' in the thicket of the phylogenetic forest. J Biol 8:59

Raymond B, Wyres KL, Sheppard SK, Ellis RJ, Bonsall MB (2010) Environmental factors determining the epidemiology and population genetic structure of the *Bacillus cereus* group in the field. PLoS Pathog 6:e1000905. doi:1000910.1001371/journal.ppat.1000905

Rendulic S, Jagtap P, Rosinus A, Eppinger M, Baar C, Lanz C, Keller H et al (2004) A predator unmasked: life cycle of *Bdellovibrio bacteriovorus* from a genomic perspective. Science 303:689–692. doi:610.1126/science.1093027

Retchless A, Lawrence JG (2010) Phylogenetic incongruence arising from fragmented speciation in enteric bacteria. Proc Natl Acad Sci 107:11453–11458. doi:11410.11073/pnas.1001291107
Rieseberg LH, Carney SE (1998) Plant hybridization. New Phytol 140:599–624
Rokas A, Williams BL, King N, Carroll SB (2003) Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature 425:798–804
Scornavacca C, Linz S, Albrecht B (2012) A first step toward computing all hybridization networks for two rooted binary phylogenetic trees. J Comput Biol 19:1227–1242. doi:1210.1089/cmb.2012.0192
Song YS, Hein J (2005) Constructing minimal ancestral recombination graphs. J Comput Biol 12:147–169
Supek F, Skunca N, Repar J, Vlahovicek K, Smuc T (2010) Translational selection is ubiquitous in prokaryotes. PLoS Genet 6:e1001004. doi:1001010.1001371/journal.pgen.1001004
Teichmann SA, Mitchison G (1999) Is there a phylogenetic signal in prokaryote proteins? J Mol Evol 49:98–107
Thomas CM, Nielsen KM (2005) Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nat Rev Micriobiol 3:711–721
Tofigh A, Hallet M, Lagergren J (2010) Simultaneous identification of duplications and lateral gene transfers. IEEE/ACM Trans Comput Biol Bioinform 8:517–535. doi:10.1109/TCBB.2010.14
Vernikos GS, Thomson NR, Parkhill J (2007) Genetic flux over time in the *Salmonella* lineage. Genome Biol 8:R100. doi:110.1186/gb-2007-1188-1186-r1100
Vulic M, Dionisio F, Taddei F, Radman M (1997) Molecular keys to speciation: DNA polymorphism and the control of genetic exchange in enterobacteria. Proc Natl Acad Sci 94:9763–9767
Wang L, Ma B, Li M (2000) Fixed topology alignment with recombination. Discrete Appl Math 104:281–300
Worning P, Jensen LJ, Nelson KE, Brunak S, Ussery DW (2000) Structural analysis of DNA sequence: evidence for lateral gene transfer in *Thermotoga maritima*. Nucleic Acids Res 28:706–709
Wozniak RAF, Waldor MK (2010) Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. Nat Rev Micriobiol 8:552–563
Zawadzki P, Roberts MS, Cohan FM (1995) The log-linear relationship between sexual isolation and sequence divergence in *Bacillus transformation* is robust. Genetics 140:917–932
Zhandong L, Venkatesh SS, Maley CC (2008) Sequence space coverage, entropy of genomes and the potential to detect non-human DNA in human samples. BMC Genomics 9:509. doi:510.1186/1471-2164-1189-1509
Zhang J (2003) Evolution by gene duplication: an update. Trends Ecol Evol 18:292–298

# Chapter 11
# Horizontal Acquisition of Prokaryotic Genes for Eukaryote Functioning and Niche Adaptation

**Maxime Bruto, Claire Prigent-Combaret, Patricia Luis, Grégory Hoff, Yvan Moënne-Loccoz and Daniel Muller**

**Abstract** Horizontal gene transfer (HGT) is a major mechanism of evolution, in that it is pervasive and can dramatically affect lifestyle by allowing adaptation to specialized niches. Although research has mostly focused on HGT within prokaryotes, examples of inter-domain transfers from prokaryotes to eukaryotes are increasing, and such inter-domain HGT is emerging as a very significant component in ecological and evolutionary terms. Here, different cases of intra- and inter-domain HGT conferring an adaptive advantage to eukaryotes are reviewed to examine novel trends and HGT paradigms. Thus, HGT appears to play an important role in eukaryotic adaptation to specific environmental conditions, including in the ecological evolution toward parasitic lifestyles and pathogenesis. The diversity of prokaryotes and their genetic potential are emerging as a vast reservoir to foster rapid eukaryote evolution.

## 11.1 Horizontal Gene Transfer and Adaptation

In nature, species need to constantly adapt to their changing environment. To maximize survival, microorganisms may acquire new functions, for instance by undergoing genetic changes such as mutations, gene duplications, and genomic

M. Bruto (✉) · C. Prigent-Combaret · P. Luis · G. Hoff · Y. Moënne-Loccoz · D. Muller
Université de Lyon, F-69622 Lyon, France
e-mail: bruto.maxime@gmail.com

D. Muller
e-mail: daniel.muller@univ-lyon1.fr

M. Bruto · C. Prigent-Combaret · P. Luis · G. Hoff · Y. Moënne-Loccoz · D. Muller
Université Lyon 1, Villeurbanne, France

M. Bruto · C. Prigent-Combaret · P. Luis · G. Hoff · Y. Moënne-Loccoz · D. Muller
CNRS, UMR5557, Ecologie Microbienne, Villeurbanne, France

P. Pontarotti (ed.), *Evolutionary Biology: Exobiology and Evolutionary Mechanisms*,
DOI: 10.1007/978-3-642-38212-3_11, 

rearrangements (Ochman et al. 2000; Raz and Tannenbaum 2010; Vial et al. 2006). This can also result from the acquisition of exogenous genes by horizontal gene transfer (HGT), using conjugation, transformation, transduction, or nanotube-based exchange (Dubey and Ben-Yehuda 2011; Pothier et al. 2008). HGT is characterized by the movement of genetic material through different cellular barriers, and takes place between organisms that are more or less related (Dunning Hotopp 2011). It is opposite to vertical transfer, which is the transmission of a gene from a parental organism to the next generation (Keeling and Palmer 2008). In certain cases, HGT may enable the acquisition of very novel genes, which is more drastic than the gradual evolutionary processes that rely on modification of pre-existing genes (Lerat et al. 2005; Ochman et al. 2000; Treangen and Rocha 2011). These novel genes can either replace a pre-existing gene (and the corresponding function) or confer a new function (Keeling and Palmer 2008).

For a long time, the scientific community believed that HGT was limited to the prokaryotic world, because (1) the eukaryotic cell barriers make it difficult to access to the nuclear genome (that includes nucleus organization and genome compaction into chromatin), and (2) in most multicellular eukaryotes the transmission of acquired genes to the offspring is made more difficult because of the separation of somatic and germ cells (limiting the heritability of genes acquired horizontally by somatic cells) (Dunning Hotopp 2011; Keeling 2009). HGT has been extensively studied in prokaryotes to decipher the transfer mechanisms (Dubey and Ben-Yehuda 2011), as well as understand its ecological importance for the acquisition of novel biotransformation, stress resistance, and biotic interaction capabilities (de la Cruz and Davies 2000; Lassalle et al. 2011). Nevertheless, an increasing number of inter-domain HGT events, especially between bacteria and eukaryotes, have been described in a large variety of eukaryotic species (Bock 2010; Dunning Hotopp 2011; Keeling 2009). Furthermore, different HGT patterns (e.g. single versus multiple transfers), with contrasted impacts on eukaryotic evolution, are now documented (Table 11.1).

The purpose of this review will be to assess the ecological significance of the acquisition of prokaryotic genes for eukaryote functioning and niche adaptation. Indeed, for an acquired gene to be maintained over generations, it is likely that it should have in most cases have a neutral or a positive effect on the recipient organism. In this context, the bacterial gene pool can be seen as a huge reservoir of potential functions for eukaryotes (Keeling 2009).

### *11.1.1 The Impact of HGT on Prokaryotic Evolution*

Large fractions of prokaryotic genomes harbor genes that show a codon usage that differs markedly from that of the rest of the genome. This anomalous composition might reflect past gene acquisitions, despite evolution of these genes toward the codon usage of the host (Lawrence and Ochman 1997). Gene gain represents a major source of genetic innovation, as new genes or operons can modify the

**Table 11.1** Examples of prokaryote-to-eukaryote horizontal gene transfers

| Eukaryotic phyla | Example of transferred functions | Possible prokaryotic donors | References |
|---|---|---|---|
| **Kingdom: Animalia or Metazoa** | | | |
| **Phylum Rotifer** | | | |
| *Adineta vaga (bdelloid)* | Oxidoreductase, arabinosidase… (biochemical diversification, new environmental niche) | Cyanobacteria, proteobacteria, bacteroidetes… | Boschetti et al. ( 2012); Gladyshev et al. (2008) |
| **Phylum Nematode** | | | |
| *Tylenchida, Aphelenchoidea, Dorylaimida, Pratylenchidae, Anguinidae, Radopholinae* | Plant cell wall-modifying proteins (adaptation to plant parasitism) | Actinomycetales, Firmicutes, and *Ralstonia* | Danchin (2010); Haegeman et al. (2011) |
| *Pristionchus genus* | Cellulase (adaptation to necromenic lifestyle) | Bacteria | Danchin (2011) |
| *Meloidogyne* species, root-knot nematode | Cellulases, polygalacturonase, and Nod factors (penetration and navigation into plant host; phytoparasitism) | *Rhizobium, Bacillus, Streptomyces, Ralstonia solanacearum*… | Scholl et al. (2003) |
| **Phylum Cnidaria** | | | |
| *Hydra magnipapillata* | Diverse metabolic pathways… | *Burkholderia, Bacillus, Clostridium*… | Chapman et al. (2010) |
| **Phylum Arthropod** | | | |
| *Hypothenemus hampei* (coffee-borer beetle) | Specific mannanase (host adaptation and new environmental niche) | *Bacillus* | Acuña et al. (2012) |
| Lepidopterian | Functions coded by 20 different genes | Bacteria | Sun et al. (2013) |
| *Tetranychus urticae* (spider mite) | Digestion and pesticide detoxification (plant–herbivore ecology) | *Streptomyces* | Grbić et al. (2011); Wybouw et al. (2012) |
| **Phylum Porifera** | | | |
| *Astrosclera willeyana* (sponge) | Spheruline (biomineralization) | Gammaproteobacteria or deltaproteobacteria | Jackson et al. (2011) |
| **Phylum Chordate** | | | |
| *Oikopleura dioica* and *Ciona intestinalis* | Cellulose synthase (role in notochord and tail morphogenesis) | *Streptomyces coelicolor* | Sagane et al. (2010) |
| **Kingdom: Fungus** | | | |
| *Saccharomyces cereviseae* | Fermentation metabolism | *Lactobacillus lactis* | Hall et al. (2005) |
| *Yarrowia lypolitica* | Arsenate reductase (detoxification) | Unknown | Marcet-Houben and Gabaldón (2010) |

(continued)

**Table 11.1** (continued)

| Eukaryotic phyla | Example of transferred functions | Possible prokaryotic donors | References |
|---|---|---|---|
| *Orpinomyces joyonii* | Glycosyl hydrolase (new environmental niche) | *Fibrobacter succinogenes* | Garcia-Vallvé et al. (2000) |
| *Botrytis cinerea* | Catalase (virulence) | *Pseudomonas syringae?* | Marcet-Houben and Gabaldón (2010) |
| *Aspergillus oryzae* | Peptidoglycan metabolism | Firmicutes? | Marcet-Houben and Gabaldón (2010) |
| **Kingdom: Chromalveolata** | | | |
| *Nyctotherus ovalis* | Diverse anaerobic and fermentative pathways (new environmental niche; pathogeny) | Firmicutes, *Bacteroides*… | Keeling and Palmer (2008); Andersson (2009) |
| *Phaeodactylum tricornutum* | Siderophore, metabolic pathway (competitiveness in aquatic environment | Cyanobacteria, proteobacteria, archeae | |
| *Phytophthora* | Cutinase | Actinobacteria | Belbahri et al. (2008) |
| **Kingdom: Plant** | | | |
| *Linaria vulgaris* | Hormone synthesis | *Agrobacterium* | Matveeva et al. (2012) |
| *Physcomitrella patens* | Diverse pathways (transition from aquatic to terrestrial environments) | Firmicutes, proteobacteria | Yue et al. (2012) |
| **Kingdom: *Excavata*** | | | |
| **Kinetoplastida** | | | |
| *Trypanosoma brucei* | | | |
| *Trypanosoma Cruzi* | | | |
| *Leishmania major* | | | |
| **Diplomonades** | | | |
| *Giardia intestinalis, Spironucleus salmonicida* | Diverse anaerobic and fermentative pathways (new environmental niche; pathogeny) | Firmicutes, *Bacteroides*… | Keeling and Palmer (2008); Andersson (2009) |
| **Parabasalia** | | | |
| *Trichomonas vaginalis* | | | |
| **Unranked in kingdom** | | | |
| **Amoebozoa** | | | |
| *Entamoeba histolytica, Acanthamoeba castelanii, Hartmannella vermiformis, Dictyostelium discoideum* | | | |

metabolic, resistance, and interactive capacities of the host organism. These acquisitions can enable the colonization of new habitats, such as arsenic-rich environments (Arsène-Ploetze et al. 2010; Muller et al. 2007), or promote symbiotic interactions (Pothier et al. 2008). HGT can occur at a high rate, as the proportion of genes affected by HGT is estimated to range from 7 to 96 % (Dagan et al. 2008; Kunin and Ouzounis 2003; Zhaxybayeva et al. 2006). For example, it appears that HGT affected nearly half the genome of the plant-associated bacterium *Azospirillum*, enabling a dramatic change in lifestyle from aquatic to terrestrial environments (Wisniewski-Dyé et al. 2011), concomitant with the radiation of land plants.

Although it has been long recognized as a major evolutionary force in prokaryotes (Koonin et al. 2001), HGT is not only adaptive. Indeed, it is thought to compensate for the high rate of gene deletion in prokaryotes (Kuo and Ochman 2009) and the irreversible accumulation of deleterious mutations in these asexual organisms, i.e. Muller's ratchet (Felsenstein 1974). Thus, HGT provides prokaryotes with a recombination tool that would be comparable to sexual recombination in eukaryotes (Muller 1964).

The occurrence of HGT also means that reconstructing bacterial phylogenetic relationships based on genetic marker analysis could lead to misleading tree topologies (Puigbo et al. 2009). In fact, a major question regarding HGT is how abundant it has been in life's evolutionary history and how well tree-like views represent bacterial relationships (Abby et al. 2012; Koonin 2009). It is clear that the importance of HGT in bacteria has opened new questions about the tree of life in bacteria (e.g., Darwinian versus Lamarckian evolution) and the bacterial species concept (Boto 2010; Koonin and Wolf 2009; Puigbo et al. 2009).

### *11.1.2 The Impact of HGT on Eukaryotic Evolution*

While HGT has been extensively studied in bacteria, it has received far less attention in eukaryotes. Gene acquisitions by eukaryotes may be difficult to detect, because (1) when originating from endosymbiotic prokaryotic donors these genes may be considered as bacterial contamination when assessing eukaryotic sequencing data (Dunning Hotopp 2011), and (2) when the donor is an eukaryote then donor–recipient similarity may be high (Keeling 2009). Despite these obstacles, striking examples of HGT were evidenced in the last two decades. Although most of them involved nuclear genes of mitochondrial or chloroplastic origin, i.e., cases of intracellular HGT (Dunning Hotopp 2011; Lang and Beatty 2007), a rising number of HGT events from eukaryote to other eukaryotes are now documented. As for prokaryotes, some eukaryotic HGTs are neutral, the acquired gene replacing a pre-existing homologous gene. This is mostly described for plastid genes, i.e., standard mitochondrial genes encoding ribosomal or respiratory proteins and that were subjected to frequent HGT between distantly related flowering plants (Archibald et al. 2003; Bergthorsson et al. 2003, 2004). However,

other eukaryotic HGT events can have important outcomes in terms of evolutionary and ecological differentiation (Keeling 2009), such as for stress adaptation or parasitic interactions. For example, certain oomycetes became successful plant parasites by multiple acquisitions of fungal genes (Richards et al. 2011; Werner et al. 2002). These transfers mainly conferred the capacity to breakdown plant cell wall and/or acquire nutrients such as sugars. The acquisition of new extracellular enzymes (endoglucanases, endoxylanases…) enabling degradation of plant cell barriers (cellulose, hemicelluloses…) can be advantageous to phytoparasites by facilitating host infection. These transfers of fungal genes are not documented with free living and animal parasitic species, which suggests that they represent specific adaptations to a phytoparasitic lifestyle (Richards et al. 2011). In addition, transfer of plant virulence determinants (like toxin-coding genes) between two phytopathogenic fungi conferred higher virulence to the recipient (Richards et al. 2011; Sanders 2006; Slot and Rokas 2011). Although these acquisitions concern generally one or few genes, certain fungi (i.e., *Alternaria alternate* and *Fusarium* strains) can acquire an entire chromosome from other strains of the same species (Ma et al. 2010), which enables host specialization and synthesis of new virulence factors (Akagi et al. 2009; Coleman et al. 2009). Many genes acquired via interkingdom or inter-domain HGT are thought to promote metabolic diversification in plants (Bock 2010; Richards et al. 2009; Yue et al. 2012), bdelloid rotifers (Boschetti et al. 2012; Gladyshev et al. 2008), and fish (Graham et al. 2008). These acquisitions modified recipient's ecology, leading to parasitic or necromenic behaviors in nematodes (Danchin 2011, 2010) and modified host specificity in insects and spider (Acuña et al. 2012; Grbić et al. 2011; Wybouw et al. 2012).

## 11.2 Inter-Domain HGT (Prokaryote–Eukaryote)

Thanks to an increasing number of sequenced genomes, more and more past HGT events are detected. HGT is invoked to explain phylogenetic results that conflict with our understanding of species relationships. In eukaryotes, the most common observation interpreted in this way is the appearance of a bacterial gene in the nuclear genome. Indeed, it appears that prokaryote-to-eukaryote inter-domain HGT is more prevalent than eukaryote-to-eukaryote HGT (Syvanen 2012).

### *11.2.1 HGT from Endosymbionts to Eukaryotes*

HGT has been extensively studied in the case of mutualistic endosymbiosis (Dunning Hotopp 2011; Nikoh et al. 2008). These transfers are facilitated by the close proximity between the eukaryotic nucleus and the cytoplasmic bacteria (Dunning Hotopp 2011; Nikoh et al. 2008). These studies include HGT from mitochondria and chloroplast (Stegemann and Bock 2009; Stegemann et al. 2003),

which themselves derive from free-living α-proteobacteria and cyanobacteria, respectively, whose endosymbiotic acquisition has represented a genomic acquisition in itself. In the case of mitochondria, many genes have been forwarded from the organelle to the eukaryotic nuclear genome (Nikoh et al. 2008). As for mitochondria, the endosymbiotic bacterium *Wolbachia* is intracellular and inherited vertically from the mother cells. *Wolbachia* infects a wide range of arthropods and filarial nematodes, and various *Wolbachia*-host HGT have been described (Dunning Hotopp et al. 2007; Fenn et al. 2006; Nikoh et al. 2011, 2010; Werren et al. 2010). These HGT concern a few genes or even large portions of the *Wolbachia* genome (Nikoh et al. 2008), i.e., 30–100 % of the entire bacterial chromosome (Dunning Hotopp et al. 2007). Some studies showed that transferred *Wolbachia* genes were expressed in the nucleus (Nikoh et al. 2008).

### *11.2.2 HGT from Free-Living Bacteria to Eukaryotes*

Evidence of HGT from non symbiotic bacteria to eukaryotes is seldom reported. Certain putative HGT transfers, such as those involving the human genome (Lander 2001), turned out to be merely artifacts (Salzberg et al. 2001). For established transfers, a majority of recipients are unicellular organisms, specifically protists (Boschetti et al. 2012; Gladyshev et al. 2008). The other recipients are fungi (Garcia-Vallvé et al. 2000; Hall et al. 2005; Klotz and Loewen 2003; Marcet-Houben and Gabaldón 2010), invertebrate animals (Acuña et al. 2012; Chapman et al. 2010; Danchin 2010; Gladyshev et al. 2008; Haegeman et al. 2011; Jackson et al. 2011; Sagane et al. 2010; Scholl et al. 2003; Wybouw et al. 2012), and plants (Aoki and Syōno 1999; Matveeva et al. 2012; Yue et al. 2012).

### *11.2.3 Bacteria-to-Eukaryotes HGT Can Be Adaptive?*

In prokaryotes, the role of HGT to promote adaptation is established (Ochman and Moran 2001; Popa et al. 2011; Prigent-Combaret et al. 2008). However, the adaptive significance of eukaryotic HGT requires further research attention in most cases, particularly in fungi and insects. So far, adaptive HGT from bacteria to multicellular eukaryotes is documented with the transfer of glycosyl hydrolase genes from cow-rumen prokaryotes to rumen fungi (Garcia-Vallvé et al. 2000), of cutinase gene (Cut family) from actinobacteria to *Phytophthora* oomycetes (Belbahri et al. 2008), of a complete aerolysin gene family from diverse bacterial phyla to distinct eukaryotes among which fungi, animals, and plants (Moran et al. 2012), and of a mannanase-encoding gene from the gut firmicute *Bacillus* to the coffee borer beetle insect *Hypothenemus hampei* (Acuña et al. 2012). In the latter case, the acquired *HhMAN1* gene enables hydrolysis of galactomannan, the most important polysaccharide source in coffee berry and a potential source of nutrient

for the insect, providing a competitive advantage compared to related species such as *Hypothenemus obscurus* (Acuña et al. 2012). Genes coding for plant cell wall-degrading enzymes in phytoparasitic nematodes and acquired from various prokaryotes are fixed in different nematode species, pointing to a key adaptive role (Danchin 2010).

### *11.2.4 Eukaryote-to-Bacteria HGT*

Cases involving eukaryotes-to-prokaryotes HGT are poorly documented. One of the few examples is the acquisition of eukaryotic genes by the endocellular bacterial pathogen *Legionella*. The acquired property (sphingosine-1-phosphate lyase) prevents the phagosome recognition by the lysozymes, allowing *Legionella* to escape the endocytic pathway (Lurie-Weinberger et al. 2010). It has been suggested that HGT might be important for the adaptation of intestinal bacteria (Xu et al. 2007). Indeed, transfers of eukaryotic metabolism genes to *Escherichia coli* (Daniels et al. 1990) and *Bacteroides* (Xu et al. 2007) have improved the ability of these bacteria to sense their gut environment and exploit the nutrients present (Xu et al. 2007). For other acquisitions of eukaryotic genes by gut bacteria, such as the ability to synthesize ubiquitin (in *Bacteroides*; Patrick and Blakely 2012) or the nuclear element L1 (in *Neisseria gonorrhoeae;* Anderson and Seifert 2011), the adaptive significance remains to be clarified. In the case of plant–microbe interactions, the antibiotic biosynthetic *phlACBDE* operon was formed in *Pseudomonas* through the acquisition of *phlACB* from Archaea and *phlD* from an unclear origin, possibly *Streptomyces*, *Vibrio* (Kidarsa et al. 2011), or even plants (Cook et al. 1995; Ramette et al. 2001; Frapolli et al. 2012).

### *11.2.5 Mechanisms of Inter-Domain HGT*

While transformation, transduction, conjugation, and nanotube-mediated mechanisms promote horizontal transfer of genetic materials across bacteria (Dubey and Ben-Yehuda 2011; Syvanen 2012), several mechanisms of gene transfer from bacteria to eukaryotes are also recognized. Bacteriovorous eukaryotes underwent the highest number of HGT events (Andersson 2009; Boschetti et al. 2012; Keeling 2009), perhaps because DNA from their prey may be released near the nucleus, which gave new meaning to the saying ‘You are what you eat’ (Doolittle 1998), but the precise mechanism is not fully established. The use of type-IV secretion systems to transfer DNA into eukaryotic cells by *Agrobacterium tumefaciens* (Kunik et al. 2001), or *Bartonella henselae* (Schröder et al. 2011) are the most studied inter-domain mechanisms. Viruses can also serve to promote gene exchanges between organisms (Filée et al. 2007). Recently discovery nuclear localization signals in bacteriophages allow import of viral proteins across the

nuclear barrier and promote effective DNA transfer from bacteria to the eukaryotic nucleus (Redrejo-Rodríguez et al. 2012). Large viruses e.g., Poxviruses are thought to have undergone multiple host changes (Hughes et al. 2010). Since they may harbor many bacterial and eukaryotic genes (Filée et al. 2007; Hughes and Friedman 2005), it means that these transfers are likely to have promoted gene transfer possibilities between distant taxa. The example of chloroplast transfer between plants, with subsequent plastid transfer to the new host genome, illustrates how sets of genes might be transferred in perhaps a lineage-specific way among eukaryotes (Stegemann and Bock 2009; Stegemann et al. 2003), but here again the transfer mechanisms involved deserve further research attention.

## 11.3 Serial Inter-Domain HGT

While it is clear that HGT has affected eukaryotic nuclear genomes (Andersson 2009; Boto 2010), the frequency of such events and the extent of their evolutionary impact have long remained unclear. Recent surveys show that gene acquisition from another domain was extensive in some archaeal and eukaryotic species.

### *11.3.1 Speciation Can Arise from Unique or Multiple Inter-Domain HGT(s)*

As described for the *H. hampei* (Acuña et al. 2012), inter-domain HGT of a single gene can lead to ecological specialization of the recipient. However, inter-domain HGT can involve more than one gene or function. Thus, one of the most dramatic inter-domain HGT was the acquisition in a putatively single HGT event of about 1,000 bacterial genes by a methanogenic ancestor of the Haloarchaea (Nelson 1999; Nelson-Sathi et al. 2012). However, cases of unique large inter-domain HGT such as those that gave rise to Haloarchaea or to the primitive eukaryotes (Nelson-Sathi et al. 2012) must be rare. Indeed, most studies showed that inter-domain HGT events occurred repeatedly, meaning that the genesis of new phyla may take place in several stages, following multiple acquisitions.

In the multicellular bdelloid rotifer *Adineta vaga*, the subtelomeric chromosomic regions are enriched in genes of foreign origin (bacteria, fungi, or plants). A quantitative analysis of HGT in the closely related species *Adineta ricciae* concluded that 10 % of genes had been horizontally acquired, and more importantly they were expressed (Boschetti et al. 2012). Most of these genes code for enzymes implicated in diverse metabolic pathways, including degradation of toxins or synthesis of antioxidants and rare metabolites. It is argued that bdelloid rotifers are prone to acquire foreign genes because they reproduce by parthenogenesis (Boschetti et al. 2012). However, sexual organisms with segregation of germlines present also multiple inter-domain transfers. For example Danchin et al. (2010) showed that multiple HGT promoted plant parasitism in nematodes. Among

arthropod genomes, traces of multiple inter-domain transfers were also found (Grbić et al. 2011; Li et al. 2011; Sormacheva et al. 2012). The genome of the spider mite *Tetranychus urticae*, a pesticide resistant agricultural pest presenting a large host range, harbors genes related to digestion and pesticide detoxification that were acquired from bacteria and fungi (Grbić et al. 2011). Similarly, in a survey of lepidopterian insect genomes found around 20 different transfers from enterobacteria, which includes endosymbionts (Sun et al. 2013). However, the most important impact of cumulative inter-domain HGT might be the transition of plants from aquatic to terrestrial environments (Yue et al. 2012), as essential plant activities such as xylem formation, plant defense, nitrogen recycling, and starch biosynthesis may involve genes of bacterial origin (Yue et al. 2012).

### *11.3.2 Dynamics of HGT*

Identifying the distribution of a transferred gene necessitates of a transferred gene is to rely on the availability of genome sequences. Thus, apparently simple cases of HGT might be more complex, for example if a bacterial gene found in a single eukaryote is also present in other, distantly related eukaryotic groups that are not sequenced (Keeling and Inagaki 2004). This punctuate gene distribution may result from paralogy or occasional maintenance in the descent once the gene was acquired by the eukaryotic ancestor. Alternatively, these bacterial genes could have been transferred to a single eukaryote, which subsequently transferred them to distantly related eukaryotes (Rogers 2007). In 2010, an exhaustive analysis of all fungal genome sequences available identified 733 HGT events from bacteria to fungi (Marcet–Houben and Gabaldón 2010). From this study, different types of transfer dynamics can be suggested. One possibility is that after a first bacterial transfer, subsequent transfer between eukaryotes occurred. A catalase-encoding gammaproteobacterial gene was evidenced in three phytopathogenic fungi, i.e., the Dothideomycetes *Stagonospora nodorum* and *Mycosphaerella fijiensis* and the Leotiomycete *Botrytis cinerea*, and it was suggested that *B. cinerea* acquired the bacterial gene from one of the Dothideomycetes (Marcet–Houben and Gabaldón 2010). Another possibility is that different eukaryotes from a same phylogenetic clade independently acquired a same gene from different bacteria. For example, a bacterial gene coding for an arsenate reductase (an important enzyme to detoxify arsenic), underwent two independent acquisitions by the fungal species *Yarrowia lipolytica* and *Rhizopus oryzae* (Marcet–Houben and Gabaldón 2010).

## 11.4 Conclusion

HGT is a widespread mechanism for acquisition of genetic variability that has modulated the evolution not only of prokaryotic domains (bacteria and archaea) but also that of eukaryotes. Although HGT importance in prokaryote evolution is

well recognized, the importance of HGT for eukaryote evolution is an emerging issue. The lack of systematic HGT analysis on newly sequenced eukaryotic genomes does not enable a comprehensive understanding of the impact of HGTs across very large taxonomic distances. The only exhaustive analysis was done on fungi, and showed that every clade was affected by inter-domain HGT (Marcet–Houben and Gabaldón 2010).

It is difficult to determine if the prevalence of bacteria as the main source of new genes in eukaryotes is due to technical biases or a biological reality. Bacteria are widespread in virtually all types of habitats and associate readily with all kinds of eukaryotes, providing ample opportunities for prokaryote–eukaryote physical interactions and eukaryotic acquisition of new genes from the bacterial pool. Most acquired genes contributed to eukaryotic adaptation, albeit to different extents. The amazing metabolic diversity of bacteria, which made them ubiquitous on earth, could also explain why bacteria are the main source of acquired genes in eukaryotes. On this basis, eukaryotic evolution needs a reappraisal by considering both vertical and HGT events, reminiscent of our current understanding of prokaryotic evolution. This is likely to highlight a more convergent picture of the evolutionary processes underpinning diversification and ecological specialization on earth.

## References

Abby SS, Tannier E, Gouy M, Daubin V (2012) Lateral gene transfer as a support for the tree of life. Proc Natl Acad Sci USA 109:4962–4967

Acuña R, Padilla BE, Flórez-Ramos CP, Rubio JD, Herrera JC, Benavides P, Lee S-J, Yeats TH, Egan AN, Doyle JJ et al (2012) Adaptive horizontal transfer of a bacterial gene to an invasive insect pest of coffee. Proc Natl Acad Sci USA 109:4197–4202

Akagi Y, Akamatsu H, Otani H, Kodama M (2009) Horizontal chromosome transfer, a mechanism for the evolution and differentiation of a plant-pathogenic fungus. Eukaryot Cell 8:1732–1738

Anderson MT, Seifert HS (2011) Opportunity and means: horizontal gene transfer from the human host to a bacterial pathogen. MBio 2:e00005–e00011

Andersson JO (2009) Gene transfer and diversification of microbial eukaryotes. Annu Rev Microbiol 63:177–193

Aoki S, Syōno K (1999) Horizontal gene transfer and mutation: *Ngrol* genes in the genome of *Nicotiana glauca*. Proc Natl Acad Sci USA 96:13229–13234

Archibald J, Rogers M, Toop M, Ishida K, Keeling P (2003) Lateral gene transfer and the evolution of plastid-targeted proteins in the secondary plastid-containing alga *Bigelowiella natans*. Proc Natl Acad Sci USA 100:7678–7683

Arsène-Ploetze F, Koechler S, Marchal M, Coppée J-Y, Chandler M et al (2010) Structure, function, and evolution of the *Thiomonas* spp. genome. PLoS Genet 6:e1000859

Belbahri L, Calmin G, Mauch F, Andersson JO (2008) Evolution of the cutinase gene family: evidence for lateral gene transfer of a candidate *Phytophthora* virulence factor. Gene 408:1–8

Bergthorsson U, Adams KL, Thomason B, Palmer JD (2003) Widespread horizontal transfer of mitochondrial genes in flowering plants. Nature 424:197–201

Bergthorsson U, Richardson A, Young G, Goertzen L, Palmer J (2004) Massive horizontal transfer of mitochondrial genes from diverse land plant donors to the basal angiosperm *Amborella*. Proc Natl Acad Sci USA 101:17747–17752

Bock R (2010) The give-and-take of DNA: horizontal gene transfer in plants. Trends Plant Sci 15:11–22

Boschetti C, Carr A, Crisp A, Eyres I, Wang-Koh Y, Lubzens E, Barraclough TG, Micklem G, Tunnacliffe A (2012) Biochemical diversification through foreign gene expression in bdelloid rotifers. PLoS Genet 8:e1003035

Boto L (2010) Horizontal gene transfer in evolution: facts and challenges. Proc R Soc B 277:819–827

Chapman JA, Kirkness EF, Simakov O, Hampson SE, Mitros T, Weinmaier T, Rattei T, Balasubramanian PG, Borman J, Busam D et al (2010) The dynamic genome of Hydra. Nature 464:592–596

Coleman JJ, Rounsley SD, Rodriguez-Carres M, Kuo A, Wasmann CC, Grimwood J, Schmutz J, Taga M, White GJ, Zhou S et al (2009) The genome of *Nectria haematococca*: contribution of supernumerary chromosomes to gene expansion. PLoS Genet 5:e1000618

Cook RJ, Thomashow LS, Weller DM, Fujimoto D, Mazzola M, Bangera G, Kim D (1995) Molecular mechanisms of defense by rhizobacteria against root disease. Proc Natl Acad Sci USA 92:4197–4201

Dagan T, Artzy-Randrup Y, Martin W (2008) Modular networks and cumulative impact of lateral transfer in prokaryote genome evolution. Proc Natl Acad Sci USA 105:10039–10044

Danchin EG (2010) Multiple lateral gene transfers and duplications have promoted plant parasitism ability in nematodes. Proc Natl Acad Sci USA 107:17651–17656

Danchin EG (2011) What nematode genomes tell us about the importance of horizontal gene transfers in the evolutionary history of animals. Mob Genet Elem 1: 269–292

Daniels SB, Peterson KR, Strausbaugh LD, Kidwell MG, Chovnick A (1990) Evidence for horizontal transmission of the P transposable element between *Drosophila* species. Genetics 124:339–355

de la Cruz F, Davies J (2000) Horizontal gene transfer and the origin of species: lessons from bacteria. Trends Microbiol 8:128–133

Doolittle W (1998) You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. Trends Genet 14:307–311

Dubey GP, Ben-Yehuda S (2011) Intercellular nanotubes mediate bacterial communication. Cell 144:590–600

Dunning Hotopp JC (2011) Horizontal gene transfer between bacteria and animals. Trends Genet 27:157–163

Dunning Hotopp JC, Clark ME, Oliveira DCSG, Foster JM, Fischer P, Muñoz Torres MC, Giebel JD, Kumar N, Ishmael N, Wang S et al (2007) Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. Science 317:1753–1756

Felsenstein J (1974) The evolutionary advantage of recombination. Genetics 78:737–756

Fenn K, Conlon C, Jones M, Quail MA, Holroyd NE, Parkhill J, Blaxter M (2006) Phylogenetic relationships of the *Wolbachia* of nematodes and arthropods. PLoS Pathog 2:e94

Filée J, Siguier P, Chandler M (2007) I am what I eat and I eat what I am: acquisition of bacterial genes by giant viruses. Trends Genet 23:10–15

Frapolli M, Pothier JF, Défago G, Moënne-Loccoz Y (2012) Evolutionary history of synthesis pathway genes for phloroglucinol and cyanide antimicrobials in plant-associated fluorescent pseudomonads. Mol Phylogenet Evol 63:877–890

Garcia-Vallvé S, Romeu A, Palau J (2000) Horizontal gene transfer of glycosyl hydrolases of the rumen fungi. Mol Biol Evol 17:352–361

Gladyshev EA, Meselson M, Arkhipova IR (2008) Massive horizontal gene transfer in bdelloid rotifers. Science 320:1210–1213

Graham LA, Lougheed SC, Ewart KV, Davies PL (2008) Lateral transfer of a lectin-like antifreeze protein gene in fishes. PLoS ONE 3:e2616

Grbić M, Van Leeuwen T, Clark RM, Rombauts S, Rouzé P, Grbić V, Osborne EJ, Dermauw W, Thi Ngoc PC, Ortego F et al (2011) The genome of *Tetranychus urticae* reveals herbivorous pest adaptations. Nature 479:487–492

Haegeman A, Jones JT, Danchin EGJ (2011) Horizontal gene transfer in nematodes: a catalyst for plant parasitism? Mol Plant-Microbe Interact 24:879–887

Hall C, Brachat S, Dietrich FS (2005) Contribution of horizontal gene transfer to the evolution of *Saccharomyces cerevisiae*. Eukaryot Cell 4:1102–1115

Hughes AL, Friedman R (2005) Poxvirus genome evolution by gene gain and loss. Mol Phylogenet Evol 35:186–195

Hughes AL, Irausquin S, Friedman R (2010) The evolutionary biology of poxviruses. Infect Genet Evol 10:50–59

Jackson D, Macis L, Reitner J, Worheide G (2011) A horizontal gene transfer supported the evolution of an early metazoan biomineralization strategy. BMC Evol Biol 11:238

Keeling P, Palmer J (2008) Horizontal gene transfer in eukaryotic evolution. Nat Rev Genet 9:605–618

Keeling PJ (2009) Functional and ecological impacts of horizontal gene transfer in eukaryotes. Curr Opin Genet Dev 19:613–619

Keeling PJ, Inagaki Y (2004) A class of eukaryotic GTPase with a punctate distribution suggesting multiple functional replacements of translation elongation factor 1alpha. Proc Natl Acad Sci USA 101:15380–15385

Kidarsa TA, Goebel NC, Zabriskie TM, Loper JE (2011) Phloroglucinol mediates cross-talk between the pyoluteorin and 2,4-diacetylphloroglucinol biosynthetic pathways in *Pseudomonas fluorescens* Pf-5. Mol Microbiol 81:395–414

Klotz MG, Loewen PC (2003) The molecular evolution of catalatic hydroperoxidases: evidence for multiple lateral transfer of genes between prokaryota and from bacteria into eukaryota. Mol Biol Evol 20:1098–1112

Koonin E (2009) Darwinian evolution in the light of genomics. Nucleic Acids Res 37:1011–1034

Koonin E, Makarova K, Aravind L (2001) Horizontal gene transfer in prokaryotes: quantification and classification. Annu Rev Microbiol 55:709–742

Koonin E, Wolf Y (2009) Is evolution Darwinian or/and Lamarckian? Biol Direct 4:42

Kunik T, Tzfira T, Kapulnik Y, Gafni Y, Dingwall C, Citovsky V (2001) Genetic transformation of HeLa cells by *Agrobacterium*. Proc Natl Acad Sci USA 98:1871–1876

Kunin V, Ouzounis CA (2003) The balance of driving forces during genome evolution in prokaryotes. Genome Res 13:1589–1594

Kuo C-H, Ochman H (2009) The fate of new bacterial genes. FEMS Microbiol Rev 33:38–43

Lander ES (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921

Lang A, Beatty J (2007) Importance of widespread gene transfer agent genes in alpha-proteobacteria. Trends Microbiol 15:54–62

Lassalle F, Campillo T, Vial L, Baude J, Costechareyre D, Chapulliot D, Shams M, Abrouk D, Lavire C, Oger-Desfeux C et al (2011) Genomic species are ecological species as revealed by comparative genomics in *Agrobacterium tumefaciens*. Genome Biol Evol 3:762–781

Lawrence JG, Ochman H (1997) Amelioration of bacterial genomes: rates of change and exchange. J Mol Evol 44:383–397

Lerat E, Daubin V, Ochman H, Moran NA (2005) Evolutionary origins of genomic repertoires in bacteria. PLoS Biol 3:e130

Li Z-W, Shen Y-H, Xiang Z-H, Zhang Z (2011) Pathogen-origin horizontally transferred genes contribute to the evolution of Lepidopteran insects. BMC Evol Biol 11:356

Lurie-Weinberger MN, Gomez-Valero L, Merault N, Glöckner G, Buchrieser C, Gophna U (2010) The origins of eukaryotic-like proteins in *Legionella pneumophila*. Int J Med Microbiol 300:470–481

Ma L-J, van der Does HC, Borkovich KA, Coleman JJ, Daboussi M-J, Di Pietro A, Dufresne M, Freitag M, Grabherr M, Henrissat B et al (2010) Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. Nature 464:367–373

Marcet-Houben M, Gabaldón T (2010) Acquisition of prokaryotic genes by fungal genomes. Trends Genet 26:5–8

Matveeva TV, Bogomaz DI, Pavlova OA, Nester EW, Lutova LA (2012) Horizontal gene transfer from genus *Agrobacterium* to the plant *Linaria* in Nature. Mol Plant-Microbe Interact 25:1542–1551

Moran Y, Fredman D, Szczesny P, Grynberg M, Technau U (2012) Recurrent horizontal transfer of bacterial toxin genes to eukaryotes. Mol Biol Evol 29:2223–2230

Muller D, Médigue C, Koechler S, Barbe V, Barakat M, Talla E, Bonnefoy V, Krin E, Arsène-Ploetze F, Carapito C et al (2007) A tale of two oxidation states: bacterial colonization of arsenic-rich environments. PLoS Genet 3:e53

Muller HJ (1964) The relation of recombination to mutational advance. Mutat Res 106:2–9

Nelson KE (1999) Evidence for lateral gene transfer between archaea and bacteria from genome sequence of *Thermotoga maritima*. Nature 399:323–329

Nelson-Sathi S, Dagan T, Landan G, Janssen A, Steel M, McInerney JO, Deppenmeier U, Martin WF (2012) Acquisition of 1,000 eubacterial genes physiologically transformed a methanogen at the origin of *Haloarchaea*. Proc Natl Acad Sci USA 109:20537–20542

Nikoh N, Hosokawa T, Oshima K, Hattori M, Fukatsu T (2011) Reductive evolution of bacterial genome in insect gut environment. Genome Biol Evol 3:702–714

Nikoh N, McCutcheon JP, Kudo T, Miyagishima SY, Moran NA, Nakabachi A (2010) Bacterial genes in the Aphid genome: absence of functional gene transfer from *Buchnera* to its host. PLoS Genet 6:e1000827

Nikoh N, Tanaka K, Shibata F, Kondo N, Hizume M, Shimada M, Fukatsu T (2008) *Wolbachia* genome integrated in an insect chromosome: evolution and fate of laterally transferred endosymbiont genes. Genome Res 18:272–280

Ochman H, Lawrence J, Groisman E (2000) Lateral gene transfer and the nature of bacterial innovation. Nature 405:299–304

Ochman H, Moran NA (2001) Genes lost and genes found: evolution of bacterial pathogenesis and symbiosis. Science 292:1096–1099

Patrick S, Blakely GW (2012) Crossing the eukaryote-prokaryote divide: a ubiquitin homolog in the human commensal bacterium *Bacteroides fragilis*. Mob Genet Elem 2:149–151

Popa O, Hazkani-Covo E, Landan G, Martin W, Dagan T (2011) Directed networks reveal genomic barriers and DNA repair bypasses to lateral gene transfer among prokaryotes. Genome Res 21:599–609

Pothier JF, Prigent-Combaret C, Haurat J, Moënne-Loccoz Y, Wisniewski-Dyé F (2008) Duplication of plasmid-borne nitrite reductase gene *nirK* in the wheat-associated plant growth-promoting rhizobacterium *Azospirillum brasilense* Sp245. Mol Plant-Microbe Interact 21:831–842

Prigent-Combaret C, Blaha D, Pothier J, Vial L, Poirier M, Wisniewski-Dyé F, Moënne-Loccoz Y (2008) Physical organization and phylogenetic analysis of *acdR* as leucine-responsive regulator of the 1-aminocyclopropane-1-carboxylate deaminase gene *acdS* in phytobeneficial *Azospirillum lipoferum* 4B and other Proteobacteria. FEMS Microbiol Ecol 65:202–219

Puigbo P, Wolf Y, Koonin E (2009) Search for a tree of life in the thicket of the phylogenetic forest. J Biol 8:59

Ramette A, Moënne-Loccoz Y, Défago G (2001) Polymorphism of the polyketide synthase gene phlD in biocontrol fluorescent pseudomonads producing 2,4-diacetylphloroglucinol and comparison of *PhlD* with plant polyketide synthases. Mol Plant-Microbe Interact 14:639–652

Raz Y, Tannenbaum E (2010) The influence of horizontal gene transfer on the mean fitness of unicellular populations in static environments. Genetics 185:327–337

Redrejo-Rodríguez M, Muñoz-Espín D, Holguera I, Mencía M, Salas M (2012) Functional eukaryotic nuclear localization signals are widespread in terminal proteins of bacteriophages. Proc Natl Acad Sci USA 109:18482–18487

Richards TA, Soanes DM, Foster PG, Leonard G, Thornton CR, Talbot NJ (2009) Phylogenomic analysis demonstrates a pattern of rare and ancient horizontal gene transfer between plants and fungi. The Plant Cell Online 21: 1897–1911

Richards TA, Soanes DM, Jones MDM, Vasieva O, Leonard G, Paszkiewicz K, Foster PG, Hall N, Talbot NJ (2011) Horizontal gene transfer facilitated the evolution of plant parasitic mechanisms in the oomycetes. Proc Natl Acad Sci USA 108:15258–15263

Rogers MB (2007) A complex and punctate distribution of three eukaryotic genes derived by lateral gene transfer. BMC Evol Biol 7:89

Sagane Y, Zech K, Bouquet J-M, Schmid M, Bal U, Thompson EM (2010) Functional specialization of cellulose synthase genes of prokaryotic origin in chordate larvaceans. Development 137:1483–1492

Salzberg SL, White O, Peterson J, Eisen JA (2001) Microbial genes in the human genome: lateral transfer or gene loss? Science 292:1903–1906

Sanders IR (2006) Rapid disease emergence through horizontal gene transfer between eukaryotes. Trends Ecol Evol 21:656–658

Scholl E, Thorne J, McCarter J, Bird DM (2003) Horizontally transferred genes in plant-parasitic nematodes: a high-throughput genomic approach. Genome Biol 4:R39

Schröder G, Schuelein R, Quebatte M, Dehio C (2011) Conjugative DNA transfer into human cells by the VirB/VirD4 type IV secretion system of the bacterial pathogen *Bartonella henselae*. Proc Natl Acad Sci USA 108:14643–14648

Slot JC, Rokas A (2011) Horizontal transfer of a large and highly toxic secondary metabolic gene cluster between Fungi. Curr Biol 21:134–139

Sormacheva I, Smyshlyaev G, Mayorov V, Blinov A, Novikov A, Novikova O (2012) Vertical evolution and horizontal transfer of CR1 Non-LTR Retrotransposons and Tc1/mariner DNA transposons in *Lepidoptera* species. Mol Biol Evol 29:3685–3702

Stegemann S, Bock R (2009) Exchange of genetic material between cells in plant tissue grafts. Science 324:649–651

Stegemann S, Hartmann S, Ruf S, Bock R (2003) High-frequency gene transfer from the chloroplast genome to the nucleus. Proc Natl Acad Sci USA 100:8828–8833

Sun BF, Xiao JH, He SM, Liu L, Murphy RW, Huang DW (2013) Multiple ancient horizontal gene transfers and duplications in lepidopteran species. Insect Mol Biol 22:72–87

Syvanen M (2012) Evolutionary implications of horizontal gene transfer. Annu Rev Genet 46:341–358

Treangen TJ, Rocha EPC (2011) Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes. PLoS Genet 7:e1001284

Vial L, Lavire C, Mavingui P, Blaha D, Haurat J, Moënne-Loccoz Y, Bally R, Wisniewski-Dyé F (2006) Phase variation and genomic architecture changes in *Azospirillum*. J Bacteriol 188:5364–5373

Werren JH, Richards S, Desjardins CA, Niehuis O, Gadau J, Colbourne JK, The *Nasonia* Genome Working Group (2010). Functional and evolutionary insights from the genomes of three parasitoid *Nasonia* species. Science 327:343–348

Werner S, Steiner U, Bechner R, Kortekamp A, Zyprian E, Deising HB (2002) Chitin synthesis during in planta growth and asexual propagation of the cellulosic oomycete and obligate biotrophic grapevine pathogen plasmopara viticola. FEMS Microbiol Letters 208: 169–173

Wisniewski-Dyé F, Borziak K, Khalsa-Moyers G, Alexandre G, Sukharnikov LO, Wuichet K, Hurst GB, McDonald WH, Robertson JS, Barbe V et al (2011) *Azospirillum* genomes reveal transition of bacteria from aquatic to terrestrial environments. PLoS Genet 7:e1002430

Wybouw N, Balabanidou V, Ballhorn DJ, Dermauw W, Grbić M, Vontas J, Van Leeuwen T (2012) A horizontally transferred cyanase gene in the spider mite *Tetranychus urticae* is involved in cyanate metabolism and is differentially expressed upon host plant change. Insect Biochem Mol Biol 42:881–889

Xu J, Mahowald MA, Ley RE, Lozupone CA, Hamady M, Martens EC, Henrissat B, Coutinho PM, Minx P, Latreille P et al (2007) Evolution of symbiotic bacteria in the distal human intestine. PLoS Biol 5:e156

Yue J, Hu X, Sun H, Yang Y, Huang J (2012) Widespread impact of horizontal gene transfer on plant colonization of land. Nat Commun 3:1152

Zhaxybayeva O, Gogarten JP, Charlebois RL, Doolittle WF, Papke RT (2006) Phylogenetic analyses of cyanobacterial genomes: quantification of horizontal gene transfer events. Genome Res 16:1099–1108

# Chapter 12
# Molecular Evolution of Disrupted Transfer RNA Genes and Their Introns in Archaea

**Akio Kanai**

**Abstract** Transfer RNA (tRNA) is one of the classical non-coding RNAs and is essential for decoding the genomic sequence into proteins. The disrupted tRNA genes in the genomes of the Archaea have been studied in recent years and three unique examples have been found: multiple-intron-containing tRNAs, split tRNAs, and permuted tRNAs. During this research, it was noted that frequent intron transpositions might have occurred among the tRNA genes in the Archaea. Because these tRNAs are encoded as precursor forms (pre-tRNAs) in the genome, they must be processed to yield mature functional tRNAs. Here, the co-evolutionary history of the tRNA gene architecture and the tRNA splicing enzymes in the domain Archaea are proposed. In this review, I discuss a possible evolutionary scenario for the disrupted tRNAs and their introns.

## 12.1 Introduction

Transfer RNA is a short non-coding RNA of approximately 70–100 bases. The principal function of tRNA is to contribute to the molecular mechanism of translation. Each tRNA is charged with the proper amino acid and the tRNAs with attached amino acids are delivered to the ribosome during protein synthesis. tRNA also acts as a primer for DNA synthesis in certain viruses (Lund et al. 2000; Mak and Kleiman 1997). Thus, tRNA is involved in every important step described in the central dogma of molecular biology.

In some cases, tRNA genes are interrupted by an intronic sequence. Many of these introns are found at the position adjacent to the anticodon, between nucleotides 37 and 38 of the precursor tRNA (37/38), known as the "canonical"

A. Kanai (✉)
Institute for Advanced Biosciences, Keio University, 403-1 Nipponkoku, Daihoji, Tsuruoka, Yamagata 997-0017, Japan
e-mail: akio@sfc.keio.ac.jp

P. Pontarotti (ed.), *Evolutionary Biology: Exobiology and Evolutionary Mechanisms*,
DOI: 10.1007/978-3-642-38212-3_12, 

position. Figure 12.1 shows the different types of tRNA introns on a phylogenetic tree of the three domains of life, the Bacteria, Eukarya, and Archaea (Belfort and Weiner 1997). All these examples of tRNAs possess a single intron in the anticodon loop region. A self-splicing type of catalytic introns, including the group I and group II introns, occurs in the Bacteria (Saldanha et al. 1993; Tanner and Cech 1996). These are ribozymes. However, a specific enzyme, called a tRNA splicing endonuclease, is required for the removal of tRNA introns in the Eukarya and Archaea (Tocchini-Valentini et al. 2005b; Trotta et al. 1997). In these cases, the tRNA intron forms a specific RNA secondary structure, the bulge–helix–bulge (BHB), with parts of the exon sequences (Marck and Grosjean 2003), and this structure is the target of the tRNA splicing endonuclease. BHB splicing motifs are further classified by their structures into the strict form (hBHBh′) and the relaxed forms (HBh′ and hBH) (Fig. 12.2). Because the positions of these tRNA introns are highly conserved in the three domains of life, the introns located in the anticodon loop region seem to be evolutionarily ancient and may have existed at the beginning of life.

In recent years, several types of disrupted tRNA genes have been reported in the Archaea and primitive Eukarya, including multiple-intron-containing tRNA genes (Sugahara et al. 2006, 2008), split tRNA genes (Chan et al. 2011; Fujishima et al. 2009; Randau et al. 2005b), and permuted tRNA genes (Chan et al. 2011; Maruyama et al. 2010; Soma et al. 2007). Figure 12.3 summarizes the disrupted tRNAs, most of which have been reported in the Archaea. The genes encoding

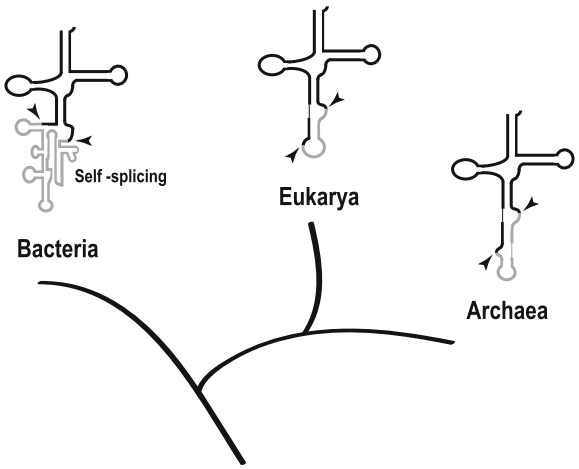



**Fig. 12.1** The three domains in life and their types of tRNA introns. A phylogenetic tree representing the separation of the Bacteria, Eukarya, and Archaea. Many organisms belonging to each domain possess intron-containing tRNAs and most introns are located in the anticodon loop region of the tRNAs. Note that the secondary structures of the eukaryotic and archaeal tRNA introns are very similar and that the introns of both are processed by a specific enzyme, called a "tRNA splicing endonuclease". However, the bacterial tRNA intron is a self-splicing type intron, such as the group I and group II introns. The two arrowheads indicate the exon–intron boundaries

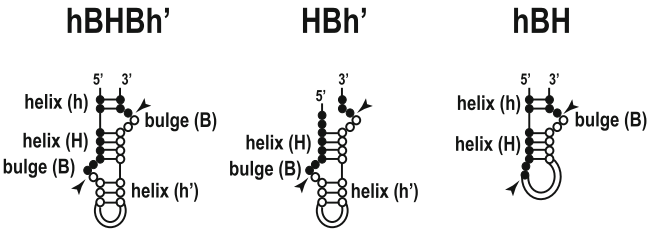



**Fig. 12.2** Three types of tRNA splicing motifs. The tRNA splicing motifs are defined based on a previous study (Marck and Grosjean 2003): the strict hBHBh' motif (*left*), the relaxed HBh' motif (*middle*), and the relaxed hBH motif (*right*). The *black circles* represent exons and the *white circles* represent introns. The two *arrowheads* indicate the exon–intron boundaries cleaved by splicing endonuclease

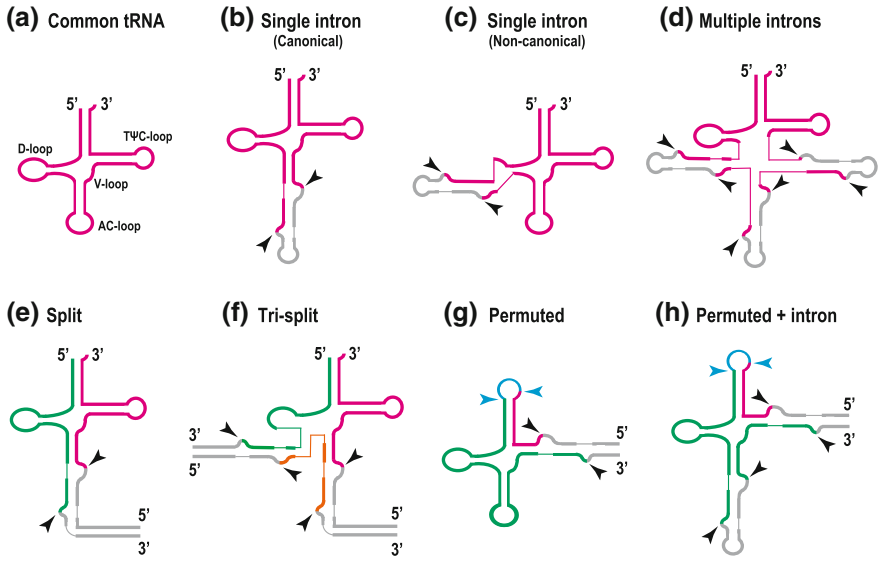



**Fig. 12.3** Molecular structures of disrupted tRNAs. **a** Common tRNA with no disruption. A description of each stem and loop is included. **b** tRNA containing a single intron at the canonical position, 37/38. The exons are indicated in *red* and an intron is indicated in *gray*. **c** Example of a tRNA containing a single intron at a non-canonical position. **d** Example of a tRNA containing multiple introns (up to three introns) at various positions. **e** Split tRNA, in which the tRNA is composed of two individual transcripts (transcript 1 with a *green* exon and transcript 2 with a *red* exon; leader sequences are indicated in *gray*). **f** Tri-split tRNA, in which the tRNA is composed of three individual transcripts (transcript 1 with a *green* exon, transcript 2 with an *orange* exon, and transcript 3 with a *red* exon; leader sequences are indicated in *gray*). **g** Permuted tRNA, in which the 3′ region (*red* exon) of the tRNA occurs upstream from the 5′ region (*green* exon). An intervening region is indicated in *blue* and the leader sequences are indicated in *gray*. **h** Intron-containing permuted tRNAs have been reported to contain an endogenous intron. The two *black arrowheads* indicate the exon–intron boundaries cleaved by tRNA splicing endonuclease. The two *blue arrowheads* indicate the processing boundaries cleaved by putative ribonucleases

these tRNAs are initially transcribed as precursor forms, pre-tRNAs. Therefore, the RNA processing steps are very important in removing the intronic sequences to generate each mature tRNA. Note that all these tRNAs contained the BHB structure(s) that is targeted by the tRNA splicing endonucleases. Other processing enzymes, possibly RNase P and tRNase Z, are also required to remove intervening regions in the permuted tRNAs (Soma et al. 2007).

## 12.2 Archaeal Phyla and tRNA Introns

Based on the analysis of 16S rRNA gene sequences in the late 1970s, Dr. Carl Woese reported that there are two groups of prokaryotes: the Bacteria and Archaea (Woese and Fox 1977). Thirty-five years later, more than 1,000 bacterial and archaeal genomes have been completely sequenced. These results support the idea that the organisms in the archaeal domain bear no close relationship to those in the bacterial domain, and are actually more closely related to the Eukarya according to the analysis of genes related to the processing of genetic information, such as those involved in DNA replication, transcription, and translation. The archaeal domain is further subdivided into two main phyla: the Euryarchaeota and the Crenarchaeota. In 2002, a new species, *Nanoarchaeum equitans*, was discovered and has been given its own phylum, the Nanoarchaeota (Huber et al. 2002). *Nanoarchaeum* appears to be an obligatory symbiont on the archaeon *Ignicoccus.* Other phyla, such as the Korarchaeota (Barns et al. 1996) and the Taumarchaeota (Brochier-Armanet et al. 2008), have also been proposed.

Most introns of the tRNA genes in the Euryarchaeota are located in the anticodon loop (37/38; Fig. 12.3b) and the strict hBHBh' is used as the tRNA splicing motif (Fig. 12.2). However, a large number of tRNA introns are found in the Crenarchaeota and many of them are located not only in the anticodon loop region but also at non-canonical positions. In addition to the strict hBHBh′ motif, the relaxed hBH motif (usually located at 37/38) and the relaxed HBh′ motif (located at non-canonical positions) are used in the phylum Crenarchaeota (Fig. 12.2).

## 12.3 Multiple-Intron-Containing tRNAs, Split tRNAs, and Permuted tRNAs

In 2006, our group developed the SPLITS software for predicting split and intron-containing tRNA genes at the genome level (Sugahara et al. 2006). At that time, the number of sequenced archaeal genomes was increasing sharply, but many of the essential tRNA genes were still missing, especially in the phylum Crenarchaeota. Because these tRNA genes are interrupted by non-canonical intron(s) or even separated across two or three genes, these types of tRNA genes can rarely be discovered

with conventional prediction software. SPLITS initially predicts the BHB splicing motif used to determine and remove the intronic regions of tRNA genes. The intron-removed DNA sequences are automatically queried in tRNAscan-SE (Lowe and Eddy 1997), one of the most useful softwares for tRNA gene prediction. Thus, SPLITS can predict tRNAs with single introns located at unconventional sites in the genes (Fig. 12.3c), tRNAs with double or triple introns (Fig. 12.3d), and split tRNAs (Fig. 12.3e). Using this software, we have initially reported many intron-containing tRNAs (with a maximum of three introns), including pre-tRNA$^{Glu}$(UUC) from *Pyrobaculum calidifontis* and pre-tRNA$^{Pro}$(UGG) from *P. islandicum* (Sugahara et al. 2008).

Split tRNAs, in which the 5′ and 3′ halves are encoded in two separate genes, were first found in the hyperthermophilic archaeon *N. equitans* (Randau et al. 2005b). An RNA-seq deep sequencing study of the tRNA half precursors in this organism further supported the *trans*-splicing of the tRNA (Randau 2012). Because *N. equitans* is a parasite, with evidence of a massive genomic reduction (Waters et al. 2003), it was unclear whether its tRNAs represented a form of tRNA unique to certain archaeal species or a later product of its genomic reduction. In 2009, we reported split tRNA genes in a free-living organism, the hyperthermo-acidophilic archaeon *Caldivirga maquilingensis*, which belongs to the deep-branching archaeal order Thermoproteales and was isolated from an acidic hot spring in the Philippines (Fujishima et al. 2009). Notably, three mature tRNA$^{Gly}$s with synonymous codons are created from one constitutive 3′ half transcript and four alternatively switched transcripts, generating a tRNA from a total of three transcripts, and designated a "tri-split tRNA" (Fig. 12.4). More recently, split tRNAs have also been found in several different species of the Desulfurococcales branch of the Crenarchaeota: tRNA$^{Asp}$(GUC) in *Aeropyrum pernix* and *Thermosphaera aggregans*, and tRNA$^{Lys}$(CUU) in *Staphylothermus hellenicus* and *S. marinus*. These results suggest that split tRNA genes are spread sporadically across a major branch of the Archaea (Chan et al. 2011). Interestingly, split tRNAs and intron-containing tRNAs share a common BHB motif around their intron/leader–exon boundaries, which can be cleaved by the same tRNA splicing endonucleases (discussed below) (Fujishima et al. 2009; Randau et al. 2005a).

Permuted tRNA is another unique form of tRNA, first discovered in the nuclear genome of a red alga, *Cyanidioschyzon merolae* (Soma et al. 2007). This tRNA is organized with the 5′ and 3′ halves of the gene positioned in reverse in the genome and is expressed via a circular RNA intermediate (Fig. 12.3g). A wide-ranging distribution of permuted tRNA genes was then identified in the genomes of photosynthetic eukaryotes (Maruyama et al. 2010), including the smallest eukaryotic genome known to date, the nucleomorph genome of the chlorarachniophyte alga *Bigelowiella natans*. In 2011, the first examples of permuted tRNA genes in the Archaea were reported: tRNA$^{iMet}$(CAU) and tRNA$^{Tyr}$(GUA) in *Thermofilum pendens* (Chan et al. 2011). So far, only permuted tRNA genes containing single introns have been found in archaeal species (Fig. 12.3f).

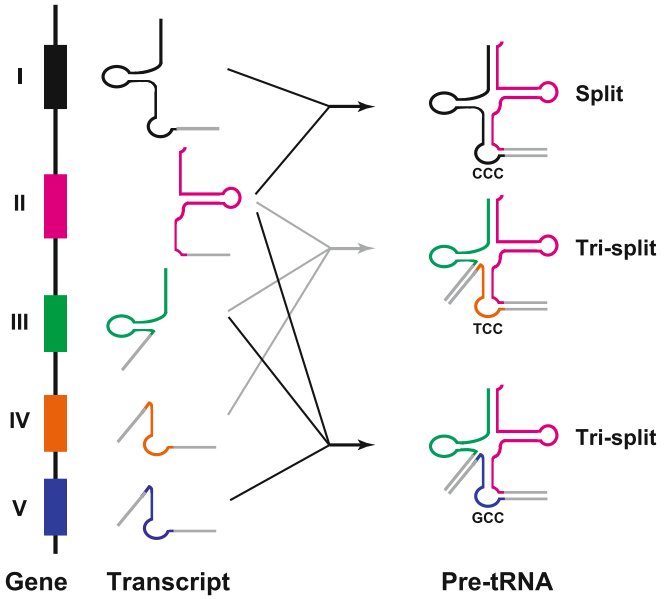



**Fig. 12.4** Schematic representation of the alternative *trans*-splicing tRNAs in a hyperthermoacidophilic archaeon *Caldivirga maquilingensis*. Combinations of two (*split*) or three (*tri-split*) transcripts generate each mature $tRNA^{Gly}$ with different anticodons

## 12.4 Co-evolution of Types of tRNA Introns and tRNA Splicing Endonucleases

Previously, three categories of tRNA splicing endonucleases have been characterized in the Archaea (Tocchini-Valentini et al. 2005b): homodimeric $\alpha_2$ and homotetrameric $\alpha_4$ enzymes in the Euryarchaeota, and a heterotetrameric $(\alpha\beta)_2$ enzyme in the Crenarchaeota. Either the homodimeric or homotetrameric enzyme requires the strict hBHBh′ motif, whereas the heterotetrameric endonucleases recognize broader substrates with relaxed BHB motifs located at non-canonical positions in Crenarchaeal genomes (Fig. 12.5). Thus, the structures of the endonucleases and their substrates appear to have evolved together (Tocchini-Valentini et al. 2005a). In other words, the co-evolution of the tRNA intron motifs and the tRNA endonuclease architecture is observed. Recently, the crystal structures of the tRNA splicing endonucleases belonging to the $(\alpha\beta)_2$ enzyme family have been determined and a Crenarchaea-specific loop structure in the enzyme is reported to play an important role in its broad substrate specificity (Hirata et al. 2011; Yoshinari et al. 2009). A conserved lysine residue in the loop is especially critical for the activity of the Crenarchaeal tRNA splicing endonucleases (Okuda et al. 2011).

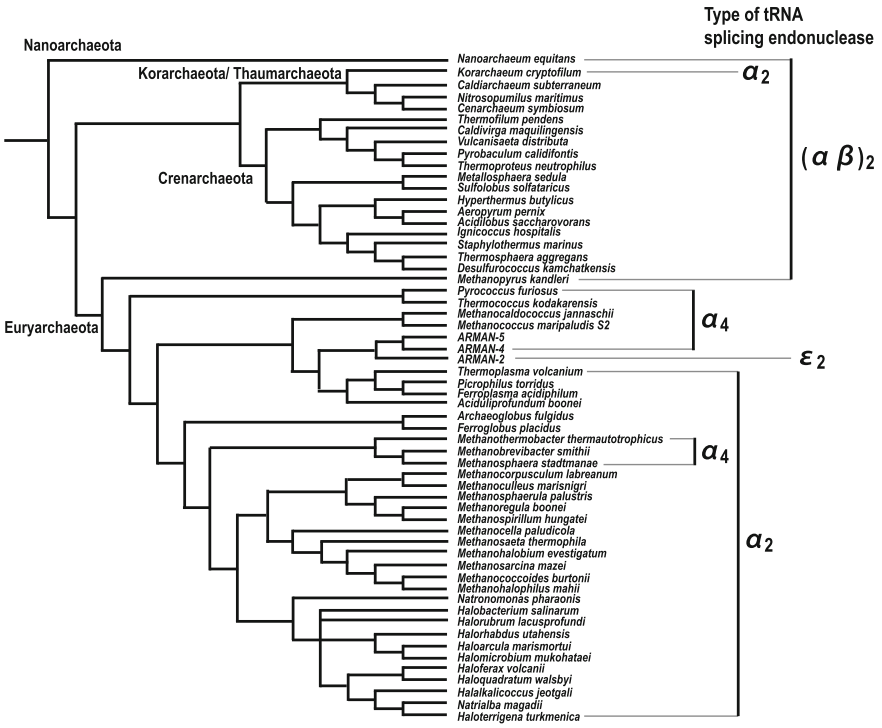



**Fig. 12.5** Phylogenetic tree of representative archaeal species and their types of tRNA splicing endonucleases. A 16S rRNA phylogenetic tree of 55 representative archaeal species (one species per genus) is shown, with the type of tRNA splicing endonuclease of each species. Phylum names are indicated on the branches. Four types of splicing endonucleases ($\alpha_4$, $\alpha_2$, $[\alpha\beta]_2$, and $\varepsilon_2$) are denoted with the corresponding archaeal species

In 2011, our group found a fourth type of tRNA splicing endonuclease ($\varepsilon_2$ type) in the genome of the acidophilic archaeon *Candidatus Micrarchaeum acidiphilum*, referred to as ARMAN-2 (Fujishima et al. 2011). The ARMAN (Archaeal Richmond Mine Acidophilic Nanoorganisms) were first discovered in an acid mine drainage site in northern California as uncultured Archaea (Baker et al. 2006) and are some of the smallest organisms known. The near-complete genomes (each about 1 Mb) of three ARMAN lineages, ARMAN-2, -4, and -5, were reconstructed from shotgun genomic sequences (Baker et al. 2010) and phylogenetic analyses of their rRNA genes suggest that the ARMAN lineages share a common ancestor with the Euryarchaeota (Fig. 12.5), but are very deeply branched in that group. However, ARMAN-2 and ARMAN-4/5 possess different types of tRNAs and tRNA splicing endonucleases. Unlike ARMAN-4/5, the intron-containing tRNA genes in ARMAN-2 are highly disrupted by introns at various positions. We found that the ARMAN-2 possess a novel tRNA splicing enzyme consisting of two duplicated catalytic units and one structural unit encoded on a single gene, representing a novel three-unit architecture (Fujishima et al. 2011). We showed that

the $\varepsilon_2$ endonuclease cleaves both canonical and non-canonical BHB motifs, similar to the $(\alpha\beta)_2$ endonuclease. Thus, the discovery of the $\varepsilon_2$ endonuclease in an archaeon deeply branched within the Euryarchaeota provides a new example of the co-evolution of tRNAs and their processing enzymes. A crystal structure analysis of the $\varepsilon_2$ homodimer supports its three-unit architecture and revealed that an ARMAN-2-type-specific loop (ASL) is involved in its broad substrate specificity (Hirata et al. 2012).

## 12.5 Translocatable tRNA Introns

Within the phylum Crenarchaeota, archaeal species in the order Thermoproteales have greater number of tRNA introns than other species (Sugahara et al. 2008). A sequence analysis of the introns in this order showed that identical introns are conserved in different tRNA genes (Fujishima et al. 2010; Sugahara et al. 2008). This result suggests that the introns translocated their positions during tRNA evolution. To estimate the exact proportions of these translocatable introns, we first collected 286 tRNA introns from seven representative archaeal species (*C. maquilingensis*, *Pyrobaculum aerophilum*, *P. arsenaticum*, *P. calidifontis*, *P. islandicum*, *Thermofilum pendens*, and *Thermoproteus neutrophilus*) in the order Thermoproteales and identified 46 intron groups containing sets of highly similar sequences (>70 %). This analysis showed that 16 of these intron groups contained sequences from evolutionarily distinct tRNA genes. The phylogeny of these 16 intron groups indicated that transposition events have occurred at least seven times throughout the evolution of the Thermoproteales (Fujishima et al. 2010). These findings suggest that frequent intron transposition occurs among the tRNA genes of the order Thermoproteales. A computational analysis revealed the limited insertion positions of these introns and the corresponding amino acid types of these tRNA genes. The limited number of positions containing these transposed introns has arisen because the BHB splicing motif is required at the newly transposed position if the pre-tRNA is to be correctly processed.

We also compared the proportions of translocatable introns at each tRNA nucleotide position with the proportions of other archaeal tRNA introns at those positions. As shown in Fig. 12.6, most of the translocatable tRNA introns are located at non-canonical positions (other than 37/38). Thus, introns located at the canonical position are not amenable to transposition. It has been shown that about half of these non-canonical introns are translocatable. The majority of non-canonical introns are distributed in the T-arm region of the tRNA. Interestingly, 10 split tRNA genes have been identified in the genome of *C. maquilingensis*, and of these, the two split $tRNA^{Ala}$s are separated at position 29/30, a known insertion site for introns (Fujishima et al. 2010). The leader sequences of these split $tRNA^{Ala}$s share 90 % sequence identity with one of translocatable introns inserted in the orthologous $tRNA^{Ala}$ gene of *P. islandicum*, indicating a clear evolutionary relationship between the intron-containing and split tRNAs (Fujishima et al. 2009).

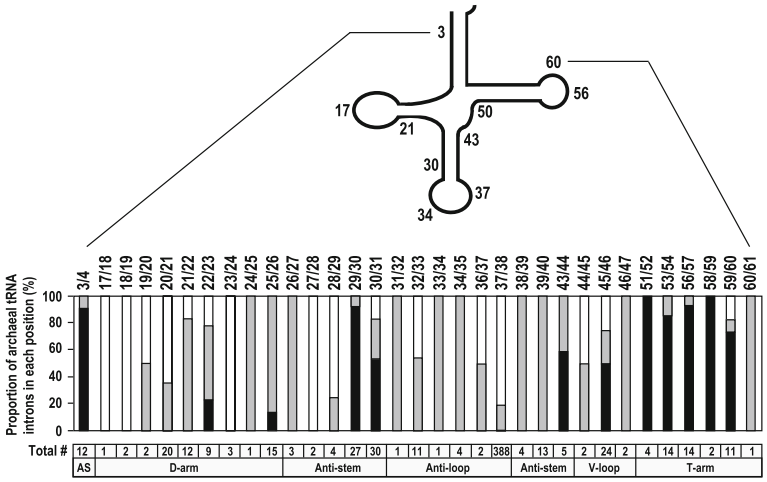



**Fig. 12.6** Proportion of transposable introns at each nucleotide position in archaeal tRNAs. The boxed numbers represent the distribution of the total 647 tRNA introns found in 58 archaeal species (see Fujishima et al. 2010 for details). The tRNA introns are classified into three categories: transposable introns in seven Thermoproteales species (*black*), other introns in seven Thermoproteales species (*gray*), and tRNA introns in the remaining 51 archaeal species (*white*). Their proportions are given for each tRNA nucleotide position. AS, acceptor stem; D-arm, dihydrouridine arm; Anti-stem, anticodon stem; Anti-loop, anticodon loop; V-loop, variable loop; T-arm, TΨC arm

An analysis of the origins of the split tRNA genes according to the trends in the gain and loss of their introns suggested that the fragmentation of the $tRNA^{Ala}$ gene may have occurred in addition to the insertion of the intron sequence at position 29/30. Recently, the group of Dr. Todd Lowe also noted examples of "recently" split tRNAs based on a comparative analysis of tRNA genes in the Desulfurococcales branch of the Archaea (Chan et al. 2011).

## 12.6 Conclusions

In the beginning of this chapter, I mentioned that the tRNA introns located in the anticodon loop region seem to be evolutionarily ancient. I also pointed out that tRNA introns located at non-canonical positions are translocatable; these introns may have been acquired at a later stage of tRNA evolution. The concept is summarized in Fig. 12.7 and suggests that tRNA introns have different backgrounds, depending on their insertion sites.

Previously, we mimicked split tRNA by artificially separating the tRNA sequences of seven primitive archaeal species at the anticodon and analyzed the sequence similarity and diversity of the 5′ and 3′ tRNA halves (Fujishima et al. 2008). Network analysis revealed specific characteristics and topological

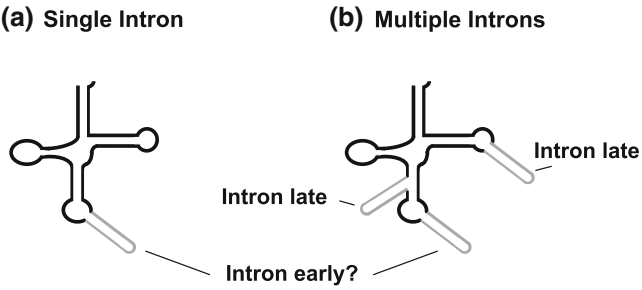



**Fig. 12.7** A model of "early" and "late" tRNA introns

differences between the tRNA halves, suggesting different evolutionary backgrounds of the 5′ and 3′ tRNA halves, although they are short fragments of a single molecule (Fujishima et al. 2008). Considering that the Archaea is a known deep-branching taxon, it is reasonable to infer that these split features of tRNAs reflect a very ancient character carried over from the early phase of life. Dr. Di Giulio hypothesized that the ancestral tRNA was encoded by two separate mini-genes, which later fused to encode the modern tRNAs (Di Giulio 2009, 2012). I suggest that the introns located in the anticodon loop might reflect the combinations of the 5′ and 3′ halves.

Dr. Lennart Randau simultaneously suggested the possibility that some disrupted tRNA genes were derived from gene separation during the process of genome rearrangement and have since been maintained as a protective mechanism against the integration of mobile genetic elements (e.g., conjugative plasmids and/or viruses), because non-intronic tRNA genes are ideal targets for site-specific recombination (Randau and Soll 2008). It has also been suggested that at least some split tRNAs were generated from intron-containing tRNAs in the later stage of tRNA evolution (Chan et al. 2011; Fujishima et al. 2010). In 2011, a composite genome of the deep-branching archaeon *Caldiarchaeum subterraneum* was reconstructed from a community genomic library (Nunoura et al. 2011). Based on a sequence analysis of the community library, it was proposed that the gain of the tRNA intron and the scattering of the tRNA fragments occurred within a short time frame via the integration and recombination of a mobile genetic element (Sugahara et al. 2012). Therefore, it is also true that the tRNA genes of organisms can be divided into fragments, even in the later stages of molecular evolution. In that case, many new introns may appear at non-canonical positions of tRNAs.

In conclusion, we have found several unique tRNAs in the Archaea and believe that further study of these tRNAs should open new doors in understanding the origin of the genetic code. However, we have so far acquired the complete genomic sequences of only 100 archaeal species, too few to answer the underlying questions about their evolutionary relationships. To identify the ancient form of tRNA and to gain a clear overview of the evolutionary transitions of the tRNA types in the Archaea, it will be necessary to collect more genomic sequences and to examine each tRNA structure through their phylogeny. The genomic sequence

of a member of the very deep-branching Archaea will be especially useful in this process. For this aim to be achieved, next-generation sequencing of various environmental samples is required.

**Acknowledgments** I thank all the members of the RNA Group at the Institute for Advanced Biosciences, Keio University, Japan, for their insightful discussions. Drs Junichi Sugahara (currently Spiber Inc., Japan) and Kosuke Fujishima (currently NASA Ames Research Center, USA), in particular, contributed to the discovery of novel tRNAs in the Archaea during their graduate studies at Keio University. This research was supported, in part, by a Grant-in-Aid for Scientific Research (B) (no. 22370066) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (to A.K.).

## References

Baker BJ, Tyson GW, Webb RI, Flanagan J, Hugenholtz P, Allen EE, Banfield JF (2006) Lineages of acidophilic archaea revealed by community genomic analysis. Science 314:1933–1935

Baker BJ, Comolli LR, Dick GJ, Hauser LJ, Hyatt D, Dill BD, Land ML, Verberkmoes NC, Hettich RL, Banfield JF (2010) Enigmatic, ultrasmall, uncultivated Archaea. Proc Natl Acad Sci USA 107:8806–8811

Barns SM, Delwiche CF, Palmer JD, Pace NR (1996) Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. Proc Natl Acad Sci USA 93:9188–9193

Belfort M, Weiner A (1997) Another bridge between kingdoms: tRNA splicing in archaea and eukaryotes. Cell 89:1003–1006

Brochier-Armanet C, Boussau B, Gribaldo S, Forterre P (2008) Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. Nat Rev Microbiol 6:245–252

Chan PP, Cozen AE, Lowe TM (2011) Discovery of permuted and recently split transfer RNAs in Archaea. Genome Biol 12:R38

Di Giulio M (2009) Formal proof that the split genes of tRNAs of Nanoarchaeum equitans are an ancestral character. J Mol Evol 69:505–511

Di Giulio M (2012) The origin of the tRNA molecule: Independent data favor a specific model of its evolution. Biochimie 94:1464–1466

Fujishima K, Sugahara J, Tomita M, Kanai A (2008) Sequence evidence in the archaeal genomes that tRNAs emerged through the combination of ancestral genes as 5′ and 3′ tRNA halves. PLoS One 3:e1622

Fujishima K, Sugahara J, Kikuta K, Hirano R, Sato A, Tomita M, Kanai A (2009) Tri-split tRNA is a transfer RNA made from 3 transcripts that provides insight into the evolution of fragmented tRNAs in archaea. Proc Natl Acad Sci USA 106:2683–2687

Fujishima K, Sugahara J, Tomita M, Kanai A (2010) Large-scale tRNA intron transposition in the archaeal order Thermoproteales represents a novel mechanism of intron gain. Mol Biol Evol 27:2233–2243

Fujishima K, Sugahara J, Miller CS, Baker BJ, Di Giulio M, Takesue K, Sato A, Tomita M, Banfield JF, Kanai A (2011) A novel three-unit tRNA splicing endonuclease found in ultrasmall Archaea possesses broad substrate specificity. Nucleic Acids Res 39:9695–9704

Hirata A, Kitajima T, Hori H (2011) Cleavage of intron from the standard or non-standard position of the precursor tRNA by the splicing endonuclease of Aeropyrum pernix, a hyper-thermophilic Crenarchaeon, involves a novel RNA recognition site in the Crenarchaea specific loop. Nucleic Acids Res 39:9376–9389

Hirata A, Fujishima K, Yamagami R, Kawamura T, Banfield JF, Kanai A, Hori H (2012) X-ray structure of the fourth type of archaeal tRNA splicing endonuclease: insights into the evolution of a novel three-unit composition and a unique loop involved in broad substrate specificity. Nucleic Acids Res 40:10554–10566

Huber H, Hohn MJ, Rachel R, Fuchs T, Wimmer VC, Stetter KO (2002) A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. Nature 417:63–67

Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:955–964

Lund AH, Duch M, Pedersen FS (2000) Selection of functional tRNA primers and primer binding site sequences from a retroviral combinatorial library: identification of new functional tRNA primers in murine leukemia virus replication. Nucleic Acids Res 28:791–799

Mak J, Kleiman L (1997) Primer tRNAs for reverse transcription. J Virol 71:8087–8095

Marck C, Grosjean H (2003) Identification of BHB splicing motifs in intron-containing tRNAs from 18 archaea: evolutionary implications. RNA 9:1516–1531

Maruyama S, Sugahara J, Kanai A, Nozaki H (2010) Permuted tRNA genes in the nuclear and nucleomorph genomes of photosynthetic eukaryotes. Mol Biol Evol 27:1070–1076

Nunoura T, Takaki Y, Kakuta J, Nishi S, Sugahara J, Kazama H, Chee GJ, Hattori M, Kanai A, Atomi H et al (2011) Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. Nucleic Acids Res 39:3204–3223

Okuda M, Shiba T, Inaoka DK, Kita K, Kurisu G, Mineki S, Harada S, Watanabe Y, Yoshinari S (2011) A conserved lysine residue in the crenarchaea-specific loop is important for the crenarchaeal splicing endonuclease activity. J Mol Biol 405:92–104

Randau L (2012) RNA processing in the minimal organism Nanoarchaeum equitans. Genome Biol 13:R63

Randau L, Soll D (2008) Transfer RNA genes in pieces. EMBO Rep 9:623–628

Randau L, Calvin K, Hall M, Yuan J, Podar M, Li H, Soll D (2005a) The heteromeric Nanoarchaeum equitans splicing endonuclease cleaves noncanonical bulge-helix-bulge motifs of joined tRNA halves. Proc Natl Acad Sci USA 102:17934–17939

Randau L, Munch R, Hohn MJ, Jahn D, Soll D (2005b) Nanoarchaeum equitans creates functional tRNAs from separate genes for their 5′- and 3′-halves. Nature 433:537–541

Saldanha R, Mohr G, Belfort M, Lambowitz AM (1993) Group I and group II introns. FASEB J 7:15–24 (official publication of the Federation of American Societies for Experimental Biology)

Soma A, Onodera A, Sugahara J, Kanai A, Yachie N, Tomita M, Kawamura F, Sekine Y (2007) Permuted tRNA genes expressed via a circular RNA intermediate in Cyanidioschyzon merolae. Science 318:450–453

Sugahara J, Yachie N, Sekine Y, Soma A, Matsui M, Tomita M, Kanai A (2006) SPLITS: a new program for predicting split and intron-containing tRNA genes at the genome level. In Silico Biol 6:411–418

Sugahara J, Kikuta K, Fujishima K, Yachie N, Tomita M, Kanai A (2008) Comprehensive analysis of archaeal tRNA genes reveals rapid increase of tRNA introns in the order thermoproteales. Mol Biol Evol 25:2709–2716

Sugahara J, Fujishima K, Nunoura T, Takaki Y, Takami H, Takai K, Tomita M, Kanai A (2012) Genomic heterogeneity in a natural archaeal population suggests a model of tRNA gene disruption. PLoS One 7:e32504

Tanner M, Cech T (1996) Activity and thermostability of the small self-splicing group I intron in the pre-tRNA(Ile) of the purple bacterium Azoarcus. RNA 2:74–83

Tocchini-Valentini GD, Fruscoloni P, Tocchini-Valentini GP (2005a) Coevolution of tRNA intron motifs and tRNA endonuclease architecture in Archaea. Proc Natl Acad Sci USA 102:15418–15422

Tocchini-Valentini GD, Fruscoloni P, Tocchini-Valentini GP (2005b) Structure, function, and evolution of the tRNA endonucleases of Archaea: an example of subfunctionalization. Proc Natl Acad Sci USA 102:8933–8938

Trotta CR, Miao F, Arn EA, Stevens SW, Ho CK, Rauhut R, Abelson JN (1997) The yeast tRNA splicing endonuclease: a tetrameric enzyme with two active site subunits homologous to the archaeal tRNA endonucleases. Cell 89:849–858

Waters E, Hohn MJ, Ahel I, Graham DE, Adams MD, Barnstead M, Beeson KY, Bibbs L, Bolanos R, Keller M et al (2003) The genome of Nanoarchaeum equitans: insights into early archaeal evolution and derived parasitism. Proc Natl Acad Sci USA 100:12984–12988

Woese CR, Fox GE (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc Natl Acad Sci USA 74:5088–5090

Yoshinari S, Shiba T, Inaoka DK, Itoh T, Kurisu G, Harada S, Kita K, Watanabe Y (2009) Functional importance of crenarchaea-specific extra-loop revealed by an X-ray structure of a heterotetrameric crenarchaeal splicing endonuclease. Nucleic Acids Res 37:4787–4798

# Chapter 13
# Memory of Temperature in the Seasonal Control of Flowering Time: An Unexplored Link Between Meteorology and Molecular Biology

**Hiroshi Kudoh and Atsushi J. Nagano**

**Abstract** Organismal response to a specific environmental signal should have a property that is sensitive to the useful cue but insensitive to other erroneous noises. Because adaptive characteristics of an organism have been shaped by natural selection, their functions should be evaluated in the natural habitats. In the study of flowering time, recent emergence of *in natura* systems biology i.e. combined applications of molecular biology and statistical modeling in natural habitats, bridges biometeorology and molecular biology. Especially, expression of genes with known function can be analyzed by conventional time-series analyses that have been frequently used in biometeorological studies. Here, by comparing knowledge from (1) meteorology, (2) molecular biology (3) *in natura* systems biology, and (4) biometeorology, we showed how signal properties in natural temperature fluctuation correspond to mechanistic property of flowering-time regulation. Meteorological data indicate that seasonal signal in temperature is a pattern over multiple weeks or months with short-term serious noises. It predicts that molecular machinery of flowering-time control should refer past temperatures for 4–6 weeks or longer. Molecular biology on the regulation of a key flowering repressor, *FLC*, in *Arabidopsis thaliana* revealed that molecular machinery of vernalization response is a representative example that corresponds to the meteorological properties of the seasonal temperature signal. *In natura* systems biology has revealed that the machinery of vernalization response serves as memory of past temperature that extract seasonal signal from natural complex fluctuation. The referring periods of past temperatures suggested by meteorology, molecular biology and *in natura* systems biology correspond to the lengths of proceeding periods during which temperature affect first flowering dates of diverse plants. Therefore, we should not take yearly variation of flowering date as a mealy passive

H. Kudoh (✉) · A. J. Nagano
Center for Ecological Research, Kyoto University, Otsu, Japan
e-mail: kudoh@ecology.kyoto-u.ac.jp

A. J. Nagano
Japan Science and Technology Agency, PRESTO, Chiyoda, Tokyo, Japan

P. Pontarotti (ed.), *Evolutionary Biology: Exobiology and Evolutionary Mechanisms*,
DOI: 10.1007/978-3-642-38212-3_13, 

developmental effect of preceding temperatures. It includes more active responses of plants that have been evolved to control flowering-time in the noisy temperature fluctuations in the natural habitats.

## 13.1 Introduction

Any organisms have their own habitats that are characterized by specific sets of environment. In the natural habitats, changing patterns of environment are generally complex. Environmental changes sometime provides useful cue for organisms to respond physiologically in short term and developmentally in long term. Such environmental responses allow organisms to acquire resources efficiently, avoid or tolerate stresses, and reproduce successfully. Environmental cues, however, often come with obstructive noises in the natural habitats. Biological functions that respond to a specific environmental cue, therefore, should have a property that is sensitive to the useful cues but insensitive to other erroneous noises.

Because adaptive characteristics of an organism have been shaped by natural selection, their functions should be evaluated in the natural habitats. In the study of gene functions, experiments are often conducted in the stable laboratory conditions. Under careful controls of environments, many biological functions have been dissected into molecular components. The identifications of genes and their interactions allowed us to introduce systems biological approaches to the study of gene functions (Richards et al. 2009), but such attempts in natural habitats have been emerged recently (Merquiol et al. 2002; Anderson et al. 2005; Mayrhofer et al. 2005; Manitašević et al. 2007; Aikawa et al. 2010; Richards et al. 2012; Nagano et al. 2012). We call such type of analyses as *'in natura'* following precedents of *'in vivo'* and *'in vitro'* contrasts in the study of biology (Quintana-Murci 2007; Shimizu et al. 2011).

One of critical novel points of gene functions suggested by recent *in natura* studies is the robustness against erroneous noises (Aikawa et al. 2010; Izawa 2012; Nagano et al. 2012). It is expected that machinery of gene regulations should corresponds to the patterns and magnitude of signals (environmental cues) relative to noises in natural habitats. In the primal aim of this chapter is to explore whether properties of natural environmental cues predict characteristics of molecular machinery of biological functions. We chose 'flowering time' as a focusing biological phenomenon. The advantage of choosing flowering time stands on great maturations of knowledge in two study areas, i.e., biometeorology and molecular biology. The former has been analyzed flowering time in relation to natural habitat environments, and the later has been analyzed molecular mechanisms underlying flowing-time control in terms of developmental biology. Although the two aspects of flowering have been studied rather separately, there should be a close link between the patterns found in two sets of studies (Fig. 13.1).

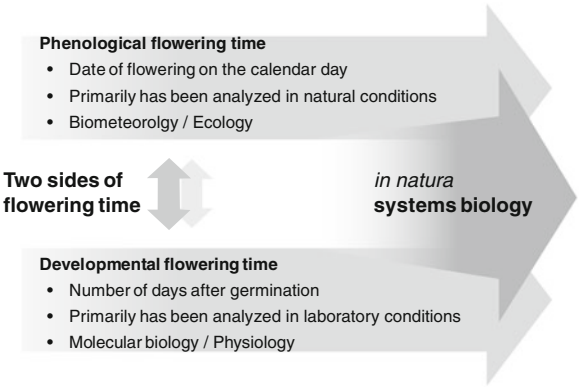



**Fig. 13.1** A schematic diagram of representing two sides of flowering time, i.e., phenological flowering time and developmental flowering time. Recent emergence of *in natura* systems biolog bridges biometeorology/ecology and molecular biology

## 13.2 Phenological and Developmental Flowering Times

In temperate regions, seasonality of flowering time is commonly observed among diverse plants (Fitter and Peat 1994; Tooke and Battey 2010). Each plant species often shows peak of flowering-time that ranges approximately 10–20 days, and we often observe sequences of flowering in different species at the community level (Ahas et al. 2000; Sparks et al. 2000). It is considered that each plant species has been evolved to bloom in a specific season when its reproductive success is maximized (Munguia-Rosas et al. 2011). To flower in specific seasons, plants utilize external cues from the outer environments, such as photoperiod, temperature and rain-fall (Lang 1952; Salisbury 1963; Bernier 1988).

There are two ways of describing flowering time, i.e., phenological flowering time and developmental flowering time (Fig. 13.1). Phenological flowering time is described as date of flowering on the calendar days and has been analyzed primarily under natural conditions in the field-oriented studies, such as biometeorology and ecology (Fig. 13.1). Developmental flowering time is described as the number of days after germination and has been analyzed under laboratory conditions in molecular biology and physiology (Fig. 13.1).

Phenological flowering time often shows yearly variation and its environmental determinants have been long interest of biometeorology (Schwartz 2003). Predicting flowering-time has been the important task in agriculture and monitoring climate changes, and data on flowering phenology over decades have been accumulated in many localities (Fitter and Peat 1994; Cleland 2007, Tooke and Battey 2010). Determinants of flowering phenology have been modeled primarily in relation to temperatures and with combinations of other environmental factors (Schwartz 2003). Recent systematic shifts of flowering time along with climate changes also suggest that temperature is the major determinant of flowering time (Walther et al. 2002; Root et al. 2003).

Developmental flowering time has been analyzed in plant physiology using diverse plant species, including important crops, and in molecular biology using the model plants (Simpson and Dean 2002; Andrés and Coupland 2012). In

*Arabidopsis thaliana*, more than 120 genes that affect flowering-time have been identified, and their functions and interactions with other molecular components have been under extensive investigations (Yanovsky and Key 2003). The response to temperature and photoperiods are two major flowering-time control systems which mechanisms have been most actively investigated in plant molecular biology (Andrés and Coupland 2012).

Although phenological and developmental flowering times should be closely connected with each other, the combined analysis and interpretation are not so straightforward. For example, early or late flowering in developmental flowering time does not necessarily correspond with early or late flowering date in phenological flowering time (Wilczek et al. 2009). Because natural populations usually are the mixture of plants with different ages, developmental flowering time may not correspond to phonological flowering time. Observations of synchronous flowering in terms of phenological flowering-time in natural plant communities make us speculate that plant responses to environments have evolved to alter developmental flowering time at individual level for synchronization of phenological flowering time between individuals at the population level.

Recent emergence of *in natura* systems biology i.e. combined applications of molecular biology and statistical modeling in natural habitats, can be regarded as the attempts to bridge biometeorology and molecular biology (Aikawa et al. 2010; Wilczek et al. 2010; Chew et al. 2012; Nagano et al. 2012). Especially, expression of genes with known function can be analyzed by conventional time-series analyses that have been frequently used in biometeorological studies (Aikawa et al. 2010; Nagano et al. 2012). Here, by comparing knowledge from (1) meteorology, (2) molecular biology (3) *in natura* systems biology, and (4) biometeorology (Table 13.1), we would like to show how signal properties in natural temperature fluctuation correspond to mechanistic property of flowering-time regulation. We first attempts to predict required characteristics of machinery of seasonal response in plants solely by the properties of meteorological temperature data. Following the prediction, we overlook what types of molecular machinery has been identified for seasonal temperature response in plant molecular biology. Then we introduce of our recent study of *in natura* systems biology that evaluate properties of a temperature-responsive flowering-time gene in the natural habitats. Finally, we revisit conventional biometeorological studies to add a novel aspect of interpretation on phenological data, which link biometeorology and molecular biology.

## 13.3 Properties of Seasonal Temperature Signals

In this section, we elucidate signal properties of seasonal temperature to predict general characteristics of the machineries for detecting seasons. Except for near to the equator, time intervals between sunrise and sunset (photoperiod) change greatly with the systematic annual-cycles that are predictable from latitude (Withrow 1959; Lee 1970). The patterns are attributable to the 23.4 degree of the

**Table 13.1** Comparisons of four related study areas in predicting properties of 'temperature memory' in the control of seasonal flowering time

| Analyses | Data | Major results | Predicted properties of 'temperature memory' |
|---|---|---|---|
| Meteorology | Meteorological temperature records | 1. Seasonal temperature is a pattern over months | 1. Sensitive to the seasonal temperatures but insensitive to short-term noises |
| | | 2. Temperature includes large short-term noises | |
| | | 3. Referring past temperatures (e.g. SMA) provide high S/N of seasonal signal | 2. Referring periods for 4-6 weeks or longer |
| Molecular biology | Functional analyses of temperature responsive genes | 1. A key regulator of vernalization in *A. thaliana*, *FLC*, shows response to long-term cold | 1. Slow response over multiple weeks |
| | | | 2. Use of cellular memories (histone modification) |
| | | 2. Multiple regulation processes allow *FLC* expression to respond slowly | 3. Stable through cell divisions |
| | | 3. Histone modifications and chromatin remodeling give the mechanistic bases of quantitative and long-term responses | |
| *In natura* systems biology | Gene expression and temperature records in natural habitats | 1. Expression of a *FLC* ortholog of a perennial *Arabidopsis* shows clear seasonal pattern | 1. Memory of past temperatures for ca. 6 weeks |
| | | 2. Variation of expression is mostly explicable by past 6-week temperature | 2. Robust against other environmental factors |
| | | 3. The regulation is robust enough to function under natural complex environmental conditions | |
| Biometeorology | Penology and temperature records for more than decades | Temperature during 3–10 weeks prior to flowering is most influential to yearly variation of flowering-time in many species | General presence of temperature memory over multiple weeks in diverse plants |

axial tilt (the angle between equatorial and orbital planes) of the earth. The day length is longest and shortest on the summer and winter solstices (approximately June 22 and December 22 in the northern hemisphere), respectively. The annual cycles of solar radiations create seasons (Trenberth 1983; Jones et al. 1999) and annual amplitudes of changes in photoperiods and temperature are large enough in temperate regions and higher latitudes to become the major determinants of plant phenology (Battey 2000).

Here, we analyze the seasonal temperature using the data at 35°00′ N and 135°00′ E (Nishiwaki, Hyogo Prefecture, Japan, located on the standard meridian of the Japan standard time) as an example. Although we use the temperature data in one location, most of arguments can be applicable in other places on the earth, especially for temperate regions (Trenberth 1983; Jones et al. 1999). Seasonal temperature at Nishiwaki can be best represented by the smoothed line (fitted by smooth spline, Wahba 1990) of 30 year's (1981–2010) averages of daily mean temperatures (Fig. 13.2a). Temperatures showed a clear seasonal pattern following changes in the day lengths with a 1–1.5 month lag (Fig. 13.2a) (Prescott and Collins 1951; Trenberth 1983).

Although the smoothed seasonal temperature in Fig. 13.2a represent average seasonal patterns, actual temperature fluctuate enormously at different time scales. Changes in daily mean temperature show large fluctuations between day by day

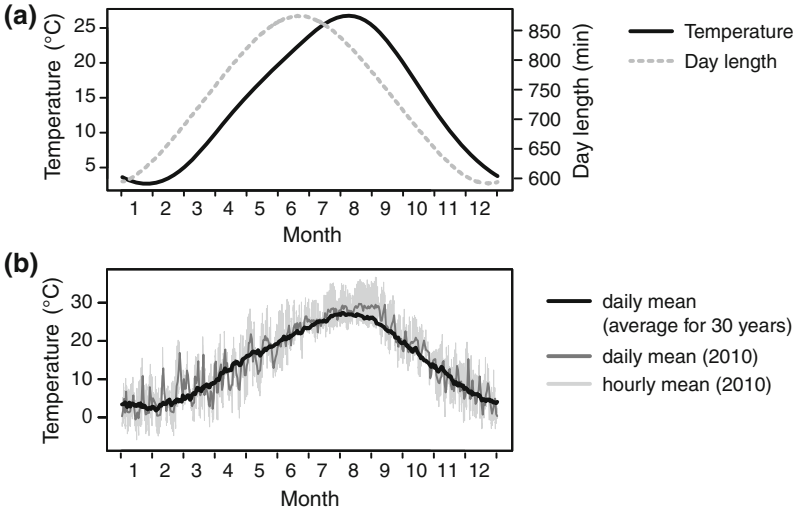



**Fig. 13.2** Typical seasonal patterns of temperature in a temperate region recorded at Nishiwaki, Japan (35°00′ N, 135°00′ E). In **a**, seasonal temperature at Nishiwaki is represented by a *thick smoothed line* (fitted by *smooth spline*) of 30 year's (1981–2010) averages of daily mean temperatures. Temperatures showed a clear seasonal pattern following changes in the day lengths (*a dotted line*) with a 1–1.5 month lag. In **b**, daily mean temperatures for 30 year's (1981–2010) averages without smoothing (*a black line*) and for 2010 (*a dark gray line*), and hourly mean temperature for 2010 (*a light gray line*) were shown. Temperature in the natural habitats involves enormous fluctuations at different time scales

and week by week (as represented by the dark gray line in Fig. 13.2b). The graph indicates that the seasonal temperature is a pattern over months (Fig. 13.2b). For example, the spring and autumn transitions are the overall upward and downward directions of temperature measurements over time, and it is common that colder days or weeks comes later during the spring transition, and vice versa in the autumn transition. If diurnal temperature fluctuations are included in the graph (the light gray line in Fig. 13.2b), we realize that temperature in the natural habitats involves enormous fluctuations. Ranges of temperature change within a single day often exceed differences in the seasonal temperature over a couple of months. Here, we list two important properties of natural temperatures and the seasonal signal.

1. Seasonal signal in temperature is a pattern over multiple weeks or months.
2. Temperature shows diurnal, day-by-day, and week-by-week fluctuations that are expected to become serious noise in the detection of seasons.

The ease of detecting a seasonal pattern depends on the amount of short-term fluctuations in the measurements—known as 'background noise', such as diurnal, day-by-day, and week-by-week temperature fluctuations. These noises can obscure the underlying seasonal signal. The relative magnitude of signal and noise is known as signal-to-noise ratio (S/N ratio). A high ratio means the signal is significantly larger than the noise, making the underlying trend of the measurements obvious. As we see in Fig. 13.2b, short-term noises of temperatures, however, is greater than the seasonal signal. It means that seasonal temperature has low signal-to-noise ratio. In general, the lower the signal-to-noise ratio, the longer the series of measurements needed to detect a reliable trend.

One way to level out the large fluctuations in a set of time-series data is to use moving averages (Lewis 1960). A simple moving average (SMA) is the unweighted mean of the previous data points. A SMA of temperature is formed by averaging past temperatures for a specific length of periods (window length). For example, 3 day SMA is the averages of temperature data for past 3 days. SMAs are calculated at any timing, and therefore, we can calculate n-day SMA for every day. Another example is daily mean temperatures, and it is equivalent for 24 h (1 day) SMA calculated at the end of each day.

We calculated a series of SMAs (1, 3, 7, 14, 28, 42, 90 and 180 days) at Nishiwaki (Fig. 13.3). Moving averages smoothed the temperature data to form a pattern following seasons, and filtered out the short-term noise. We can see that longer the referring periods of SMA, the seasonal signal became conspicuous and noises became less remarkable. We still see large spring noises and moderate autumn noises in the 7 and 14 day SMAs. In 28 day SAM, small noises exist in early spring but autumn noise becomes negligible. In 42 day SAM, spring noises also became negligible.

We quantified the strengths of seasonal signal and noises of SMAs by calculating variances. For a specific sequential temperature data, we fitted a smooth line using the smooth spline method to represent seasonal cycles (thick lines in Fig. 13.4). We defined the variance explained by the seasonal cycle line as a signal

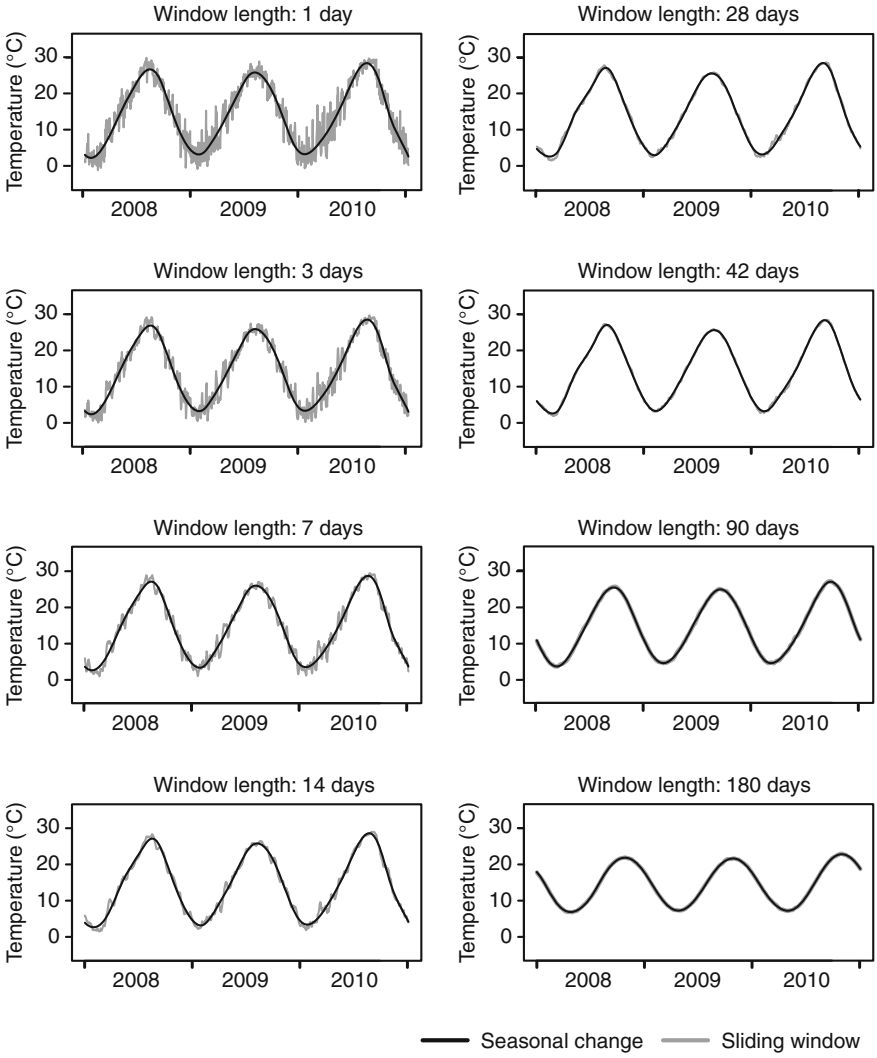



**Fig. 13.3** Simple moving averages (SMAs) of temperatures calculated for 2008–2010 at Nishiwaki with different referring periods (widow length) of 1, 3, 7, 14, 28, 42, 90 and 180 days. Moving averages calculated for everyday are in *gray lines*, and seasonal patterns (fitted by *smooth spline*) are in *dark lines*

and the residual variance as a noise. Signal and noise for SMAs with different referring periods were calculated (Fig. 13.4). Longer the referring periods of SMA, both seasonal signal and noises become smaller. However, the declining patterns are largely different. Decline of signal is small up to 42 days, but signal strength fall rapidly as averaging periods approach 90 days (Fig. 13.4a). On the other hand, as referring periods of SMA become longer, noise declines rapidly from the

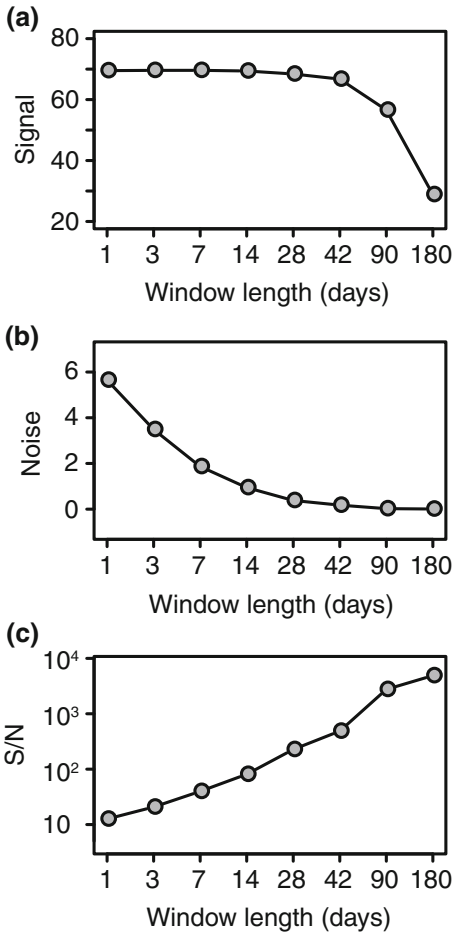



**Fig. 13.4** Levels of seasonal signal (**a**), noises (**b**) and S/N (**c**) of SMAs with different window lengths (1, 3, 7, 14, 28, 42, 90 and 180 days). Signal and noise were quantified by calculating variances (see text for the definitions of seasonal signal and noises)

beginning (Fig. 13.4b). Noise becomes considerably small if referring period is 28 days or longer. As a result, S/N ratio increases dramatically as referring periods of SMA increase (Fig. 13.4c). Principally, longer the referring periods for past temperature, seasonal signal becomes easier to detect. Although S/N increases, however, referring longer past temperatures in SMA has two other effects. First, absolute magnitude of signals becomes smaller in longer-days SMAs. Second, because SMA refers past, it causes lag of seasonal signal, and the lag become larger when referring periods become longer.

Here, we predict properties of the seasonal-response machinery of plants (Table 13.1). Because the machinery needs to be sensitive to the seasonal cycle but insensitive to short-term fluctuations, it should refer past temperatures just like SMA does. The referring periods are expected to be 28–42 days (4–6 weeks) or longer. However, referring periods should have upper limit because of three reasons; (1) signal needs to be larger than sensitivity of putative temperature sensors

of plants; (2) lag of signals should not obscure response timing; and (3) the machinery should involve a certain kind of memory to refer past temperature, but the error of memory itself is expected to become larger when referring periods become longer. Overall, we predicted that molecular machinery of flowering-time control should refer past temperatures for 4–6 weeks or longer, but not for too long (Table 13.1).

## 13.4 Molecular Characteristics of Temperature-Dependent Flowering-Time Control

In general, organismal responses to temperature at the cell level often occur within short time at the scale of seconds, minuets, and hours (Penfield 2008). Responses longer than 4–6 weeks at the cell level are likely to involve a specific mechanism of memory. Furthermore, because growth is expected to occur during weeks, the memory should be transmitted through cell divisions.

There is one mechanism that is known to serve long-term memory at the cell level. The mechanism is referred as 'cellular memory' and is based on regulation of gene expression through chromatin remodeling (Turner 2002; Ringrose and Paro 2004). Up- and down-regulation of genes corresponds to facultative changes between euchromatin and heterochromatin states, respectively. The cellular memory is regulated by epigenetic modifications of histones through chromatin-remodeling protein complexes, and the modification states can be transmitted through cell divisions (Lanzuolo and Orlando 2012). The cellular memory initially studied as the mechanism for cell linages to remember their developmental identities. For example, the homeobox (*Hox*) genes in *Drosophila* govern the body plan by controlling activity of the many downstream target genes based on memories of segment identities (Pearson et al. 2005). Expression of *Hox* genes is coordinated along the body axis by a cascade of maternal and zygotic transcription factors in early embryo (Akam 1987). Early signals decay shortly after the establishment of expression patterning in *Hox* gene, but the patterns are maintained throughout later developmental stages (Lanzuolo and Orlando 2012).

Interestingly, this type of gene regulation has been reported in vernalization, a major temperature-dependent flowering-time control in plants (Andrés and Coupland 2012; Song et al. 2012). Experiencing prolonged cold promotes floral induction in some group of plants, and vernalization refers either such treatment or response of plants (Chouard 1960; Salisbury 1963). It is considered as a mechanism that allow plants flower in spring following the exposure to low winter temperatures. A key regulator gene of vernalization identified in *A. thaliana* is *FLOWERING LOCUS C (FLC)* that encodes a MADS-box transcription factor (Sheldon et al. 1999; Michaels and Amasino 1999). *FLC* strongly suppresses floral induction by directly binding to the promoter regions of floral integrator genes that initiate the early stages of flowering processes. These floral integrator genes

include *FLOWERING LOCUS T (FT)* and *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1)* (Searle et al. 2006; Li et al. 2008). Therefore, when *FLC* shows high expression, plant remains in vegetative phase and shoot apical meristem continually produce leaves. After prolonged cold, *FLC* suppression is maintained even temperature becomes warmer, and then *A. thaliana* plants flower and complete seed production. Therefore, the suppressed state of *FLC* has been referred as 'winter memory' (Michaels and Amasino 2000; Amasino 2004).

In *A. thaliana* plants under the laboratory cold condition (usually 4 or 5 °C constant temperature), suppression of *FLC* transcription decrease to the lowest level within 2–3 weeks (Swiezewski et al. 2009), but it requires 4–8 week cold for the completer suppression of *FLC* that stably lasts after cold (Sheldon et al. 2000; Shindo et al. 2006). Once it is suppressed completely, the genomic region containing *FLC* form facultative heterochromatin through histone modification and it provides memory of suppression (Bastow 2004; De Lucia et al. 2008; Schmitz et al. 2008) (Fig. 13.5). The establishment of complete *FLC* suppression consists of multiple processes that function at different time scale during cold exposures (Song et al. 2012).

Within 2–3 weeks of cold exposure, transcripts of a non-coding RNA, cold induced long antisense intragenic RNA (COOLAIR) increase and reach to a peak of accumulation (Swiezewski et al. 2009) (Fig. 13.5). COOLAIR is an anti-sense transcript which transcription starts from the proximal downstream of poly-A site of *FLC* sense sequence (Swiezewski et al. 2009). After reaching the peak, the COOLAIR transcript decleases. In turn, a non-coding transcript from the sense strand of *FLC* first intron, COLDAIR (cold assisted intronic noncoding RNA), increases (Heo and Sung 2011). The transcript levels of COLDAIR reach the maximum level approximately one week after the peak of COOLAIR (Fig. 13.5).

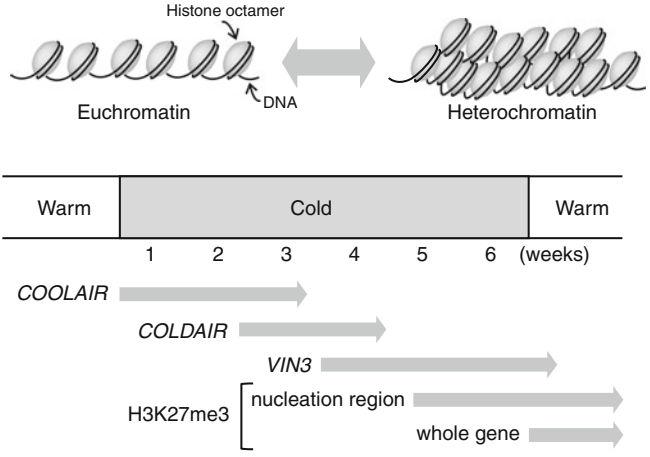



**Fig. 13.5** Facultative heterochromatin formation of *FLC* gene region in response to prolonged cold (vernalization), and timings of different processes involved in repression of *FLC*

The level of COLDAIR transcript returns to the pre-cold level by 4 weeks after cold (Heo and Sung 2011).

Repressive histone modification requires accumulations of *Polycomb* group (PcG) protein complexes in a localized region of the genome (Margueron and Reinberg 2011; Bemer and Grossniklaus 2012). Most prominent histone mark in the supressed *FLC* region is the trimethylation of histone H3 at lysine 27 (H3K27me3). Polycomb repressive complex 2 (PRC2) composed of subunits encoded by *VERNALIZATION2 (VRN2)*, *SWINGER (SWN)*, *FERTILIZATION-INDEPENDENT ENDOSPERM (FIE)*, and *MULTICOPY SUPPRESSOR OF IRA1 (MSI1)* are required to catalyze H3K27me3 at the *FLC* region (De Lucia et al. 2008). Curly leaf (CLF) acts partially redundantly with SWN in the regulation of *FLC* (Chanvivattana et al. 2004), and COLDAIR was reported to physically associate with CLF (Heo and Sung 2011). Therefore, COLDAIR is considered to play a role in the recruitment of PRC2 to *FLC* region (Heo and Sung 2011, but see Bemer and Grossniklaus 2012).

Prolong cold results in the association of plant homeodomain (PHD) proteins with the PRC2 to form PHD–PRC2 at a tightly localized nucleation region in the first exon and the first intron of *FLC* (Heo and Sung 2011). The genes encoding the PHD proteins include VERNALIZATION INSENSITIVE3 (VIN3), VERNALIZATION5 (VRN5) and VIN3-LIKE 2 (VIL2) (Sung and Amasino 2004; Wood et al. 2006; Greb et al. 2007; De Lucia et al. 2008). *VIN*3 expression is induced by cold exposure for 3 weeks and it becomes highly active by 6 week cold (Sung and Amasino 2004; Heo and Sung 2011) (Fig. 13.5). During the cold, the levels of H3K27me3 in the nucleation region show a quantitative rise that correlates with an increase in expression of *VIN3*. Cold-induced H3K27me3 in the nucleation region accumulates to a maximum level after 4–6 weeks exposure to cold (Angel et al. 2011) (Fig. 13.5).

After plants return to warm, the H3K27me3 spreads across the rest of the *FLC* locus according to the length of the cold period (Angel et al. 2011) (Fig. 13.5). The increase of H3K27me3 levels correlates with the level of *FLC* suppression (Angel et al. 2011). Substantial increase of H3K27me3 across the whole gene is required for stable silencing throughout the rest of development (Finnegan and Dennis 2007; De Lucia et al. 2008; Angel et al. 2011). There are additional three proteins that are required for stable repression of *FLC* after plants return to warm condition from the cold. Two of them are chromatin associating proteins, LIKE HETEROCHROMATIN PROTEIN 1 (LHP1) (Mylne et al. 2006; Sung et al. 2006; Turck et al. 2007) and VERNALIZATION1 (VRN1) (Levy et al. 2002), and the other is a type II protein arginine methyltransferase, *Arabidopsis* PROTEIN ARGININE METHYLTRANSFERASE 5 (ATPRMT5) (Schmitz et al. 2008). LHP1 has been shown to preferentially associates with H3K27me3 (Sung et al. 2006; Turck et al. 2007).

Overall, molecular biology on the regulation of a key flowering repressor, *FLC*, in *A. thaliana* revealed characteristics of the molecular machinery of temperature-dependent flowering time control. Although experiments were conducted under constant temperatures using small number of specific strains of *A. thaliana*, we can

still infer following three characteristics of molecular machinery at least in this specific example. (1) The sequential multiple process allows gene regulation to respond slowly over the scales of 4–8 weeks. (2) The response is quantitative at least for the range of 2–6 weeks of cold periods, and the mechanistic bases of quantification is the level of epigenetic histone marks in the gene region. (3) The response is mitotically stable and thus can be transmitted through cell divisions over multiple weeks. Whether or not the machinery serves as memory of seasonal temperature is required to be tested by the *in natura* systems biology.

## 13.5 Robust Functions of Temperature-Dependent Flowering Time Gene *in natura*

Using quantification methods of gene expression, such as real-time quantitative PCR, microarray and RNA-seq, it is possible to take a systems biological approach to evaluate function and control of genes or gene-networks in the natural habitats (Aikawa et al. 2010; Nagano et al. 2012). In the study *in natura*, for example, time-series data of gene expression obtained under the natural biotic and abiotic factors. Obviously, different set of questions from those in molecular genetics are asked in the approach. (1) How environmental signal is perceived and how noise is filtered out by organisms? (2) The regulatory system is robust enough to function under the natural range of whole set of fluctuations of abiotic and biotic factors? The analyses might add a novel interpretation especially in terms of robustness of the systems against existing noise fluctuations of diverse factors in the natural habitats.

The floral repressor *FLC* is conserved in other related species of *A. thaliana* (Lin et al. 2005; Wang et al. 2009; Aikawa et al. 2010; Kemi et al. 2013), and serves temperature-dependent flowering-time controls in both annual and perennial life cycle of plants (Albani and Coupland 2010). Although *A. thaliana FLC* suppression occurs quantitatively corresponding with the length of cold exposure (Sheldon et al. 2000), supression is maintained until the end of its life-span as an annual plant (Choi et al. 2009). Therefore, it is expected that seasonal control of *FLC* regulation ceases by the phase transition from vegetative growth to reproduction. In a perennial species, *Arabis alpina*, expression of *PERPETUAL FLOWERING 1 (PEP1)*, an ortholog of *FLC*, is downregulated by prolonged cold and this repression correlates with increases in H3K27me3 levels within the locus, as in *A. thaliana* (Wang et al. 2009). Interestingly, *PEP*1 repression is not stably maintained after exposure of the plants to warm temperatures, providing the molecular basis for the perennial lifecycle that requires repeated responses to the seasonal temperature cycles (Wang et al. 2009). It was an open question whether the regulatory mechanisms of *FLC* orthologs in perennial plants serve as a monitoring machinery of seasonal temperature.

Aikawa et al. (2010) conducted a time-series analyses of *FLC* expression in a natural population of a perennial *Arabidopsis*. Primary question asked in the study

was how long plants refer past temperatures to sense seasons? The study population of *A. halleri* subsp. *gemmifera* located near to Nishiwaki where plants experience seasonal temperature signals with enormous shorter-term noises (see Sect. 13.3). Real-time quantitative PCR was used to quantify the *FLC* ortholog, *AhgFLC*, for 96 timings at one-week interval over two years. Expression of *AhgFLC* showed a clear seasonal pattern (Fig. 13.6). Marked changes of the gene expression occurred autumn–winter–spring seasons, and the expression decreased in autumn and increased in spring following the temperature transitions. The expression levels were kept high from May to November.

Expression of *AhgFLC* successfully extract the seasonal pattern from noisy temperatures. There was also a lag between seasonal patterns of *AhgFLC* and temperature (Fig. 13.6). Flowering occurred using this lag when it became wormer but the level of *AhgFLC* was still low. These characteristics of *AhgFLC* patterns

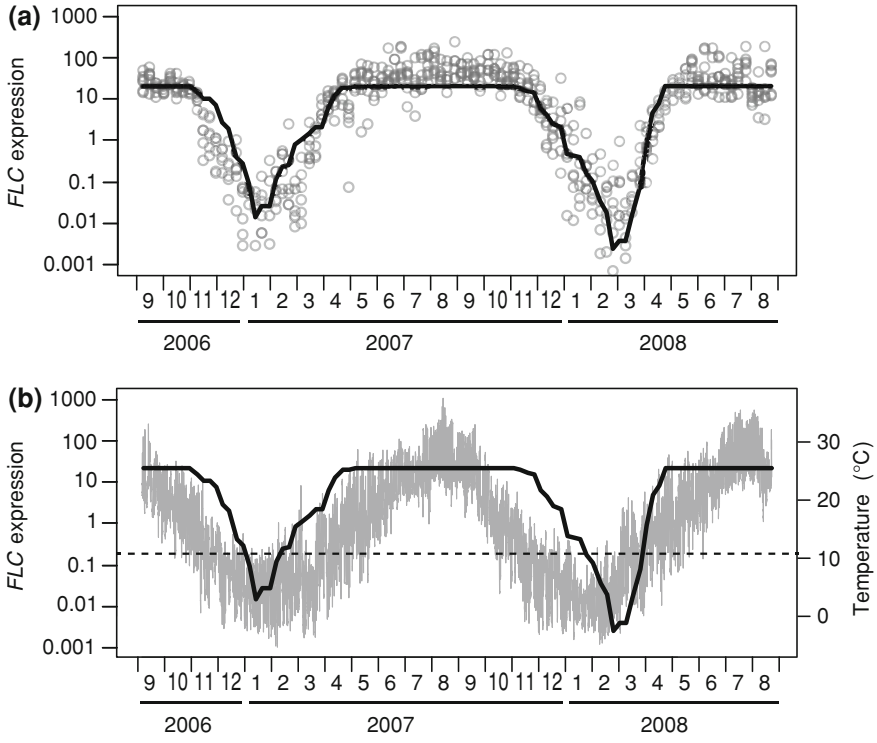



**Fig. 13.6** Time-series data of expression of *FLC* ortholog in *Arabidopsis halleri* subsp. *gemmifera* (*AhgFLC*) for two years in the natural habitat. In **a** actual data (6 plants × 96 time samplings) are shown by *open circles* and expected gene expression predicted by the most likely model was shown by *a thick line*. In **b**, the expected gene expression (*the thick line*) is compared with hourly mean temperatures (*a gray line*). The *FLC* expression is modeled to be determined by temperatures lower than a threshold temperature, 10.5 °C (*a broken line*), for past 42 days. The figure is produced using the data obtained by Aikawa et al. (2010)

are similar to what we found in SMA of temperatures (see Sect. 13.3). To infer how long of past temperature affected the gene expression, they conducted time series analyses using the chilling unit model. The chilling unit is the sum of temperatures lower than a certain threshold over a certain past periods. The most likely values turned out to be 10.5 °C threshold temperature and 42 days (6 weeks) past referring period (Aikawa et al. 2010). Surprisingly, the selected model, i.e. solely past 6 week temperature, explained 83 % of 2 years variation of *AhgFLC* expression. Referring periods were turned out to be important, and shorter and longer periods greatly reduced power of model explanation.

This estimate of 6 weeks referring periods of past temperature corresponds well to what we have expected for properties of season-detecting machinery solely by meteorological temperature data (Table 13.1). It also fit to the time scale of responses reported in the cellular memory mechanisms, i.e., regulation of gene expression through chromatin conformations (Table 13.1). Therefore, we consider that *FLC* and its homologs have been evolved as more general machinery to detect seasonal temperature.

Another importance of the study was that it was conducted in the natural conditions. In the natural habitats, factors other than temperatures also fluctuate largely. Actually samples analyzed in Aikawa et al. (2010) experienced flooding, snow cover and desiccation stress, but large part of *AhgFLC* regulation was simply dependent on past temperatures. This indicates that regulatory system of *AhgFLC* not only detects the seasonal temperature but also equip robustness that can function under complex natural conditions.

## 13.6 Revisiting Biometeorological Analyses

In biometeorology, flowering time has been analyzed using phenology data over decades (Schwartz 2003). Records of first flowering date (FFD) of a particular species or a set of species have been recorded every year (Fitter and Peat 1994; Tooke and Battey 2010). Because FFD is observed once a year, it requires data for nearly 20 years or longer to statistically test how environmental factors influence yearly variation of flowering time. Frequently practiced analyses were to correlate FFD against the indices that represent temperature variation between years. Effects of climate change on phenology have often been analyzed, and the global warming has reported to significantly shift spring phenology forward (Bradley et al. 1999; Penuelas and Filella 2001; Walther et al. 2002; Root et al. 2003).

The kind of study we would like to revisit here is those which aimed to predict FFD based on environments during proceeding weeks or months (e.g. Marletto et al. 1992; Censi and Ceschia 2000). The major question asked in these studies was which environmental factors and how long of proceeding periods affect variation of FFD. For example, Fitter et al. (1995) analyzed FFD of 243 species of flowering plants in central England over a period of 36 years (1954–1989). They found strong temperature effects on FFD, and 219 of the 243 species showed FFD

variation that were explicable largely by temperatures of the preceding months. FFD of early flowering species (January–April) mostly affected by temperature 1–2 months before flowering, and temperatures up to 4 months previously were important for summer flowering species (May onwards). Censi and Ceschia (2000) successfully modeled FFD of 57 species during 23 years (1960–1982) at central Italy using preceding temperatures and rainfall. The length of periods of environmental effects for most fit models was 20–38 days for the species that flowered between February and May and was 41–83 days (but mostly within 60 days) for the species that flowered between June and September, respectively. In other studies that modeled FFD in a single or multiple species, many found that the temperature records preceding 1–2 months to the flowering were critical dependent variables (Marletto et al. 1992; Rodríguez-Rajo et al. 2003; Roberts 2008; Luedeling et al. 2012).

Overall, although it varies between species and locations, 3–10 weeks of past temperatures turned out to be influence yearly variation of FFD (Table 13.1). One of conventional views is that the length of proceeding periods in which temperature affect FFD represents the length of developing process that is susceptible to temperature. A novel view is that plants refer past temperature to flower in a specific season in the noisy natural conditions, probably for equivalent periods that have been reported in the biometeorological analyses. Therefore, we should not take yearly variation of flowering date as a mealy passive developmental effect of preceding temperatures. It may include more active responses of plants that have been evolved to control flowering-time in the noisy temperature fluctuations in the natural habitats. Obviously, we need further combined studies of biometeorology, gene expression *in natura*, and flowering-time control mechanism of non-model species.

## 13.7 Context-Dependent Evolution of Long-Term Memory of Temperature

We found a corresponding property, i.e. long-term memory of past temperature, that are expected from patterns of temperature signals, as well as in molecular characteristics of vernalization response in *A. thaliana* and its relatives. We hypothesized that the correspondence is the result of evolution in the natural habitats. Because the property of seasonal temperature signals is rather general one that is experienced by majority of temperate plants, we expect parallel evolution to occur if the hypothesis is true. This is testable because vernalization has evolved multiple times in phylogenetically distinctive group of plants.

In cereals, different sets of non-orthologous genes are known to be involved in the vernalization compared those in *Arabidosis* and its relatives (Trevaskis et al. 2007). The vernalization responses of wheat (*Triticum aestivum*) and barley (*Holdeum vulgare*) are quantitative and require more than six weeks of cold for

saturated responses (Takahashi and Yasuda 1971; Flood and Halloran 1984). Vernalization response is controlled by the different MADS-box gene, *VRN1*, that encodes an *APETALA1*-like MADS box transcription factor (*FRUITFULL* is the most closely related gene in *A. thaliana*) (Yan et al. 2003; Preston and Kellogg 2006). Interestingly, *VRN1* expression is induced by prolonged cold and enhances inflorescence initiation by suppressing a downstream flowering repressor (Trevaskis et al. 2007). Furthermore, temperature sensitivity is likely a feature of many other MADS-Box genes (Hemming and Trevaskis 2011). In peach, several *SHORT VEGETATIVE PHASE (SVP)*-like MADS-box genes have been reported to be regulated by exposure to cold and to associate with bud dormancy (Bielenberg et al. 2008; Li et al. 2009, 2010). The repeated uses of MADS-box genes are supportive to our hypothesis, and future analysis should ask whether or not the regulatory system of these genes serve as a long-term memory of temperature to detect seasons.

Recent studies *in natura* indicate that robustness may be common among the gene regulation. Nagano et al. (2012) conducted time-series analyses of whole transcriptome of rice (*Oryza sativa*) using microarray techniques. Although they sampled plants outside where multiple environmental factors fluctuated with a large complexity, expression patterns of 97 % of expressed genes were successfully modeled by single environmental factors. In near future, transcriptome analyses under natural conditions will rapidly increase using diverse plants by the introduction of RNA-seq techniques. We expect that regulatory mechanism of a gene is generally sensitive only to a few particular environmental factors and robust against noises that exist in the natural conditions. The modeling of transcriptome *in natura* will allow us to understand how gene regulatory mechanisms provide robustness of gene functions under natural noisy conditions.

## References

Aikawa S, Kobayashi MJ, Satake A, Shimizu KK, Kudoh H (2010) Robust control of the seasonal expression of the *Arabidopsis FLC* gene in a fluctuating environment. Proc Natl Acad Sci USA 107:11632–11637

Ahas R, Jaagus J, Aasa A (2000) The phenological calendar of Estonia and its correlation with mean air temperature. Int J Biometeorol 44:159–166

Akam M (1987) The molecular basis for metameric pattern in the *Drosophila* embryo. Development 101:1–22

Albani MC, Coupland G (2010) Comparative analysis of flowering in annual and perennial plants. Curr Top Dev Biol 91:323–348

Amasino R (2004) Vernalization, competence, and the epigenetic memory of winter. Plant Cell 16:2553–2559

Anderson JV, Gesch RW, Jia Y, Chao WS, Horvath DP (2005) Seasonal shifts in dormancy status, carbohydrate metabolism, and related gene expression in crown buds of leafy spurge. Plant Cell Environ 28:1567–1578

Andrés F, Coupland G (2012) The genetic basis of flowering responses to seasonal cues. Nat Rev Genet 13:627–639

Angel A, Song J, Dean C, Howard M (2011) A Polycomb-based switch underlying quantitative epigenetic memory. Nature 476:105–108

Bastow R et al (2004) Vernalization requires epigenetic silencing of *FLC* by histone methylation. Nature 427:164–167

Battey NH (2000) Aspects of seasonality. J Exp Bot 51:1769–1780

Bemer M, Grossniklaus U (2012) Dynamic regulation of Polycomb group activity during plant development. Curr Opin Plant Biol 15:523–529

Bernier G (1988) The control of floral evocation and morphogenesis. Annu Rev Plant Physiol 39:175–219

Bielenberg DG, Wang Y, Li Z, Zhebentyayeva T, Fan S, Reighard GL, Scorza R, Abbott AG (2008) Sequencing and annotation of the evergrowing locus in peach [*Prunus persica* (L.) Batsch] reveals a cluster of six MADS-box transcription factors as candidate genes for regulation of terminal bud formation. Tree Genet Genomes 4:495–507

Bradley NL, Leopold AC, Ross J, Huffaker W (1999) Phenological changes reflect climate change in Wisconsin. Proc Natl Acad Sci USA 96:9701–9704

Censi CA, Ceschia M (2000) Forecasting of the flowering time for wild species observed at Guidonia, central Italy. Int J Biometeorol 44:88–96

Chanvivattana Y, Bishopp A, Schubert D, Stock C, Moon Y-H, Sung ZR, Goodrich J (2004) Interaction of Polycomb-group proteins controlling flowering in *Arabidopsis*. Development 131:5263–5276

Chew YH, Wilczek AM, Williams M, Welch SM, Schmitt J, Halliday K (2012) An augmented *Arabidopsis* phenology model reveals seasonal temperature control of flowering time. New Pytol 194:654–665

Choi J, Hyun Y, Kang MJ, In Yun H, Yun JY, Lister C, Dean C, Amasino RM, Noh B, Noh YS, Choi Y (2009) Resetting and regulation of *FLOWERING LOCUS C* expression during *Arabidopsis* reproductive development. Plant J 57:918–931

Chouard P (1960) Vernalization and its relations to dormancy. Annu Rev Plant Physiol 11:191–238

Cleland EE, Chuine I, Menzel A, Mooney HA, Schwartz MD (2007) Shifting plant phenology in response to global change. Trends Ecol Evolut 22:357–365

De Lucia F, Crevillen P, Jones AME, Greb T, Dean C (2008) A PHD-polycomb repressive complex 2 triggers the epigenetic silencing of *FLC* during vernalization. Proc Natl Acad Sci USA 105:16831–16836

Fitter AH, Peat HJ (1994) The ecological flora database. J Ecol 82:415–425

Fitter AH, Fitter RSR, Harris ITB, Williamson MH (1995) Relationships between first flowering date and temperature in the flora of a locality in central England. Func Ecol 9:55–60

Finnegan EJ, Dennis ES (2007) Vernalization-induced trimethylation of histone H3 lysine 27 at *FLC* is not maintained in mitotically quiescent cells. Curr Biol 17:1978–1983

Flood RG, Halloran GM (1984) Basic development rate in spring wheat. Agro J 76:260–264

Greb T, Mylne JS, Crevillen P, Geraldo N, An H, Gendall AR, Dean C (2007) The PHD finger protein VRN5 functions in the epigenetic silencing of *Arabidopsis FLC*. Curr Biol 17:73–78

Hemming MN, Trevaskis B (2011) Make hay when the sun shines: the role of MADS-box genes in temperature-dependant seasonal flowering responses. Plant Sci 180:447–453

Heo JB, Sung S (2011) Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. Science 331:76–79

Izawa T (2012) Physiological significance of the plant circadian clock in natural field conditions. Plant Cell Environ 35:1729–1741

Jiménez S, Reighard GL, Bielenberg DG (2010) Gene expression of *DAM5* and *DAM6* is suppressed by chilling temperatures and inversely correlated with bud break rate. Plant Mol Biol 73:157–167

Jones PD, New M, Parker DE, Martin S, Rigor IG (1999) Surface air temperature and its changes over the past 150 years. Rev Geophys 37:173–199

Kemi U, Niittyvuopio A, Toivainen T, Pasanen A, Quilot-Turion B, Holm K, Lagercrantz U, Savolainen O, Kuittinen H (2013) Role of vernalization and of duplicated *FLOWERING LOCUS C* in the perennial *Arabidopsis lyrata*. New Phytol 197:323–335

Lang A (1952) Physiology of flowering. Ann Rev Plant Physiol 3:265–306

Lanzuolo C, Orlando V (2012) Memories from the Polycomb group proteins. Annu Rev Gen 46:561–589

Levy YY, Mesnage S, Mylne JS, Gendall AR, Dean C (2002) Multiple roles of *Arabidopsis* VRN1 in vernalization and flowering time control. Science 297:243–246

Lee R (1970) Latitude and photoperiodism. Arch Met Geoph Biokl Ser B 18:325–332

Lewis P (1960) The use of moving average in the analysis of time-series. Weather 15:121–126

Li D, Liu C, Shen L, Wu Y, Chen H, Robertson M, Helliwell CA, Ito T, Meyerowitz E, Yu H (2008) A repressor complex governs the Integration of flowering signals in *Arabidopsis*. Dev Cell 15:110–120

Li Z, Reighard GL, Abbott AG, Bielenberg DG (2009) Dormancy-associated MADS genes from the *EVG* locus of peach [*Prunus persica* (L.) Batsch] have distinct seasonal and photoperiodic expression patterns. J Exp Bot 60:3521–3530

Lin S, Wang J, Poon S, Su C, Wang S, Chiou T (2005) Differential regulation of expression by vernalization *FLOWERING LOCUS C* in cabbage and *Arabidopsi*s. Plant Physiol 137:1037–1048

Luedeling E, Kunz A, Blanke MM (2012) Identification of chilling and heat requirements of cherry trees—a statistical approach. Int J Biometeorol. doi:10.1007/s00484-012-0594-y

Manitašević S, Dunđerski J, Matić G, Tucić B (2007) Seasonal variation in heat shock proteins Hsp70 and Hsp90 expression in an exposed and a shaded habitat of *Iris pumila*. Plant Cell Env 30:1–11

Marletto V, Branzi GP, Sirotti M (1992) Forecasting flowering dates of lawn species with air temperature: Application boundaries of the linear approach. Aerobiologia 8:75–83

Mayrhofer S, Teuber M, Zimmer I, Louis S, Fischbach RJ, Schnitzler J-P (2005) Diurnal and seasonal variation of isoprene biosynthesis-related genes in grey poplar leaves. Plant Physiol 139:474–484

Merquiol E, Pnueli L, Cohen M, Simovitch M, Rachmilevitch S, Goloubinoff P, Kaplan A, Mittler R (2002) Seasonal and diurnal variations in gene expression in the desert legume *Retam raetam*. Plant Cell Env 25:1627–1638

Margueron R, Reinberg D (2011) The Polycomb complex PRC2 and its mark in life. Nature 469:343–349

Mylne JS, Barrett L, Tessadori F, Mesnage S, Johnson L, Bernatavichute YV, Jacobsen SE, Fransz P, Dean C (2006) LHP1, the *Arabidopsis* homologue of heterochromatin protein 1, is required for epigenetic silencing of *FLC*. Proc Natl Acad Sci USA 103:5012–5017

Michaels SD, Amasino RM (1999) *FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering. Plant Cell 11:949–956

Michaels SD, Amasino RM (2000) Memories of winter: vernalization and the competence to flower. Plant Cell Environ 23:1145–1153

Munguia-Rosas MA, Ollerton J, Parra-Tabla V, De-Nova JA (2011) Meta-analysis of phenotypic selection on flowering phenology suggests that early flowering plants are favoured. Ecol Letters 14:511–521

Nagano AJ, Sato Y, Mihara M, Antonio BA, Motoyama R, Itoh H, Nagamura Y, Izawa T (2012) Deciphering and prediction of transcriptome dynamics under fluctuating field conditions. Cell 51:1358–1369

Pearson JC, Lemons D, McGinnis W (2005) Modulating *Hox* gene functions during animal body patterning. Nat Rev Gen 6:893–904

Penfield S (2008) Temperature perception and signal transduction in plants. New Phytol 179:615–628

Peñulelas J, Filella I (2001) Responses to a warming world. Science 294:793–794

Prescott JA, Collins JA (1951) The lag of temperature behind solar radiation. Quart J Roy Meteorol Soc 77:121–126

Preston JC, Kellogg EA (2006) Reconstructing the evolutionary history of paralogous *APETALA1/FRUITFULL*-like genes in grasses (Poaceae). Genetics 174:421–437

Quintana-Murci L, Alcaïs A, Abel L, Casanova J-L (2007) Immunology *in natura*: clinical, epidemiological and evolutionary genetics of infectious diseases. Nat Immunol 8:1165–1171

Richards CL, Hanzawa Y, Katari MS, Ehrenreich IM, Engelmann KE, Purugganan MD (2009) Perspective on ecological and evolutionary systems biology. Annu Plant Rev 35:331–351

Richards CL, Rosas U, Banta J, Bhambhra N, Purugganan MD (2012) Genome-wide patterns of *Arabidopsis* gene expression in nature. PLoS Genet 8:e1002662

Ringrose L, Paro R (2004) Epigenetic regulation of cellular memory by the poycomb and trithorax group proteins. Annu Rev Gen 38:413–443

Roberts AMI (2008) Exploring relationships between phenological and weather data using smoothing. Int J Biometeorol 53:463–470

Rodríguez-Rajo FJ, Frenguelli G, Jato MV (2003) Effect of air temperature on forecasting the start of the *Betula* pollen season at two contrasting sites in the south of Europe (1995–2001). Int J Biometeorol 47:117–125

Root TL, Price JT, Hall KR, Schneider SH, Rosenzweig C, Pounds JA (2003) Fingerprints of global warming on wild animals and plants. Nature 421:57–60

Salisbury FB (1963) The flowering process. Pergamon, Oxford

Schmitz RJ, Sung S, Amasino RM (2008) Histone arginine methylation is required for vernalization-induced epigenetic silencing of *FLC* in winter-annual *Arabidopsis thaliana*. Proc Natl Acad Sci USA 105:411–416

Schwartz MD (2003) Phenology: an integrative environmental science. Kluwer Academic Publishers, Dordrecht

Shimizu KK, Kudoh H, Kobayashi MJ (2011) Plant sexual reproduction during climate change: gene function in natura studied by ecological and evolutionary systems biology. Ann Bot 108:777–787

Searle I, He Y, Turck F, Vincent C, Fornara F, Kröber S, Amasino RA, Coupland G (2006) The transcription factor FLC confers a flowering response to vernalization by repressing meristem competence and systemic signaling in *Arabidopsis*. Genes Dev 20:898–912

Sheldon CC, Burn JE, Perez PP, Metzger J, Edwards JA, Peacock WJ, Dennis ES (1999) The *FLF* MADS box gene: a repressor of flowering in *Arabidopsis* regulated by vernalization and methylation. Plant Cell 11:445–458

Sheldon CC, Rouse DT, Finnegan EJ, Peacock WJ, Dennis ES (2000) The molecular basis of vernalization: the central role of *FLOWERING LOCUS C* (*FLC*). Proc Natl Acad Sci USA 97:3753–3758

Shindo C, Lister C, Crevillen P, Nordborg M, Dean C (2006) Variation in the epigenetic silencing of *FLC* contributes to natural variation in *Arabidopsis* vernalization response. Genes Dev 20:3079–3083

Simpson G, Dean C (2002) *Arabidopsis*: the Rosetta stone of flowering time? Science 296:285–289

Song J, Angel A, Howard M, Dean C (2012) Vernalization—a cold-induced epigenetic switch. J Cell Sci 125:1–9

Sparks TH, Jeffree EP, Jeffree CE (2000) An examination of the relationship between flowering times and temperature at the national scale using long-term phenological records from the UK. Int J Biometeorol 44:82–87

Sung S, Amasino RM (2004) Vernalization in *Arabidopsis thaliana* is mediated by the PHD finger protein VIN3. Nature 427:159–164

Sung S, He Y, Eshoo TW, Tamada Y, Johnson L, Nakahigashi K, Goto K, Jacobsen SE, Amasino RM (2006) Epigenetic maintenance of the vernalized state in *Arabidopsis thaliana* requires LIKE HETEROCHROMATIN PROTEIN 1. Nat Genet 38:706–710

Swiezewski S, Liu F, Magusin A, Dean C (2009) Cold-induced silencing by long antisense transcripts of an *Arabidopsis* Polycomb target. Nature 462:799–802

Takahashi R, Yasuda S (1971) Genetics of earliness and growth habit in barley. In: Nilan RA (ed) Barley genetics II (proceedings of the second international barley genetics symposium). Washington State University Press, Pullman, pp 388–408

Tooke F, Battey NH (2010) Temperate flowering phenology. J Exp Bot 61:2853–2862

Trenberth KE (1983) What are the seasons? Bull Amer Meteorol Soc 64:1276–1282

Trevaskis B, Hemming MN, Dennis ES, Peacock WJ (2007) The molecular basis of vernalization-induced flowering in cereals. Trends Plant Sci 12:352–357

Turck F, Roudier F, Farrona S, Martin-Magniette M-L, Guillaume E, Buisine N, Gagnot S, Martienssen RA, Coupland G, Colot V (2007) *Arabidopsis* TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. PLoS Genet 3:e86

Turner BM (2002) Cellular memory and the histone code. Cell 111:285–291

Wahba G (1990) Spline model for observation data. Society for Industrial and Applied Mathematics, Philadelphia

Walther G-R, Post E, Convey P, Menzel A, Parmesan C, Beebee TJC, Fromentin J-M, Hoegh-Guldberg O, Bairlein F (2002) Ecological response to recent climate change. Nature 416:389–395

Wang R, Farrona S, Vincent C, Joecker A, Schoof H, Turck F, Alonso-Blanco C, Coupland G, Albani MC (2009) PEP1 regulates perennial flowering in *Arabis alpina*. Nature 459:423–427

Wilczek AM, Roe JL, Knapp MC, Cooper MD, Lopez-Gallego C, Martin LJ, Muir CD, Sim S, Walker A, Anderson J, Egan JF, Moyers BT, Petipas R, Giakountis A, Charbit E, Coupland G, Welch SM, Schmitt J (2009) Effects of genetic perturbation on seasonal life history plasticity. Science 323:930–934

Wilczek AM, Burghardt LT, Cobb AR, Cooper MD, Welch SM, Schmitt J (2010) Genetic and physiological bases for phenological responses to current and predicted climates. Philos Trans R Soc B 365:3129–3147

Withrow RB (1959) A kinetic analysis of photoperiodism. In: Withrow RB (ed) Photoperiodism and related phenomena in plants and animals. American Association for Advancement of Science, Washington, pp 439–471

Wood CC, Robertson M, Tanner G, Peacock WJ, Dennis ES, Helliwell CA (2006) The *Arabidopsis thaliana* vernalization response requires a polycomb-like protein complex that also includes VERNALIZATION INSENSITIVE 3. Proc Natl Acad Sci USA 103:14631–14636

Yan L, Loukoianov A, Tranquilli G, Helguera M, Fahima T, Dubcovsky J (2003) Positional cloning of the wheat vernalisation gene *VRN*1. Proc Natl Acad Sci USA 100:6263–6268

Yanovsky MJ, Kay SA (2003) Living by the calendar: how plants know when to flower. Nat Rev Mol Cell Biol 4:265–275

# Chapter 14
# Current Approaches in Spatial Genetics

**V. Montano, A. Eriksson, A. Manica and Y. Moodley**

**Abstract** Spatial and landscape genetics have been recently undergone rapid advancement, although their theoretical development can be dated back to the origin of theoretical population genetics. The extent at which the spatial distribution of a population influences the short- and long-term population dynamics can be investigated by the application of several genetic and statistical approaches. In the present chapter, we offer a perspective on the different aims in spatial and landscape genetic researches and the tools to achieve them, reviewing some of the most important software that have been implemented to date.

## 14.1 Introduction

The work of an evolutionary biologist can be compared to that of a historian trying to understand the past processes that led populations and species to their current biological status, whether that status is defined ecologically, morphologically, adaptively, phylogenetically, or genetically. In most cases, the evolutionary events are not experimentally reproducible. Apart from the occasional fossil, which can be thought of as a surviving historical document, the evolutionary path through time can only be observed indirectly, in the resulting variation of present-day organisms observed in nature. Computer simulations provide a way to investigate the mechanisms of evolution, and to estimate the most likely evolutionary

V. Montano (✉) · Y. Moodley
Integrative Biology and Evolution, University of Veterinary Medicine Vienna,
Savoyanstaße 1A, A-1160 Vienna, Austria
e-mail: valeria.montano@vetmeduni.ac.at

A. Eriksson · A. Manica
Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK

P. Pontarotti (ed.), *Evolutionary Biology: Exobiology and Evolutionary Mechanisms*,
DOI: 10.1007/978-3-642-38212-3_14, 

scenarios on the basis of our current knowledge, but require a deep understanding of the mechanisms underlying biological phenomena.

The field of population genetics concerns how dynamic evolutionary forces affect allele frequencies in groups of individuals within a species. The seminal framework developed by the theoretical population geneticists Fisher, Wright and Haldane represented a Modern Synthesis of the principles of Mendelian inheritance with Darwinian selection. Since then, population genetics has assumed a central role in biology, exploring and developing ever more sophisticated mathematical and empirical tools to draw conclusions from the distribution of genetic polymorphisms, making use of parallel technological advances in the generation of high-resolution genetic data.

The ability to construct testable evolutionary scenarios that simulate the distribution of genetic variation has become an essential approach to the understanding of the dynamics that have led to the natural populations we observe today. To achieve this objective, modelling demographic histories through time and space is a major goal for the inference of both neutral and selective processes in population genetics. Intuitively, the spatial dimension, quantified as the habitat where natural populations live and evolve, has an influence on the observed distribution of allele frequencies. Its role can be interpreted in terms of purely geographic distances (spatial genetics; see Guillot et al. 2009 for a review) or as the dimensions that make up an ecological niche, thus defining the ecological differences within the habitat range of a species (landscape genetics, originally defined by Manel et al. 2003).

The analysis of the spatial distribution of genetic variation can be carried out by exploring the patterns in the observed data or by modelling scenarios which take into account the spatial dimension to simulate and predict the distribution of genetic variability. The former analysis may be model free or model based, the latter of which requires the explicit definition of demographic models for population expansion, migration, and subdivision. The study and interpretation of the role of space in genetics have become increasingly common during the last decade with the development of appropriate statistical tools, some of which will be summarised and discussed in the present chapter (see also Guillot et al. 2009 for a review).

## 14.2 Exploring Population Genetic Structure and Spatial Patterns

Genetic populations were formally defined by Godfrey Harold Hardy, an English mathematician, and Wilhelm Weinberg, a German physician, who independently developed analogous models in 1908. Their formulation describes a panmictic, randomly reproducing group of individuals of constant population size across generations, leading to constant allele frequencies through time—the so-called Hardy–Weinberg equilibrium. Although the conditions described by this formulation are rarely found in nature, violations of the demographic assumptions

mostly lead to negligible deviations from Hardy–Weinberg equilibrium, making this simple model a useful tool for detecting candidate loci for non-neutral, evolutionary processes (which, conversely, readily lead to significant deviations from equilibrium). Moreover, this model represents a convenient mathematical starting point for the theoretical exploration of all the processes that actually occur in natural populations (genetic drift, subdivision, migration, and natural selection).

The population is a key concept in both ecology and evolution, although its definition in these two fields differs. In ecology, the demographic cohesion and the ecological homogeneity of the habitat occupied (*deme*) by individuals are emphasised. A set of subdivided populations occupying different demes interconnected by gene flow is referred to as a *metapopulation* (Levins 1969). On the other hand, evolutionary definitions focus rather on the reproductive cohesion of individuals which must show reproductive continuity over the generations (Waples and Gaggiotti 2006). Beside the theoretical definitions, quantifying population subdivision is a non-trivial problem that depends mostly on local effective population sizes and the intensity of the gene flow connecting demes (Waples and Gaggiotti 2006).

In population genetic studies, populations are often defined a priori on the basis of biological criteria (e.g. sampling locations, ethnic affiliation) and the structure so determined can be explored using a number of basic genetic statistics. Fixation indices ($F_{ST}$, $F_{IT}$, $F_{IS}$) were first proposed by Wright (1949) as useful statistics to quantify genetic differentiation among populations based on the partitioning of genetic variance among different levels, i.e. individuals (*I*), demes (*S*), and the entire population (*T*). Weir and Cockerham (1984) later formalised this approach by recasting the fixation indices in terms of population demographic parameters. Since this seminal work, a number of new descriptors of genetic differentiation have been suggested, with the most popular being $G'_{ST}$ ($G_{ST}$ standardised by the theoretical maximum $G_{ST}$ at a given locus based on heterozygosity; Hedrick 2005), *D,* and $D_{EST}$ (fraction of allelic variation among populations and its bias correction; Jost 2007, 2008). While the magnitude of these indices has a clear biological interpretation in terms of the degree of divergence among populations and the level of inbreeding within populations, testing their significance against a null hypothesis of no population structure can be challenging. Bootstrap approaches are often employed, but the resulting *p* values may be meaningless if the sample sizes are very large (Jost 2007). $G'_{ST}$ and $D_{EST}$ were shown to be good estimators of population differentiation and are in strong agreement in ranking populations (8 % of discrepancies on observed datasets; Heller and Siegismund 2009), while raw $G_{ST}$ estimates appear mathematical inconsistent (Jost 2007).

When a priori population assignment is not available or when it is desirable to explore the population structure of a genetic dataset without strict constraints, the distribution of allele frequencies can be ordered in *K* populations through the application of clustering methods. Usually relying on the assumption that individual genotypes are composed of unlinked loci, Bayesian clustering methods estimate the likelihood of a specific genotype of belonging to one or another population as the joint probability for each allele at a locus of belonging to that

particular population. The higher the difference in allele frequency distribution among different populations, the easier for the clustering algorithms to correctly assign each individual to its 'actual' population. The results of such algorithms are usually graphically visualised by plotting each individual as a line partitioned into coloured segments that are proportional to the probabilities of assignment of the individual to different clusters (see Distruct software; Rosenberg et al. 2002).

Bayesian or multivariate clustering approaches have been implemented in several genetic software packages during the last decade (e.g. STRUCTURE, BAPS, GENELAND, PARTITION, GENECLUST, TESS, DAPC; Falush et al. 2003, 2007; Corander and Marttinen 2006; Cheng et al. 2013; Guillot et al. 2005; Dawson and Belkhir 2001; François et al. 2006; Durand et al. 2009; Jombart et al. 2010). The use of statistical clustering methods to infer the best population structure of a dataset allows optimising the number of populations (or demes) contained in an observed sample, although the optimal number of clusters may change according to the model specified (François and Durand 2010). Ideally, larger numbers of individuals and higher resolution genetic data are needed when population genetic structuring is weak (François and Durand 2010).

In spatial genetics, the emphasis is placed in understanding whether population structure can be explained in terms of the spatial distribution of the samples. Some samples may not be spatially structured as is the case where assortative mating occurs within the same habitat. A well-known example of this in humans is the Indian caste system, which led to layered genetic populations within the same area due to social stratification (Bamshad et al. 2001).

A common approach is the Mantel's test, which evaluates the correlation of a matrix of population differentiation (see above) or genetic distances (see Chap. 6 in Templeton 2006 for a discussion on the methods for genetic distances and the conceptual difference with population differentiation approaches) with a matrix of geographic distances among demes using permutations to generate an empirical p-value for the observed coefficient of correlation between the matrices (Mantel 1967).

For the same purpose, some methods for structure analysis, such as GENECLUST, TESS and BAPS, allow the inclusion of individual geographic coordinates as a prior, so that geographic neighbours are more likely to be assigned to the same cluster. A different approach is used in GENELAND where the spatial units are represented by cells and the number of cells and individuals per cell can be explicitly defined.

The performance of spatial and nonspatial clustering algorithms has been explored with simulations. The migration model (see next paragraph) and diffusion processes that originally led to the spatial pattern have been found to potentially affect the number of clusters and the assignment of individuals, with different software resulting in contrasting assignments under the same circumstances. In general, Bayesian approaches have been shown to underperform when population structure results from a stepping stone model of migration, when the isolation by distance (IBD) among locations is very high or when the $F_{ST}$ values are $\leq 0.01$ (Schwartz and McKelvey 2009; Frantz et al. 2009; Jombart et al. 2010), while multivariate methods tend to be somewhat more robust under these conditions

(Jombart et al. 2010). STRUCTURE, TESS and GENECLUST were shown to be reliable when dealing with island models and with spatial overlapping clusters, while GENELAND appears to be a better option for detecting recent isolation without spatial continuity (Chen et al. 2007). These differences are due to the specific model considered by each package and highlight the importance of comparing the outcome of multiple software packages when investigating a given dataset.

The methods discussed, thus far, are useful tools for the understanding of the dynamics driving the evolution of a species, but they do not represent an explicit hypothesis test, since different evolutionary histories may shape the same distribution of genetic variation. Patterns of IBD (Wright 1943), the presence of spatial discontinuities due to natural barriers or even habitat fragmentation, could produce a significant correlation of genetics to geography. However, the most useful information provided by the multivariate approach (DAPC) relies in the clusters' coordinates onto the discriminant components. This information indeed shows the reciprocal positions of the inferred groups in the variance space and patterns can suggest the presence of specific migration schemes, potentially offering the opportunity of formulating evolutionary hypotheses (Jombart et al. 2010; Montano et al. 2013).

## 14.3 Analysis of Spatial Gradients of Genetic Variation

When individuals are scattered along a spatial gradient with steep but not discrete genetic subdivision, the performance of clustering approaches may be strongly influenced by the sampling scheme, and could lead to contradictory results (Schwartz and McKelvey 2009). Although the difference between gradients and clusters is more of a grey zone than a clear dichotomy and clusters are an appealing descriptor in population genetics, clustering methods are not always suitable to the study of spatial or landscape genetics (Schwartz and McKelvey 2009). The presence of genetic gradients requires tools which treat geographic space as a continuum (see Barton et al. 2013 for a theoretical overview).

A compromise in Bayesian cluster analysis is the implementation of admixture models, which can be spatially (BAPS, TESS) or non-spatially explicit (STRUCTURE). These approaches consider the allele frequency distribution as organised in a gradient or cline where the alleles composing an individual genotype can be assigned to different parental populations on the basis of their correlation. As such, the retrieved groupings should not be referred to as 'clusters' (François and Durand 2010). The addition of spatial information in admixture models allows the consideration of the spatial autocorrelation of the allele frequencies (Sokal and Oden 1978; François and Durand 2010).

Spatial autocorrelation analysis is a valuable approach to the exploration of the continuities as it estimates the degree of genetic affinity among samples depending on their geographic distance. A statistic often used to quantify spatial autocorrelation is the Moran Index, which can either be positive or negative depending on

the tendency of a variable or a set of variables to show more similar or more different allele frequencies in geographically closer entities.

Spatial autocorrelation, therefore, offers more insight into the interpretation of spatial genetic patterns compared to other approaches. An interesting development that summarizes the spatial information contained in the Moran Index and the distribution of the genetic variance among localities is the spatial principal component analysis (sPCA), which estimates genetic similarities and dissimilarities of each entity with its spatial neighbours based on a scheme of spatial weights (Jombart et al. 2008). Unlike cluster analyses, the Moran Index and sPCA explore spatial to genetic distance relationships by treating space as a continuum, without focus on revealing the presence of genetic barriers or units. Another interesting development in the exploration of spatial gradients is spatial ancestry analysis (SPA), which models allele frequencies along a two-dimensional gradient assuming that their distribution is continuous (Yang et al. 2012). Despite its simplified model, SPA may allow reconstructing the geographic gradient of allele frequencies precisely enough to locate an individual's geographic origin from its observed genotype or vice versa. An analogous approach was developed for microsatellite loci by Amos and Manica (PNAS 2006).

## 14.4 Migration Models and Explicit Spatial Models of Evolution

Thus far, we have presented some useful approaches for the exploration and description of observed spatial patterns in a given genetic dataset. However, they do not test explicit hypothesis about the processes that underlie the demographic history of populations. In order to do so, the distribution of genetic variation needs to be explored under explicit spatial scenarios.

The spatial dimension is intuitively (and simply) introduced into a demographic model by taking into account more than one population or deme connected to each other through a migration matrix. Wright (1943) was first to explore the properties of a model with spatial subdivision, introducing his "island model" along with the above-mentioned concept of IBD (see also Kimura and Weiss 1964; Malecot 1973; Barbujani 1987; Slatkin 1993). Here, populations were considered as islands of approximate random mating, except for a portion of migrating individuals, which were fully connected in a complete migration matrix. From this, Kimura (1964) derived his "stepping-stone model", where the populations interchanged migrants only if they were located in adjacent positions along the matrix. In this model, the effective population sizes of the demes were considered constant as they tended to infinity, so that the portion of incoming or outgoing migrating individuals did not affect the population size of the demes.

Discriminating the evolutionary dynamics that led to specific patterns of distribution of observed genetic variation requires a priori definition of explicit

models. Then, once different evolutionary hypotheses are formulated, they must be parameterised and implemented in software that simulates the evolution of the genetic data under those predefined scenarios. In this way, it is possible to compare and evaluate the fit of the simulated data with the observed data in order to choose which among the tested models best describes the evolutionary history of the analysed populations. Model comparison can be achieved either by using a full-likelihood approach, or more recently by an approximate Bayesian computation (ABC, Beaumont et al. 2002) framework which allows the consideration of more complex demographic models for which the likelihood cannot be formally defined.

Several coalescence-based software packages, such as IMa2, Migrate-n and Lamarc (Beerli and Palczewski 2010; Kuhner and Smith 2007), estimate the likelihood of sets of genetic data in subdivided populations (see Box 1 for a brief summary on coalescent theory). Although all these packages do not explicitly take into account the extent of spatial subdivision in quantitative terms, they work on defined genetic demes as discrete neighbouring units that exchange migrants.

An overarching problem with demographic inference is that of dimensionality, where the amount of data needed to accurately infer the posterior distributions of parameters increases exponentially as more parameters are added to the model. Furthermore, each coalescence-based approach has its specific limitations. While IMa2 is able to distinguish between common ancestry and gene flow, it only allows up to 10 populations and does not allow a user-defined migration matrix to compare different subdivision models. On the other hand, Migrate-n does not include population divergence, thus treating the genetic structure as stable through time. Lamarc (which includes Migrate-n and other software) has recently implemented a user-defined population divergence scheme, allowing the comparison of different models of divergence and migration. However, all three coalescence-based packages consider only a simple stable or exponential demographic model of growth.

Other software packages (e.g. simcoal, MSMS, easypop; Laval and Excoffier 2004; Ewing and Hemirsson 2010; and Balloux 2001) are able to simulate genetic data for multiple subdivided populations under complex demographic scenarios and model choice can be conducted by ABC. Full-likelihood approaches, such as IMa2 and related software, are statistically more consistent than ABC, which risks introducing an arbitrary factor when selecting the summary statistics for model comparison (Aeschbacher et al. 2012), but are definitely less versatile.

An important contribution to landscape genetic simulations was made by the developers of SPLATCHE (Ray et al. 2010), which models complex spatial demographic scenarios on an explicit lattice of squares taking into account both geographic and ecological information. The approach implemented in SPLATCHE is divided into two main steps, the first of which simulates the forward ‘ecological history’ of the demes included in the model. These simulations are aimed at estimating the carrying capacities of the demes across generations resulting in a non-genetic estimate of population size—an important parameter in any demographic simulation. Using the forward simulations as data, coalescent backward simulations are carried out to reproduce the observed genetic variability.

This kind of model requires a complex and probably challenging user definition and the parameters used to define the landscape of the species under study are crucial. Furthermore, it is currently limited to a single stepping stone migration scheme (Ray and Excoffier 2009).

Nevertheless, the simulation of demographic histories with explicit spatial models may have very important implications for the detection of events of positive selection. It has been shown that long-term range expansions can lead to patterns of genetic variability close to those expected under positive selection, a process known as gene surfing (Excoffier and Ray 2008). Distinguishing neutral spatial demographic processes from selection remains one of the most important and challenging aspects in the study of species evolution (Crisci et al. 2012).

## Box 1. The Coalescent

The original definition of coalescence is of a time-dependent process that looks backward in time to reconstruct the genealogy from currently observed lineages to ancestral ones, ending up with a most recent common ancestor of the whole sample. The unit of time is usually coded in generations and some important properties of the genealogical tree, such as the total tree length ($L$) and its depth ($D$) (that is the distance of the leaves from the original ancestor of the whole genealogy), depend on the temporal distribution of the coalescent events.

In its simplest formulation, the coalescent is built up independently from the mutational process, since by definition neutral mutations do not influence the number of offspring of each individual and therefore do not shape the genealogical tree (Kingman 1982a, b; Hudson 1990). Given this assumption, neutral mutations can be superimposed onto the genealogy afterwards (Hudson 1990; Nordborg 2000). Population genetic processes, such as population subdivision or migration, can be estimated assuming a neutral model of mutation.

A coalescent event happens when two lineages (represented by a single locus) merge into a unique parental lineage across generational time. Proceeding backwards in time, a finite number of observed offspring led to a single common parent. Different genetic unlinked loci are considered to have independent genealogies. However, neutral loci with equal effective population sizes are expected to fit closely related genealogies mirroring the history of a population.

Coalescence is a continuous time Markov process with a diffusion approximation (i.e. terms higher than the first order are neglected), where the probabilities of all the individuals belonging to a certain parent $N$ from the

previous generation are equal and independent (Wright–Fisher model of random mating). Considering a haploid population with size $N$ as large and constant, the probability of two individuals having the same ancestor in the previous generation is $1/N$ and the probability Pr(2) of having different ancestors is $1 - 1/N$. Adding a third individual, the number of possible different parents is $N-2$ and so the probability Pr(3) of having an ancestor different from the previous two is $(N-2)/N = 1-2/N$. Extending the case to $n$ multiple individuals, the probability Pr($n$) for of them all having different ancestors (i.e. no coalescent event) is the product of all the $i$ in the sampled $n - 1$ lineages of the quantity $(1 - i/N)$:

$$\Pr(n) = \prod_{i=1}^{n} 1 - \frac{1}{N} \tag{14.1}$$

This probability can be described using a binomial coefficient with parameters $n$ and 2 that expresses all the possible different combinations of two individuals within the total sample $n$. Ignoring terms of order $1/N^2$ or higher, one obtains

$$\Pr(n) = 1 - \frac{\binom{n}{2}}{N} = 1 - n\frac{(n-1)}{2N} \tag{14.2}$$

Thus the probability of a coalescent event happening is

$$\Pr(c) = 1 - \Pr(n) = \frac{\binom{n}{2}}{N} \tag{14.3}$$

To obtain a tree of coalescent events, this probability must be estimated across generations. After $t$ generations the probability of a coalescent event becomes $(1 - \binom{n}{2}/N)^t$ and the probability of one single coalescent event occurring after $t + 1$ generations is $\binom{n}{2}/N$ which means that a coalescent event happens in $t + 1$ generations as the product of probability of no coalescent events occurring for $t$ generations and one event occurring after $t + 1$:

$$\Pr(c, t) = \Pr(n)^t(1 - \Pr(n)) = \left(1 - \frac{1}{N}\right)^t n\frac{(n-1)}{2}\frac{1}{N} \tag{14.4}$$

The exponential of the second term (probability of a coalescent event occurred) is equal to one, since it is the probability of a coalescent event occurring in 1 generation, that is the $t + 1$ generation. This probability is very small compared to the number of trials necessary (for this reason there are no multiple coalescent events in one generation), and thus the binomial distribution can be approximated to a Poisson distribution with rate parameter $\binom{n}{2}/N$ and mean $N/\binom{n}{2}$

$$\Pr(c,\ t)=\frac{\binom{n}{2}}{N}e^{-\frac{\binom{n}{2}}{N}t} \tag{14.5}$$

This probability expresses the time until the first coalescent event occurs. From this formulation, it is clear that under a general Wright–Fisher model and a neutral mutation model, the genealogies only depend on the effective population size $N(e)$, that is equal to $N$ in neutral model with a constant population, and its depth is associated to a large variance. Scaling the time in generations helps simplify the model for the times of coalescent events with rate $\lambda$:

$$\lambda=\frac{\binom{n}{2}}{N}\Rightarrow \Pr(c,\ t)=\lambda e^{-\lambda t} \tag{14.6}$$

After each coalescent event, the number of possible ancestor to follow is $n-1$, since the coalescent rate is determined by the number of possible combinations $\binom{n}{2}$, the time to the next coalescent event becomes higher and higher with the decreasing of the number of molecules and if $N \gg n$ its distribution is approximately exponential with mean $N/\binom{n-1}{2}$ which changes after each coalescent event.

The waiting times till the last coalescent event, and hence follows a generalised Erlang distribution (hypo-exponential) with $\lambda_k$ coefficients for $k$ events of coalescence, with maximum $k = n - 1$. Models with deterministic varying effective population size across the generations have been explored by several authors (refs) and are implemented in methods like the generalised skyline plot to estimate the best demographic history (Pybus et al. 2000; Strimmer and Pybus 2001; Drummond et al. 2005).

$$\Pr(t_{n,k}) = S\, c_{i,k}\lambda_{n-i+1}e^{-\lambda n-i+1\ \mathrm{tn},k} \tag{14.7}$$

The mathematical formulation of the coalescent model presented so far can be applied to a diploid system substituting the term $N$ with 2 $N$.

The large variance associated to the statistical distribution of the waiting times for a genealogy of $n$ samples does not allow inferences without taking into account the information contained in the current genetic variability. The current distribution of the genetic variation within a certain sample or population is indeed the direct result of both the molecular evolution and the demographic history, which are usually not reducible to a model with time stable population size, and these two processes cannot be analysed independently when looking backwards in time. All the methods implemented for the estimation of population demographic history and its parameters are based on the observed genetic variation within the sample of interest and a low number of genetic polymorphisms usually prevent a good posterior estimate.

The previous theoretical work on inferences based on reconstructed genealogies (Fu 1994; Nee et al. 1995; Pybus et al. 1999, 2000; Strimmer and Pybus 2001; and Drummond et al. 2005) is probably one of the most explicative examples on the statistical and inferential connection of the mutational process and the effective population size in coalescence. These models are usually based on the assumption of a molecular clock of evolution and consider models with varying population sizes through the generations, without previous assumptions on the demography of the population.

## References

Aeschbacher S, Beaumont MA, Futschik A (2012) A novel approach for choosing summary statistics in approximate Bayesian computation. Genetics 192:1027–1047

Amos W, Manica A (2006) Global genetic positioning: evidence for early human population centers in coastal habitats. Proc Natl Acad Sci USA 103:820–824

Balloux F, (2001) EASYPOP (Version 1.7): A Computer Program for Population Genetics Simulations. J Hered 92:301–302

Barbujani G, (1987) Autocorrelation of gene frequencies under isolation by distance. Genetics 117:777–782.

Bamshad M, Kivisild T, Watkins WS, Dixon ME, Ricker CE, Rao BB, Naidu JM, Prasad BVR, Reddy PG, Rasanayagam A, Papiha SS, Villems R, Redd AJ, Hammer MF, Nguyen SV, Carroll ML, Batzer MA, Jorde LB (2001) Genetic evidence on the origins of Indian caste populations. Genome Res 11:994–1004

Barton NH, Etheridge AM, Véber A (2013) Modelling evolution in a spatial continuum. J Stat Mech 2013:P01002

Beaumont MA, Zhang W, Balding DJ (2002) Approximate bayesian computation in population genetics. Genetics 162:2025–2035

Beerli P, Palczewski M (2010) Unified framework to evaluate panmixia and migration direction among multiple sampling locations. Genetics 185:313–326

Chen C, Durand E, Forbes F, François O (2007) Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. Molecular Ecology Notes 7:747–756

Cheng L, Connor TR, Sirén J, Aanensen DM, Corander J (2013) Hierarchical and spatially explicit clustering of DNA sequences with BAPS software. Mol Biol Evol. doi:10.1093/molbev/mst028

Corander J, Marttinen P (2006) Bayesian identification of admixture events using multilocus molecular markers. Mol Ecol 15:2833–2843

Crisci JL, Poh Y-P, Bean A, Simkin A, Jensen JD (2012) Recent progress in polymorphism-based population genetic inference. J Hered 103:287–296

Dawson KJ, Belkhir K, (2001) A Bayesian approach to the identification of panmictic populations and the assignment of individuals. Genetics Research 78:59–77

Drummond AJ, Rambaut A, Shapiro B, Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. Mol Biol Evol 22:1185–1192

Durand E, Jay F, Gaggiotti OE, François O (2009) Spatial inference of admixture proportions and secondary contact zones. Mol Biol Evol 26:1963–1973

Excoffier L, Ray N (2008) Surfing during population expansions promotes genetic revolutions and structuration. Trends Ecol Evol (Amst) 23:347–351

Ewing G, Hermisson J (2010) MSMS: a coalescent simulation program including recombination, demographic locus. Structure and selection at a single. Bioinformatics 26:2064–2065
Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–1587
Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol. Ecol. Notes 7:574–578
François O, Ancelet S, Guillot G (2006) Bayesian clustering using hidden markov random fields in spatial population genetics. Genetics 174:805–816
François O, Durand E (2010) Spatially explicit Bayesian clustering models in population genetics. Mol Ecol Resour 10:773–784
Frantz AC, Cellina S, Krier A, Schley L, Burke T (2009) Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: clusters or isolation by distance? J Appl Ecol 46:493–505
Fu YX (1994) A phylogenetic estimator of effective population size or mutation rate. Genetics 136:685–692
Guillot G, Leblois R, Coulon A, Frantz AC (2009) Statistical methods in spatial genetics. Mol Ecol 18:4734–4756
Guillot G, Mortier F, Estoup A (2005) Geneland: a computer package for landscape genetics. Mol Ecol Notes 5:712–715
Hedrick PW (2005) A Standardized Genetic Differentiation Measure. Evolution 59:1633–1638
Heller R, Siegismund HR (2009) Relationship between three measures of genetic differentiation GST, DEST and G'ST: how wrong have we been? Mol Ecol 18:2080–2083
Hudson, R.R., 1990. Gene genealogies and the coalescent process. Oxford Surveys in Evolutionary Biology, Vol. 7. D. Futuyma and J. Antonovics (Eds)
Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet 11:94
Jombart T, Devillard S, Dufour A-B, Pontier D (2008) Revealing cryptic spatial patterns in genetic variability by a new multivariate method. Heredity 101:92–103
Jost L (2007) Partitioning diversity into independent alpha and beta components. Ecology 88:2427–2439
Jost L (2008) GST and its relatives do not measure differentiation. Molecular Ecology 17:4015–4026
Kimura M, Weiss GH (1964a) The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics 49:561–576
Kimura M, Weiss GH (1964b) The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics 49:561–576
Kingman JFC (1982a) On the genealogy of large populations. J Appl Probab 19A:27
Kingman JFC (1982b) The coalescent. Stochast Process Appl 13:235–248
Kuhner MK, Smith LP (2007) Comparing likelihood and Bayesian coalescent estimation of population parameters. Genetics 175:155–165
Laval G, Excoffier L (2004) SIMCOAL 2.0: a program to simulate genomic diversity over large recombining regions in a subdivided population with a complex history. Bioinformatics 20:2485–2487
Levins R (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. Bull ESA 15:237–240
Malecot G (1973) Isolation by distance. In: Morton NE (ed) Genetic structure of populations. University of Hawaii Press, Honolulu, pp 72–75
Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. Trends Ecol Evol 18:189–197
Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209–220
Montano V, Marcari V, Pavanello M, Anyaele O, Comas D, Destro-Bisol G, Batini C (2013) The influence of habitats on female mobility in Central and Western Africa inferred from human mitochondrial variation. BMC Evol Biol 13:24

Nee S, Holmes EC, Rambaut A, Harvey PH (1995) Inferring population history from molecular phylogenies. Phil Trans R Soc Lond B 349:25–31
Nordborg M (2000) Coalescent theory. In: Balding DJ, Bishop M, Cannings C (eds) Handbook in statistical genetics. Wiley, Chichester, pp 179–212
Pybus OG, Holmes EC, Harvey PH (1999) The mid-depth method and HIV-1: a practical approach for testing hypotheses of viral epidemic history. Mol Biol Evol 16:953–959
Pybus OG, Rambaut A, Harvey PH (2000) An integrated framework for the inference of viral population history from reconstructed genealogies. Genetics 155:1429–1437
Ray N, Currat M, Foll M, Excoffier L (2010) SPLATCHE2: a spatially explicit simulation framework for complex demography, genetic admixture and recombination. Bioinformatics 26:2993–2994
Ray N, Excoffier L (2009) Inferring past demography using spatially explicit population genetic models. Hum Biol 81:141–157
Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, Zhivotovsky LA, Feldman MW (2002) Genetic structure of human populations. Science 298:2381–2385
Schwartz MK, McKelvey KS (2009) Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. Conserv Genet 10:441–452
Slatkin M (1993) Isolation by Distance in Equilibrium and Non-Equilibrium Populations. Evolution 47:264
Sokal RR, Oden NL (1978) Spatial autocorrelation in biology: 1. Methodology. Biol J Linn Soc 10:199–228
Strimmer K, Pybus OG (2001) Exploring the demographic history of DNA sequences using the generalized skyline plot. Mol Biol Evol 18:2298–2305
Templeton AR (2006) Population genetics and microevolutionary theory, 1st edn. Wiley-Blackwell, New Jersey
Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Mol Ecol 15:1419–1439
Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38:1358
Wright S (1943) Isolation by distance. Genetics 28:114–138
Wright S (1949) The genetical structure of populations. Ann Eugenics 15:323–354
Yang W-Y, Novembre J, Eskin E, Halperin E (2012) A model-based approach for analysis of spatial structure in genetic data. Nat Genet 44:725–731

# Chapter 15
# Consequences of Segregation and Genetic Exchange on Adaptability in Arbuscular Mycorrhizal Fungi (AMF)

**Caroline Angelard and Ian R. Sanders**

**Abstract** Arbuscular mycorrhizal fungi (AMF) form symbioses with the majority of land plants, improving the nutrition and productivity of plants. The fungi are coenocytic, grow clonally, and no sexual stage in their life cycle is known. Recent evidence suggests that AMF are heterokaryotes, i.e., that genetically different nuclei coexist within a single cytoplasm. In this chapter, we present the last studies that have investigated the ecological and evolutionary consequences of an AMF heterokaryotic system on both sides of the symbiosis: Fungi and plants. We first present two important mechanisms, namely segregation and genetic exchange, two processes directly related to the heterokaryotic state of AMF. We then present the consequences of segregation and genetic exchange on AMF/plant interactions. Finally, we discuss the role of heterokaryosis in the enhancement of AMF adaptability and how AMF have maintained genetically different nuclei.

## 15.1 Introduction

### *15.1.1 AMF and Their Importance in Ecosystems*

The majority of terrestrial plants form symbioses with arbuscular mycorrhizal fungi (AMF). The fungi provide minerals to the plants, mostly phosphate, in exchange for carbohydrates. AMF improve plant growth and protect them against pathogens and herbivores (Bennett et al. 2006, Bennett and Bever 2007, Newsham

C. Angelard (✉) · I. R. Sanders
University of Lausanne, Department of Ecology and Evolution, Biophore Building, 1015 Lausanne, Switzerland
e-mail: Caroline.Angelard@unil.ch

I. R. Sanders
e-mail: Ian.Sanders@unil.ch

P. Pontarotti (ed.), *Evolutionary Biology: Exobiology and Evolutionary Mechanisms*,
DOI: 10.1007/978-3-642-38212-3_15, 

et al. 1995). The contribution of the symbiosis to plant growth is of enormous importance for the productivity of the major crops that humans rely on for nutrition. Additionally, the symbiosis plays a key role in nutrient cycling in natural ecosystems, in ecosystem productivity, and plant diversity (van der Heijden et al. 1998b). It has recently been proposed that the symbiosis between plants and fungi is very old, in fact as old as the first land plants. Indeed, plants colonized land 400–500 millions years ago and AMF-like structures were present inside the roots of those first land plants (Remy et al. 1994). It has therefore been proposed that AMF played a key role in colonization of land by plants (Redecker et al. 2000).

### *15.1.2 Life Cycle of AMF*

AMF (phylum Glomeromycota) are believed to be monophyletic and diverged from other fungal lineages several hundred millions years ago (James et al. 2006). The fungi are believed to grow clonally. Indeed, no sexual stage in their life cycle is known. The ancient divergence of AMF and the fact that they grow clonally has lead to the hypothesis that AMF are ancient asexuals (Judson and Normark 1996).

The life cycle of an AMF occurs belowground and starts with the germination of a spore (Fig. 15.1). The hypha produced by the germinated spore colonizes a host root where it forms arbuscules in cells of the root cortex, as well as vesicles, and hyphae. Plants and fungi exchange nutrients via the arbuscules. The fungus starts to produce new hyphae once the symbiosis has been established. Those new hyphae grow out of the root, where they can act as an extension of the root system by taking up nutrients and transporting them back to the plants. The surrounding soil and other potential plant roots can then become colonized, forming a large mycelial network. The hyphal network produced by a single AMF can exhibit a diameter of at least 10 m (Rosendahl and Stukenbrock 2004), resulting in a single fungus simultaneously living in quite different biotic and abiotic conditions. At the

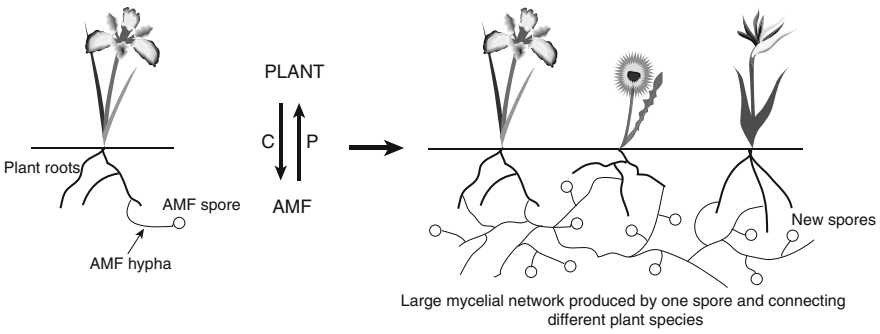



**Fig. 15.1** Life cycle of AMF

end of hyphal tips new spores are formed. These new spores can serve as resting stages and as dispersal units (Mangan and Adler 2002).

Two morphological features of AMF distinguish them from Basidio- and Ascomycetes: (1) Spores of AMF are multinucleate and generally contain from a few hundred to several thousand of nuclei (Bianciotto and Bonfante 1992); (2) The hyphae of AMF are not septate (not divided into cells) and are therefore coenocytic. It has been recently shown that the nuclei composition of newly formed spores has two origins: massive amounts of nuclear that migrate into the new spores and the nuclei arising from mitosis inside the new spores (Marleau et al. 2011). As a consequence, AMF lack the strict genetic bottleneck of a single-nucleus stage present in all other known eukaryotic organism.

In this work, an AMF individual or an AMF line corresponds to all the clonal material (new hyphae and new spores) that has been initiated from one spore.

### *15.1.3 Genetics of AMF and Heterokaryosis*

Given the importance of AMF, it is surprising that remarkably little is known about the genetics and genomics of the fungi. One limitation to the study of AMF genetics is that the fungi are obligate biotrophs, which means that they have to grown with plant roots to develop. AMF have not been successfully cultured without host roots. It is, therefore, difficult to obtain a large amount of clean material for molecular studies. A clean culture system can be obtained with in vitro cultures containing an artificial media and roots transformed with *Agrobacterium tumefaciens* (Becard and Piche 1992, StArnaud et al. 1996). However, few species have been successfully established in this axenic system. The AMF species *Rhizophagus irregularis* grows well in this culture system, and most of our knowledge is, therefore, limited to this species. *R. irregularis* was previously ascribed to the species *Glomus intraradices* and subsequently to *G. irregulare* (Stockinger et al. 2009). The fungus has, however, recently been renamed as *R. irregularis* (Kruger et al., 2012). Another limitation in the study of AMF genetics is the long time of AMF cultivation compared to other microorganisms. The generation time, that corresponds to the time from inoculation of roots with spores until the production of new viable spores, is a minimum of about 3 months.

No genome sequence data are currently available, although a project to sequence a complete genome of an AMF species is ongoing and a recent large survey of the transcriptome has been published (Martin et al. 2008; Tisserant et al. 2012). Levels of ploidy are not known for most species. *R. irregularis* was shown to be haploid (Hijri and Sanders 2004) and evidence also supports haploidy in *Scutellospora castanea*, an AMF species with a very high nuclear DNA content (Hijri and Sanders 2005). Cytological descriptions of nuclear fusion, mitosis, or meiosis are not available to date.

Some of the recent findings concerning AMF genetics and genome organization are so particular to this organism that conventional models of evolution and of

mendelian genetics are difficult to apply. Most organisms have one nuclei per cell and all the nuclei are genetically identical among cells. As previously mentioned in Sect. 15.1.2 of this chapter, AMF are coenocytic (they are not compartmentalized in cells) and hundreds of nuclei coexist. Although controversial (Pawlowska and Taylor 2004), direct and indirect evidences suggest that AMF contain population of genetically different nuclei (nucleotypes) coexisting in a single AMF (Angelard et al. 2010; Hijri and Sanders 2005; Kuhn et al. 2001). AMF are totally unique organism for their genetic organization and reproductive system. In addition to being ancient asexuals, AMF contain genetically different nuclei coexisting in a single organism and this genetic diversity is transmitted to the next generation, without a reduction of one nuclei per cell during reproduction (Marleau et al. 2011). This heterokaryotic system has been thought not to occur because of the potential problems related to this genetic organization. Indeed, how can an organism function while receiving information from different nuclei? What about genomic/genetic conflict and selfish gene/nuclei? How can the heterokaryotic system maintained for several million years, and why?

## 15.2 Genetic Exchange and Segregation in AMF

### *15.2.1 Genetic Exchange*

It is well-known that different hyphae of a same AMF individual can anastomose (fuse) and create new cytoplasmic connections (Giovannetti et al. 1999, 2001). A large and interconnected mycelium can be built in that way from a single AMF individual, and even connect roots of different plant species (Giovannetti et al. 2004). However, fusion of hyphae between genetically different individuals of the same AMF species was thought not to occur due to potential genetic incompatibilities. This kind of fusion between two individuals having different nucleotypes can have important consequences for the dynamism and maintenance of AMF genetic diversity and for the functioning of the symbiosis itself.

A recent study presents direct evidence for anastomoses between hyphae of genetically different individuals of the AMF species *R. irregularis* (Croll and Sanders 2009). The nuclei of the two AMF individuals can have been mixed in the hyphal network produced after the fusion. The authors analyzed the progeny produced after the hyphal fusion and showed that the offspring exhibited biparental inheritance (Croll and Sanders 2009). Moreover, some offspring differed in fitness-related traits (spore and hyphal densities) compared to their parents and compared to other progeny (Croll and Sanders 2009). Exchange of nuclei among AMF lines can lead, therefore, to the emergence of unique and novel AMF phenotypes. This process of exchange of nuclei between two AMF individuals and resulting in AMF offspring having a mix of parental nuclei has been called "genetic exchange" (Fig. 15.2a). Genetic exchange is a direct consequence of

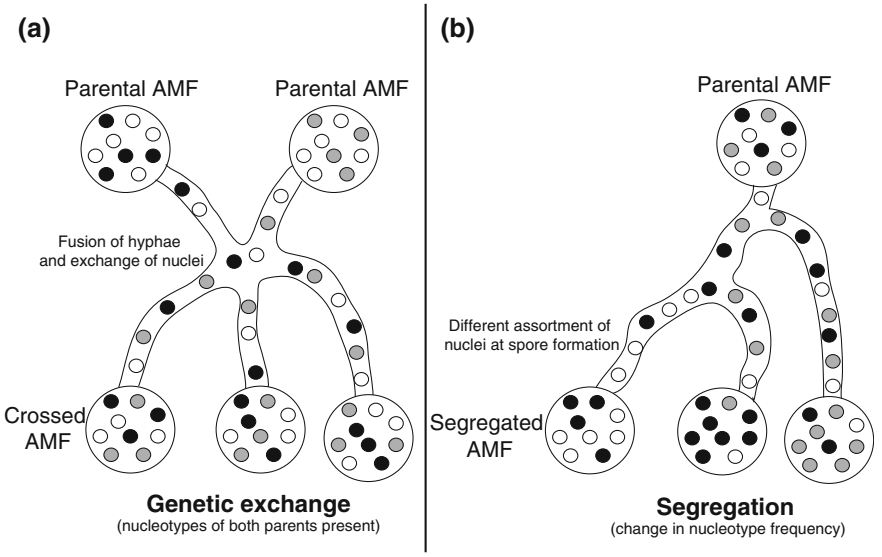



**Fig. 15.2** Two potential consequences of heterokaryosis, potentially resulting in a change in the relative nucleotype frequencies in the offspring. **a** *Genetic exchange* hyphae of two genetically different parental AMF can fuse, allowing the exchange of nuclei; offspring (here crossed AMF) can contain an assortment of genetically different nucleotypes. **b** *Segregation* genetically different nuclei of one parental AMF can be differently segregated in offspring (here segregated AMF)

AMF heterokaryosis, as it allows the mixing of nucleotypes within single spores. AMF individuals that fuse their hyphae and exchange nuclei are called here "parental AMF" or "parental lines".

A potential consequence of genetic exchange is that genetically different individuals of a population can connect and form a large, genetically diverse hyphal network. Moreover, exchange of genetically different nuclei between AMF individuals can have important impacts on the maintenance of heterokaryosis in AMF (Bever and Wang 2005; Hijri and Sanders 2005). Another important consequence of genetic exchange is that the mixing of nuclei through genetic exchange would increase genetic diversity of nucleotypes in the offspring (called here "crossed lines").

### *15.2.2 Segregation*

Another consequence of heterokaryosis in AMF is that one mother spore can produce new spores with different nucleotype contents due to segregation of nucleotypes at spore formation (Fig. 15.2b). The different assortment of nuclei at spore formation is called "segregation". Progeny produced through the process of segregation are called "segregated lines". An important potential consequence of

segregation in AMF is that one spore could produce new spores with a lower genetic diversity due to the loss of rare nucleotypes.

Segregation was thought not to occur until very recently. Previous study investigated segregation hypothesis by looking at inheritance of 13 variants of a gene (PLS) that were known to co-exist in the AMF *G. etunicatum* (Pawlowska and Taylor 2004). The authors fund the 13 variants of the PLS gene in all offspring after one generation and conclude that all the 13 variants must be present in each nucleus and that *G. etunicatum*, a presumed haploid species, could be highly polyploid. However, homokaryosis in this AMF species was then challenged with experimental and theoretical approaches. It has also been shown, with simulations, that hyphal fusions among germinated spores of a same individual could explain the absence of segregation of the 13 variants of the PLS gene (Bever and Wang 2005).

The authors studied the segregation of variants at only one locus and concluded that segregation did not occur in AMF. A multi-locus study would have been more appropriate to detect segregation in only one generation. Moreover, the authors only looked at "total segregation", which refers to the total disappearance of some variants in offspring. Segregation could have occurred, but in a quantitative way, with spores receiving different relative proportions of the variants. This quantitative segregation has been called "partial segregation" and corresponds to a change in frequency of different nucleotypes in newly formed spore, but without a complete loss of nucleotypes (Sanders and Croll 2010).

The occurrence of segregation has only been recently shown with the AMF species *R. irregularis* (Angelard et al. 2010). The likelihood of segregation occurring is dependent on the genetic diversity of nucleotypes present in the mother spore. As said in the previous (Sect. 15.2.1), genetic exchange should lead to new AMF individuals (crossed lines) with higher genetic diversity compared to their parents because of mixing the nuclei. Therefore, the investigation into the process of segregation has been performed by isolating and cultivating separately single spores coming from a same crossed line. The culture of single spores gave rise to new AMF lines called segregated lines. The authors then compared the genetic content between segregated and crossed lines, and among segregated lines (see Fig. 15.3 for a general understanding of how segregation can be detected). They used genetic markers to detect total segregation and other genetic markers to detect partial segregation. They showed that markers disappeared and/or fluctuated in frequency among segregated lines and between segregated lines and crossed lines. The disappearance or fluctuation of nucleotypes means that the new spores formed from the same AMF individual, can contain a different assortment or a different proportion of nucleotypes compared to parental or other sibling spores. Therefore, segregation, partial or total, occurs in the AMF species *R. irregularis* (Angelard et al. 2010). Moreover, segregated lines can differ in fitness-related traits (spore and hyphal densities) compared to their parents and compared to other segregated lines (unpublished data). Segregation is, therefore, another mechanism, in addition to genetic exchange, that can lead to the emergence of unique and novel AMF phenotypes and genotypes.

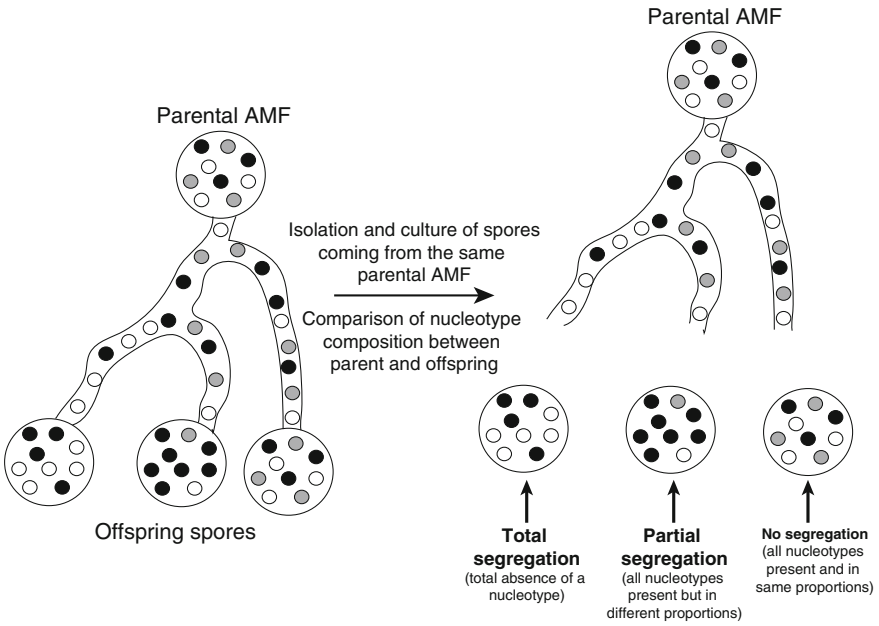



**Fig. 15.3** Detection of segregation. Comparison of nucleotype frequencies between siblings and the parent

## 15.3 Consequences of Segregation and Genetic Exchange on AMF–Plant Interactions

### *15.3.1 Relevance of Genetic Exchange and Segregation*

The impact of AMF diversity on plant diversity and plant productivity has been well studied at the AMF species level (Hart and Reader 2002; Stampe and Daehler 2003; van der Heijden et al. 1998a, b, 2003), as well as the effect of different plant species on AMF species spore production and on the coexistence of diverse AMF communities (Burrows and Pfleger 2002). However, the consequences of the genetic diversity caused by to heterokaryosis in AMF and its effect on the interaction of the fungus with plants have been given much less attention.

In the study revealing genetic exchange, the fusion of hyphae between genetically different AMF individuals was a rare event compared to the fusion among hyphae of a same AMF individual (Croll and Sanders 2009). It is important to note that this study has been performed in in-vitro conditions. The mycelial networks existing in natural environments can be very large and can connect different plant species at the same time. Contacts between hyphae of different individuals are, therefore, highly likely to happen often. Consequently, the fusion between hyphae of different individuals could have impact on the mixing of nuclei in the hyphal

network. Similarly, nuclear segregation at spore formation is likely to be a very common event, because generally spores are produced frequently and in high quantity. Consequently, alterations in relative nucleotype frequencies in progeny obtained through genetic exchange or segregation are likely to be common events in mycelial networks.

The impact of both genetic exchange and segregation could be very important for the symbiosis. Indeed, if different nuclei express different genes, then alteration of nuclear frequencies could have important impacts on plant/AMF interactions. We can hypothesize that progeny obtained through genetic exchange and segregation can have different effect on plants, such as on plant growth and plant gene transcription. Moreover, those effects can be plant specific, meaning that the growth response of one plant species will not be the same than the one of another plant species. On the other hand, plants species can also differentially affect the genotypes and phenotypes of AMF lines, and the effect can be different between offspring and their parental AMF. The following sections focus on those hypotheses and present works that have been performed those last years to answer those questions.

### *15.3.2 Effects of Segregation and Genetic Exchange on Plant Growth and Plant Gene Transcription*

Two recent studies investigated the effect of genetic exchange and segregation on plant growth and plant gene transcription (Angelard et al. 2010; Colard et al. 2011). Those studies have been performed with "real" plants in the greenhouse (as opposed to the studies performed in in-vitro cultures with transformed roots). The authors inoculated plants with spores of segregated, crossed, and parental AMF lines. After few months of growth in greenhouse, they studied the plant dry mass and the transcription of plant genes. Both studies showed that genetic alterations via segregation and genetic exchange have important effects on both rice growth and rice gene transcription.

Even though rice establishes symbioses with AMF, no published study, to our knowledge, shows a clear positive impact of AMF inoculation on rice growth, despite to other studies showing positive responses to AMF of hundreds of plant species. However, the studies conducted by Angelard et al. (2010), Colard et al. (2011) reveal that rice growth responses to AMF are not fixed and can be altered by simple manipulations of the genetics of the fungi by genetic exchange and segregation. One of these studies even revealed a growth increase of up to five times when rice was inoculated with some segregated lines, compared with rice inoculated with parental line or other segregated lines (Angelard et al. 2010). Results of those studies also show that segregation and genetic exchange can have large effects on rice gene expression. Genes involved in plant growth were affected, as well as early and late transcribed genes that are specific to the

establishment of the symbiosis, and fundamental mycorrhizal specific genes implicated in nutrient acquisition in plants.

Additionally, the effects found on plant growth in the study of Angelard et al. (2010) were plant species specific. Indeed, the authors used two different plant species, rice and plantain, and showed that differences in plant growth due to inoculation with different AMF lines were not the same depending on the plant species. For example, some segregated AMF lines increased the growth of rice compared to other AMF lines, but no positive growth effect was observed for plantain growing with the same AMF lines. Other studies found that plants did not respond in the same way when inoculated with different AMF (Johnson et al. 1997; Koch et al. 2006; StreitwolfEngel et al. 1997, van der Heijden et al. 1998a), and that AMF are not always beneficial for all plant species (Johnson et al. 1997, Klironomos 2003, Koch et al. 2006). However, those studies reveal the importance of AMF diversity at the community or population scale. The results of Angelard et al. (2010), Colard et al. (2011) reveal that manipulating the genetics of AMF at the individual level has important consequences for plants.

### *15.3.3 Effects of a Change of Environment on AMF Phenotype and Genotype*

In the previous Sect. 15.3.2, we have seen that segregation and genetic exchange can lead to new individuals that differentially alter the growth of different plant species. Here, we will discuss how different plant species can differentially alter the phenotypes and genotypes of AMF offspring obtained by segregation and genetic exchange.

In the study of Angelard et al. (2010), rice and plantain were inoculated with segregated and crossed lines. The authors found important effects of segregation and genetic exchange on plant growth and plant gene transcription. Angelard and Sanders (Angelard et al. 2010) revealed that the phenotypes (percentage of arbuscules, vesicles, and hyphae) of the fungi within the roots of the plants used in the study of Angelard et al. (2010) were also affected by (1) segregation, (2) genetic exchange, and by (3) the different plant species. Indeed, segregated (and crossed) AMF lines had different phenotypes compared to their parents and compared to each other. Moreover, those differences among AMF lines were not the same depending on the plant species. Combined with the study of Angelard et al. (2010), the results suggest that specific interactions could occur between different plant species and different AMF genotypes.

Another study conducted by Angelard et al. (unpublished data) investigated the effect of different plant species on the genotypes and phenotypes of segregated and crossed AMF lines. The study was performed in in-vitro cultures to be able to conduct clean molecular analyses of AMF genotypes. The design of the experiment allowed to test the effect of a change of environment, here a change of host

species, on the genotype and phenotype of AMF. Indeed, AMF lines (segregated, crossed and parental AMF) were maintained for several generations with transformed roots of one plant species, and then half of the replicates were transferred with roots of another plant species, while the remaining half were maintained with roots of the first plant species. The authors tested two hypotheses: (1) the genetic composition of AMF can rapidly change following a change of host due to alteration of nucleotype frequencies, (2) the genetic changes will be different depending on the AMF lines. Results revealed genetic changes in all the AMF lines following a change of host. The genetic changes were different between offspring (segregated or crossed) and their parents and among offspring. Moreover, the phenotypes of the AMF lines were also affected by the change of host. The changes were different between siblings and their parents and among siblings.

Taken together, the results of those studies suggest that AMF can create new individuals with different genetic content via segregation and genetic exchange, in a very short time, and those new offspring can respond genetically and phenotypically differently to a change in the environment compared to their parents. Moreover, the results also suggest that AMF can produce a large variety of offspring that are potentially able to respond differently to a change in the environment. As a consequence, some offspring could perform better in the new environment and can be differentially selected. Therefore, heterokaryosis and changes of genetic content via natural processes (genetic exchange and segregation) could be mechanisms by which AMF increase their adaptability to changing environment. This is particularly relevant because AMF grow in almost all natural ecosystems and are able to form large mycelial networks connecting plant of different species, even at the same time. AMF, therefore, encounter a large variety of different abiotic (pH, nitrogen, phosphate etc.), and biotic (such as different plant species) factors and need to respond and adapt rapidly to those different conditions.

## 15.4 How and Why AMF Maintain Different Nucleotypes?

Complex among species interactions and fitness feedbacks have been proposed as mechanisms allowing the coexistence of plant and AMF species diversity in natural ecosystems (Bever 2002; Bever et al. 2009; Hart et al. 2003). Studies presented in this chapter suggest that such mechanisms could also act at a lower level than AMF species, and can potentially explain the genetic and phenotypic diversity found in an AMF population. Moreover, such mechanisms (i.e., specificity of associations between host species and AMF genotypes) could also act at the AMF individual scale and participate in the maintenance of the intra-individual genetic diversity (heterokaryosis) in AMF.

Genetic exchange and segregation can have important effects on the dynamics of the nucleotype contents of spores. Indeed, those mechanisms can lead, in a very short time, to progeny with novel symbiotic effects. Some of those new AMF

offspring, due to their particular and novel nucleotype composition, could be better suited in case of a change in the environment than other AMF, and could therefore be selected.

To conclude, environmental heterogeneity, such as different host species, could affect the fate of different nucleotypes following genetic exchange and segregation in AMF, and could potentially contribute to the maintenance of heterokaryosis in AMF. On the other hand, heterokaryosis could also be a way to enhance the adaptability of AMF by creating, in a very short time, a large variety of new AMF progeny on which selection can act.

**Acknowledgments** C. Angelard is funded by a Marie Curie Outgoing International Fellowship within the 7th European Community Framework Program (PIOF_GA_2009-251712).

## References

Angelard C, Colard A, Niculita-Hirzel H, Croll D, Sanders IR (2010) Segregation in a mycorrhizal fungus alters rice growth and symbiosis-specific gene transcription. Curr Biol 20:1216–1221

Becard G, Piche Y (1992) Establishment of vesicular-arbuscular mycorrhiza in root organ-culture—review and proposed methodology. Method microbiol 24:89–108

Bennett AE, Alers-Garcia J, Bever JD (2006) Three-way interactions among mutualistic mycorrhizal fungi, plants, and plant enemies: hypotheses and synthesis. Am Nat 167:141–152

Bennett AE, Bever JD (2007) Mycorrhizal species differentially alter plant growth and response to herbivory. Ecology 88:210–218

Bever JD (2002) Negative feedback within a mutualism: host-specific growth of mycorrhizal fungi reduces plant benefit. Proc R Soc Lond, Ser B: Biol Sci 269:2595–2601

Bever JD, Wang M (2005) Arbuscular mycorrhizal fungi—hyphal fusion and multigenomic structure. Nature 433

Bever JD, Richardson SC, Lawrence BM, Holmes J, Watson M (2009) Preferential allocation to beneficial symbiont with spatial structure maintains mycorrhizal mutualism. Ecol Lett 12:13–21

Bianciotto V, Bonfante P (1992) Quantification of the nuclear-dna content of 2 arbuscular mycorrhizal fungi. Mycol Res 96:1071–1076

Burrows RL, Pfleger FL (2002) Host responses to AMF from plots differing in plant diversity. Plant Soil 240:169–179

Colard A, Angelard C, Sanders IR (2011) Genetic exchange in an arbuscular mycorrhizal fungus results in increased rice growth and altered mycorrhiza-specific gene transcription. Appl Environ Microbiol 77

Croll D, Sanders IR (2009) Recombination in glomus intraradices, a supposed ancient asexual arbuscular mycorrhizal fungus. BMC Evol Biol 9

Giovannetti M, Azzolini D, Citernesi AS (1999) Anastomosis formation and nuclear and protoplasmic exchange in arbuscular mycorrhizal fungi. Appl Environ Microbiol 65:5571–5575

Giovannetti M, Fortuna P, Citernesi AS, Morini S, Nuti MP (2001) The occurrence of anastomosis formation and nuclear exchange in intact arbuscular mycorrhizal networks. New Phytol 151:717–724

Giovannetti M, Sbrana C, Avio L, Strani P (2004) Patterns of below-ground plant interconnections established by means of arbuscular mycorrhizal networks. New Phytol 164

Hart MM, Reader RJ (2002) Taxonomic basis for variation in the colonization strategy of arbuscular mycorrhizal fungi. New Phytol 153:335–344
Hart MM, Reader RJ, Klironomos JN (2003) Plant coexistence mediated by arbuscular mycorrhizal fungi. Trends Ecol Evol 18:418–423
Hijri M, Sanders IR (2004) The arbuscular mycorrhizal fungus Glomus intraradices is haploid and has a small genome size in the lower limit of eukaryotes. Fungal Genet Biol 41:253–261
Hijri M, Sanders IR (2005) Low gene copy number shows that arbuscular mycorrhizal fungi inherit genetically different nuclei. Nature 433:160–163
James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, Cox CJ et al (2006) Reconstructing the early evolution of Fungi using a six-gene phylogeny. Nature 443:818–822
Johnson NC, Graham JH, Smith FA (1997) Functioning of mycorrhizal associations along the mutualism-parasitism continuum. New Phytol 135:575–586
Judson OP, Normark BB (1996) Ancient asexual scandals. Trends Ecol Evol 11:A41–A46
Klironomos JN (2003) Variation in plant response to native and exotic arbuscular mycorrhizal fungi. Ecology 84:2292–2301
Koch AM, Croll D, Sanders IR (2006) Genetic variability in a population of arbuscular mycorrhizal fungi causes variation in plant growth. Ecol Lett 9:103–110
Kruger M, Kruger C, Walker C, Stockinger H, Schussler A (2012) Phylogenetic reference data for systematics and phylotaxonomy of arbuscular mycorrhizal fungi from phylum to species level. New Phytol 193:970–984
Kuhn G, Hijri M, Sanders IR (2001) Evidence for the evolution of multiple genomes in arbuscular mycorrhizal fungi. Nature 414:745–748
Mangan SA, Adler GH (2002) Seasonal dispersal of arbuscular mycorrhizal fungi by spiny rats in a neotropical forest. Oecologia 131:587–597
Marleau J, Dalpe Y, St-Arnaud M, Hijri M (2011) Spore development and nuclear inheritance in arbuscular mycorrhizal fungi. BMC Evol Biol 11:51
Martin F, Gianinazzi-Pearson V, Hijri M, Lammers P, Requena N, Sanders IR et al (2008) The long hard road to a completed Glomus intraradices genome. New Phytol 180:747–750
Newsham KK, Fitter AH, Watkinson AR (1995) Arbuscular mycorrhiza protect an annual grass from root pathogenic fungi in the field. J Ecol 83:991–1000
Pawlowska TE, Taylor JW (2004) Organization of genetic variation in individuals of arbuscular mycorrhizal fungi. Nature 427:733–737
Redecker D, Kodner R, Graham LE (2000) Glomalean fungi from the Ordovician. Science 289:1920–1921
Remy W, Taylor TN, Hass H, Kerp H (1994) 4-hundred-million-year-old vesicular-arbuscular mycorrhizae. Proc Natl Acad Sci USA 91:11841–11843
Rosendahl S, Stukenbrock EH (2004) Community structure of arbuscular mycorrhizal fungi in undisturbed vegetation revealed by analyses of LSU rDNA sequences. Mol Ecol 13:3179–3186
Sanders IR, Croll D (2010) Arbuscular mycorrhiza: the challenge to understand the genetics of the fungal partner. Annu Rev Genet 44:271–292
Stampe ED, Daehler CC (2003) Mycorrhizal species identity affects plant community structure and invasion: a microcosm study. Oikos 100:362–372
StArnaud M, Hamel C, Vimard B, Caron M, Fortin JA (1996) Enhanced hyphal growth and spore production of the arbuscular mycorrhizal fungus Glomus intraradices in an in vitro system in the absence of host roots. Mycol Res 100:328–332
Stockinger H, Walker C, Schussler A (2009) 'Glomus intraradices DAOM197198', a model fungus in arbuscular mycorrhiza research, is not Glomus intraradices. New Phytol 183:1176–1187
StreitwolfEngel R, Boller T, Wiemken A, Sanders IR (1997) Clonal growth traits of two Prunella species are determined by co-occurring arbuscular mycorrhizal fungi from a calcareous grassland. J Ecol 85:181–191

Tisserant E, Kohler A, Dozolme-Seddas P, Balestrini R, Benabdellah K, Colard A et al. (2012) The transcriptome of the arbuscular mycorrhizal fungus glomus intraradices (DAOM 197198) reveals functional tradeoffs in an obligate symbiont. New Phytol 193

van der Heijden MGA, Boller T, Wiemken A, Sanders IR (1998a) Different arbuscular mycorrhizal fungal species are potential determinants of plant community structure. Ecology 79:2082–2091

van der Heijden MGA, Klironomos JN, Ursic M, Moutoglis P, Streitwolf-Engel R, Boller T et al (1998b) Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. Nature 396:69–72

van der Heijden MGA, Wiemken A, Sanders IR (2003) Different arbuscular mycorrhizal fungi alter coexistence and resource distribution between co-occurring plant. New Phytol 157:569–578

# Chapter 16
# Mitochondrial Genes, Sex Determination and Hermaphroditism in Freshwater Mussels (Bivalvia: Unionoida)

**Donald T. Stewart, Walter R. Hoeh, Gerhard Bauer and Sophie Breton**

**Abstract** Many bivalves, including both marine and freshwater mussels, have an unusual system of "doubly uniparental inheritance" (DUI) of mitochondrial (mt) DNA. In species with DUI, males are heteroplasmic for a male-transmitted (M-type) mitochondrial genome that is dominant in sperm and a female-transmitted (F-type) mitochondrial genome. The F-type genome is found in somatic tissues of both sexes as well as in eggs. In unionoid freshwater mussels, dioecious species have been shown to have both M and F-type genomes. Sequence analysis of mt genomes has shown that there are gender-specific, novel, open reading frames (*orfs*) in the F- and M-type genomes, respectively, in all dioecious freshwater mussel species examined (i.e., an *F-orf* and an *M-orf*, respectively). In contrast, hermaphroditic freshwater mussels lose the M-type genome completely (and, concomitantly, the *M-orf* is lost as well) and the remaining F-type undergoes macromutations in the *F-orf*. We refer to this highly modified *orf* in hermaphrodites as the "*H-orf*". The functions of the *F-orf*, *M-orf*, and *H-orf* genes and their protein products have not yet been determined. However, immunoelectron microscopy has demonstrated the *F-ORF* protein in mitochondria (which is consistent with a mitochondrial encoded protein) but also on the nuclear membrane and in the nucleoplasm (which is not thought normal for a mitochondrial encoded

D. T. Stewart (✉)
Department of Biology, Acadia University, Wolfville, NS B4P 2R6, Canada
e-mail: don.stewart@acadiau.ca

W. R. Hoeh
Department of Biological Sciences, Kent State University, Kent, OH 44242, USA
e-mail: randy.hoeh@gmail.com

G. Bauer
Biology I, University Freiburg, Hauptstrasse 1, D-79104 Freiburg, Germany
e-mail: Gerhard.Bauer@biologie.uni-freiburg.de

S. Breton
Département de sciences biologiques, Université de Montréal, Pavillon Marie-Victorin 90, avenue Vincent-d'Indy, Montréal, QC H2V 2S9, Canada
e-mail: s.breton@umontreal.ca

P. Pontarotti (ed.), *Evolutionary Biology: Exobiology and Evolutionary Mechanisms*,
DOI: 10.1007/978-3-642-38212-3_16, 

protein). The presence of this protein in extra-mitochondrial locations suggests a non-oxidative phosphorylation function for the *F-ORF* protein and has led to the hypothesis that this protein is involved in the genetic regulatory network specifying dioecy (separate males and females) in freshwater mussels. Future research should focus on the period of transitioning from dioecy to hermaphroditism to characterize key mutational changes in the *orf* gene(s) associated with this change in reproductive strategy.

## 16.1 A Review of Doubly Uniparental Inheritance of Mitochondrial DNA in Bivalves

In contrast to the paradigm of strict maternal inheritance (SMI) of mitochondrial DNA, many marine (orders Mytiloida and Veneroida) and freshwater (order Unionoida) bivalves have a fundamentally different mode of mtDNA transmission referred to as "Doubly Uniparental Inheritance" or DUI. In "Power, Sex, Suicide—Mitochondria and the Meaning of Life," Lane (2005) hypothesizes that uniparental inheritance has repeatedly been reproductively advantageous in animals and plants because it allows nuclear genomes to test their compatibility with specific mitochondrial genomes and limits the likelihood of "cytoplasmic warfare" between competing selfish genetic elements. Under the exceptional system of DUI, males receive both their mother and father's mtDNAs (reviewed in Breton et al. 2007) and this system would appear to challenge aspects of Lane's hypothesis. However, as male mussel embryos mature, their father's mtDNA comes to dominate in sperm and their mother's mtDNA comes to dominate in somatic tissues (Garrido-Ramos et al. 1998). In females, the father's mitochondria disappear soon after fertilization (but exceptions have been noted, see Sano et al. 2010 and Zouros 2012 and references therein). Because this system is "doubly uniparental", DUI results in two gender-associated mtDNA lineages; a male-transmitted lineage or M-type mitochondrial genome that is primarily, but not exclusively involved in functioning in sperm (Dalziel and Stewart 2002), and a female-transmitted lineage or F-type mitochondrial genome that functions in all female tissues, and in the somatic tissues of males (Garrido-Ramos et al. 1998). Understanding the evolutionary and ecological forces that have resulted in the origin and maintenance of DUI in bivalves will provide key insights into the fundamental cellular processes associated with standard maternal inheritance as found in all other animals (Stewart et al. 2009; Zouros 2012). One of the leading hypotheses to explain the origin and maintenance of DUI is that a male-transmitted genome might have evolved because it plays a role in sex determination (Breton et al. 2011; Zouros 2012) and this could benefit a "selfish" paternally transmitted cytoplasmic genome (Burt and Trivers 2006).

The standard M and F-type mitochondrial genomes in some freshwater mussels (e.g., families Unionidae and Margaritiferidae) have been on separate evolutionary trajectories for over 200 million years (Hoeh et al. 2002; Doucet-Beaupré et al. 2010). It was recently suggested by Breton et al. (2011) that novel genes found in the M and F mt genomes of freshwater mussels might be involved in sex determination. The existence of several species of freshwater mussels in North America that are obligate simultaneous hermaphrodites allowed Breton et al. (2011) to examine the relationship between mitochondrial genome evolution and a sexual state other than dioecy (i.e., separate sexes). By sequencing complete M and F mt genomes as well as the hermaphrodite's (H) mt genome, present in congeneric hermaphroditic sister taxa of dioecious freshwater mussels (Fig. 16.1), it was shown that: (1) the F mt genome contains a unique gene, referred to as the *F-orf* or F-open reading frame, with no evident homology to any other known proteins; (2) the M mt genome contains a separate *M-orf* that is not obviously homologous to the *F-orf* or to any other known protein; (3) hermaphrodites retain a modified *F-orf* (referred to as the *H-orf*) and they lose the M genome completely; (4) *F-orf* sequence similarity is relatively high among unionoid FWMs but the *H-orfs* exhibit extremely variable nucleotide sequences and deduced amino acid hydrophobicity profiles; and (5) immunoelectron microscopy has shown that the *F-ORF* protein is transported out of the mitochondria (thought unique for mtDNA encoded proteins) and is imported to the cell nucleus. This last observation indicates that the *F-ORF* protein has a non-oxidative phosphorylation function (again, thought unique for mitochondrially encoded proteins) and is consistent with the hypothesis that this gene plays a role in sex determination.

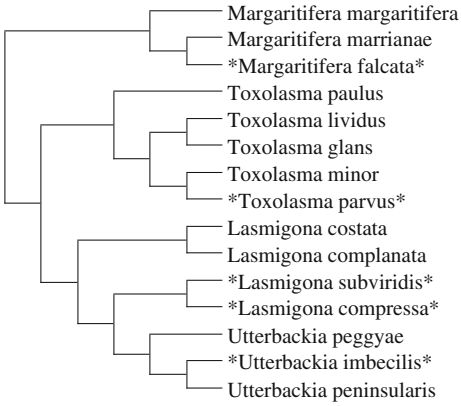



**Fig. 16.1** Simplified phylogeny for a collection of dioecious and hermaphroditic freshwater mussel species (family Unionidae and Margaritiferidae). Species marked by asterisks (*) are hermaphrodites which have (1) lost the M-transmitted genome, (2) reverted to uniparental inheritance, and (3) experienced considerable molecular evolution in the transition from an *F-orf* gene present in a dioecious ancestor to an *H-orf* gene in the extant hermaphrodite (Adapted from Breton et al. 2011)

## 16.2 Hermaphroditism in Freshwater Mussels

In freshwater mussels, as in various invertebrates, dioecious species occasionally evolve into hermaphrodites (also known as ambisexual or monoecious species). It has been hypothesized that selection favors the evolution of hermaphrodites in situations where mates are dispersed and gametes of the opposite sex might be difficult to locate (Ghiselin 1969). This has been proposed as a specific reproductive strategy of freshwater mussels living in low-density populations especially those with low rates of water flow such that gamete distribution is limited (e.g., Bauer 1987).

Of the 300 or so species of freshwater mussels in North America, only seven have been shown to be simultaneous hermaphrodites (Hoeh et al. 1995; Johnston et al. 1998) with sporadic hermaphroditism found in many others (van der Schalie 1970; Heard 1975). All of the obligate simultaneous hermaphrodites identified among the new world freshwater are relatively recently evolved species and character reconstruction indicates that hermaphroditism is a derived character in the family Unionidae (van der Schalie 1970; Hoeh et al. 1995). The inference that hermaphroditism is derived and independently evolved repeatedly within the large cohort of North American freshwater mussel species suggests that it is a rather ephemeral evolutionary state. Just as parthenogenic asexual species are evolutionarily short-lived due to lack of recombination and loss of adaptive potential to respond to novel environmental or biotic challenges (e.g., Bentsson 2009), hermaphroditism, particularly when it involves high-levels of selfing, is predicted to be an evolutionary dead-end (e.g., Takebayahi and Morrell 2001). The available evidence suggests that selfing is common in North American hermaphroditic unionoid species (Hoeh et al. 1998; Johnston et al. 1998).

We are interested in understanding the relationship among the inheritance of gender-associated lineages and sexual determination and gonad development of males, females, and hermaphrodites among freshwater mussels. There is likely a fundamental difference between processes of sexual development and, more specifically, gametogenic apportionment in the ovotestis of obligate simultaneous hermaphrodites and the accidental sporadic hermaphrodites periodically identified. For example, obligate simultaneous syngonic hermaphrodites contain an ovotestis in which sperm and ova are present in the same organ (Davison 2006). The syngonic ovotestis tissue of freshwater mussels identified as obligate hermaphrodites apparently contain discrete acini (or follicles) that produce either spermatogonia or oogonia but not both gametes and these acini are distributed in a consistent pattern within the ovotestis across individuals within a population. In contrast, accidental hermaphroditic individuals are characterized by random appearance of small amounts of gonad tissue or even random cells containing gametes of the opposite sex to the predominant gametes within a gonad (e.g., a very small number of sperm bearing follicles found in a predominantly egg bearing gonad) (van der Schalie 1970). It has been suggested that the accidental hermaphroditic condition is due to a failure of the sex-determining mechanism to function normally (van der Schalie 1970). While the developmental pathway(s) of the obligate versus accidental

hermaphrodites might be fundamentally different in the sense that the obligate hermaphroditic developmental pattern has become reified, the difference between these two categories of hermaphrodites could in fact be rather subtle. Obligate hermaphrodites presumably originate due to a breakdown in the normal sexual determination pathway as is hypothesized to be the case for accidental hermaphrodites. These breakdowns in the normal developmental mode then happen to be favored under certain ecological conditions as predicted by the low-density model (Ghiselin 1969). Natural selection can, of course, act to either optimize and reify the genetic controls of ovotestis development or remove these genetic variants from the population if they are of reduced fitness compared to the genes controlling dioecious development. The nature of selection (positive or negative) on a mutation causing accidental hermaphroditism will reflect the specific ecological details (e.g., community structure, water flow parameters) for that particular population. For example, Johnston et al. (1998) found a negative correlation between male tissue allocation within the ovotestes of hermaphroditic *Utterbackia imbecillis* from seven populations in the eastern United States and rates of self-fertilization (as measured by allozyme electrophoretic assessment of genetic variation at nine protein-coding loci). These results suggest that the control of production of male versus female acini within the ovotestis is susceptible to precise evolutionary manipulation. Because the seven (or so) appearances of hermaphroditism among the North American freshwater mussel fauna represent independent evolutionary events, the specific mutational events that lead to breakdown in dioecy in the ancestors to each of these species might not have been exactly the same in each case and indeed, broad comparative studies of plants and animals have shown that sex determination mechanisms are notoriously labile (e.g., Charlesworth and Mank 2010). We are interested in searching for commonalities in developmental control changes among these potentially diverse genetic systems. The most obvious commonalities associated with the independent evolution of hermaphroditism in North American freshwater mussels are listed above, namely, loss of the male mitochondrial genome entirely (and consequently the *M-orf* gene) and molecular "degeneration" of the *F-orf* gene and its transformation into a highly divergent *H-orf* gene. It is worth emphasizing that the *H-orf* is a functional category as the *H-orfs* in all of the freshwater mussel taxa examined by Breton et al. (2011) are examples of convergent evolution as opposed to being derived from a single ancestral gene. In other words, the *H-orf*s are not homologues.

## 16.3 The Next Step: Identification of Key Mutational Changes Associated with the Transition from Dioecy to Hermaphroditism in Freshwater Mussels

The period of transition from a population having separate sexes to exhibiting only hermaphrodites is critical to understanding the role of the novel *F-orf* and *M-orf* genes in sex determination in freshwater mussels. We know that accidental

hermaphroditism is widespread throughout the Unionoida but we are interested in identifying populations that are in the process of evolving into obligate hermaphrodites. Ideally, we will eventually identify multiple such populations including populations from various species of freshwater mussels. The literature on hermaphroditism in freshwater mussels suggests that populations living in small, recently isolated bodies of water are most likely to experience the selective pressures conducive to favoring obligate hermaphroditism (Ghiselin 1969; van der Schalie 1970; Bauer 1987). Our objective is to identify such populations and characterize their mitochondrial genomes early in the process described above. Eventually, these specimens would be used to perform western blot and immunoelectron-microscopical studies of females, males, and hermaphrodites from these populations and characterize expression patterns of the *F-orf*, *M-orf*, and proto-*H-orf* genes in these transitional populations. In concert, we are currently sequencing the nuclear genome of a dioecious unionoid species to study the co-evolution of mitochondrial and nuclear genes and, in particular, identify nuclear genes that are working in concert with mitochondrial genes to control the developmental aspects of DUI and sex determination (e.g., Zouros 2012).

As an example of our strategy to study the evolutionary transition from dioecy to hermaphroditism, we herein present some of our recent findings. We initially chose to focus on the transition from dioecy to hermaphroditism in species of *Anodonta* (Unionidae) from Germany. This decision was made based on a classic paper written by Weisensee (1916) that described potentially recent transitions to hermaphroditism in populations of *Anodonta* mussels living in small, isolated bodies of water recently cut off from the Rhine River in Germany (e.g., recently formed "ox-bow" ponds or other isolated bodies of water). We collected and analyzed samples from these potentially transitioning populations.

In brief, we followed the approach described in Breton et al. 2011, to identify, amplify, and sequence various mitochondrial genes in dioecious and hermaphroditic *Anodonta* individuals from slow versus moving waterways associated with the Rhine River. *Anodonta* mussels were obtained from Flückigersee, Freiburg, Germany (48.010004, 7.817888) and Lake Konstanz, Germany (47.666312, 9.181824). Following the method of Bauer 1987, needle punctures of the gonad were examined microscopically for the presence of either eggs or sperm and/or sperm morulae (indicative of a dioecious female or male specimen, respectively) or both eggs and sperm morulae (indicative of hermaphrodites). The presence of multinucleated sperm morulae indicates the process of intrinsic, atypical spermatogenesis (e.g., Shepardson et al. 2012) and is a better indicator of "maleness" (or hermaphroditism) of a given specimen than the presence of sperm, which could be from an extrinsic source (Bauer 1987).

To confirm the species identity of the collected *Anodonta* individuals, we PCR-amplified a portion of the cytochrome oxidase subunit 1 gene (*cox*1) using modified Folmer et al. 1994 primers LCO22me and HCO700dy2 (Walker et al. 2006a, b, 2007). We then used the Blast search routine to compare the resulting sequences with data available on GenBank (Altschul et al. 1997). To test for the presence of an *F-orf* or *H-orf* gene in females or hermaphrodites, respectively, we amplified a

portion of this open reading frame using the newly designed primers AnoForfF 5′-ctcggagaagatttgccttg-3′ and AnoForfR 5′-tycagrtgrttagycaattaga-3′. To test for the presence of an *M-orf* in males or hermaphrodites, we used the primer pair MORFunioF and MORFunioR (Breton et al. 2011). The resulting amplicons were either sequenced directly or cloned and sequenced.

Specimens from the Flückigersee, Freiburg, Germany, were identified as hermaphrodites based on the presence of both eggs and sperm morulae (Fig. 16.2). Specimens from the Lake Konstanz, Germany, were identified as either males or females based on the exclusive presence of sperm morulae or eggs, respectively. Based on the results of a Blast search, the Flückigersee samples were identical or nearly identical to published *Anodonta cygnea* sequences and the Lake Konstanz samples were identical to published *A. anatina* sequences. Amplification and sequencing with the AnoForfF and AnoForfR primers produced an amplicon from female and somatic tissues of male *Anodonta anatina* that was clearly homologous to published unionid *F-orfs* (Fig. 16.3). These same primers produced an amplicon that was approximately 800 bp long from hermaphrodites. Consistent with previous studies (Breton et al. 2011), the *A. cygnea H-orf* was highly divergent from the *A. anatina F-orf*. Indeed, the *H-orf* from *A. cygnea* could not be confidently aligned with the congeneric *F-orf* from *A. anatina*. Amplification with the M-ORFUnioF and M-*ORF*-UnioR primers failed to produce an amplicon from any female or hermaphroditic individuals. As expected with male samples, these same primers yielded a sequence that was clearly homologous to the unionoid *M-orf* sequences on GenBank (Fig. 16.4).

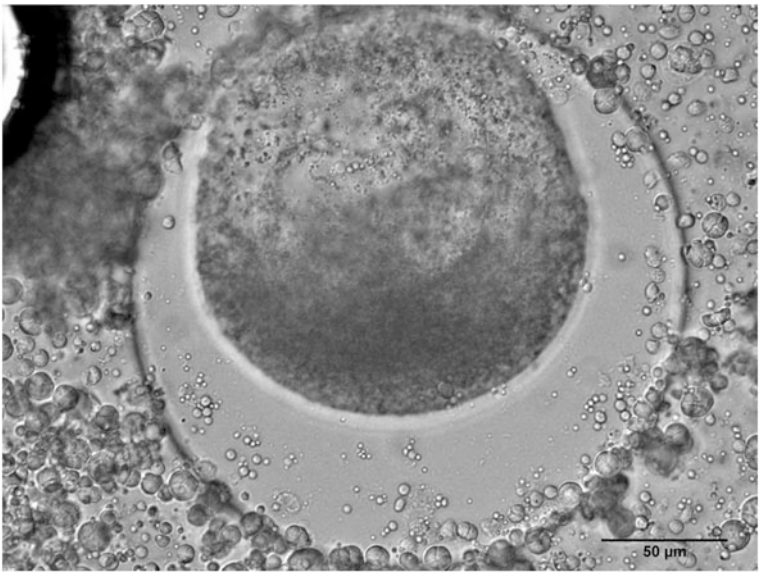



**Fig. 16.2** Micrograph image of a gonad pluck from a mussel identified as a hermaphroditic *Anodonta cygnea* showing both an egg and multiple sperm morulae

```
F-ORF_Anodonta_anatina          MKTILLILSIIILTILLAQAIQMFLTSSELWKIDQILCSMDLGSTIPQPE 50
F-ORF_Strophitus_undulatus      LKLFLLICLFSLLMLSFYQIIQVLLVPDEFWVASQILCSMDLGDTFPQPT 50
                                :* :***  : :* : : * **::*...*:*  .*********.*:***

F-ORF_Anodonta_anatina          TSDHPVIQSLASTDLTKPVIKP 72
F-ORF_Strophitus_undulatus      TSDHPVIQSPAKTDLTKSDAKP 72
                                ********* *.*****.  **
```

**Fig. 16.3** Alignment of the amino acid sequences for the F-ORF proteins from Anodonta anatina and Strophitus undulatus (GenBank Ref. no. AEC14102.1). An asterisk (*) indicates amino acid positions that are identical, a colon (:) indicates two biochemically strongly similar amino acids and a period (.) two biochemically weakly similar amino acids

Our initial objective was to study the molecular transitions in the unusual system of doubly uniparental inheritance of mitochondrial DNA associated with the shift from dioecy to hermaphroditism in unionoid mussels. We had hoped to identify individuals of *Anodonta anatina* that had recently made the evolutionary shift from separate sexes (i.e., dioecy) to hermaphroditism. However, the hermaphroditic individuals that we analyzed turned out to be from the species *A. cygnea*. This inference comes from DNA barcoding analysis of the *cox1* gene. However, our analytical results from these European dioecious and hermaphroditic freshwater mussel populations are highly important as they confirm patterns that had been previously documented for North American unionoids: (1) the dioecious species, *A. anatina,* has both an M and an F gender-associated mitochondrial genome, (2) the M and F genomes of *A. anatina* contain the unique open reading frames referred to as the *M-orf* and the *F-orf*, respectively, (3) the congeneric, obligately hermaphroditic species *A. cygnea* appears to have lost the male-transmitted genome as we found no evidence of an M-genome *cox1* gene or the unique *M-orf* in samples from that species, and (4) the fragment of the mitochondrial genome amplified by the AnoForfF and AnoForfR primers, which amplified the *F-orf* in *A. anatina,* amplified a highly divergent PCR product, the putative *H-orf*, from the hermaphroditic *A. cygnea*. As suggested by Breton et al. (2011), these patterns are consistent with the hypothesis that the *M-orf* and *F-orf* play a role in specifying dioecy in unionoid freshwater mussels and hence these genes are lost and drastically modified,

```
M-ORF_Anodonta_anatina          MDLLNDLITFVKWIKNCFALSPYVTLTMLFVFMMVIVGILRGIRLYGYDL 50
M-ORF_Utterbackia_peninsularis  ---MENFSVFFKWVKDCVMVSPYVTFVMFMVVMLLVLGFVRNVYLYWYEV 47
                                   :::: .*.**:*:*. :*****:.*::*.*::::*::*.: ** *::

M-ORF_Anodonta_anatina          YKKFMLLMVGSSWFTEKSGEAVKDEKSISDGLNPS-EGQKVSGVDGPKGE 99
M-ORF_Utterbackia_peninsularis  YKAFMTLMVGKSWFTDDPSVTTKETDTTNRSFDPSGEPKVVSNVKIVEDS 97
                                ** ** ****.****:... :.*: .: . .::** * : **.*.  :..

M-ORF_Anodonta_anatina          GGSIKVEDISSLKNMISEAVKEAIKEAMKDLVVKEAKKKKEKVVSGDGTS 149
M-ORF_Utterbackia_peninsularis  ----KVVDIT---GGVSTSVDSGIKDNK---------------SGS-DA 123
                                    ** **:   . :* :*...**:                 **.  :

M-ORF_Anodonta_anatina          PKKKSKKSLEVGEAVGDTVVVVSSKKKSSPKKKELVVKEEPSKEN 194
M-ORF_Utterbackia_peninsularis  PK----------EAVGLPKVENKNDLKSTIKEAVKEALEE----- 153
                                **          **** . *  ... **: *:    . **
```

**Fig. 16.4** Alignment of the amino acid sequences for the *M-ORF* proteins from *Anodonta anatina* and *Utterbackia peninsularis* (GenBank Ref. no. NC_015477.1). An asterisk (*) indicates amino acid positions that are identical, a colon (:) indicates two biochemically strongly similar amino acids and a period (.) two biochemically weakly similar amino acids

respectively, in hermaphrodites. Specifically, hermaphrodites lose the *M-orf* by losing the entire M-genome while they "lose" the *F-orf* gene via its "degeneration" into a highly modified *H-orf* that does not retain the conserved hydrophobicity profiles of or sequence similarity to the *F-orfs* (Breton et al. 2011).

In this particular study, we were not able to use the data collected to examine the immediate consequences of the transition from dioecy to hermaphroditism. *Anodonta cygnea* is not likely a recent convert to hermaphroditism. Indeed, *Anodonta anatina* and *Anodonta cygnea* might not be sister species. Recent phylogenetic analysis of north western European unionid mussels (Källersjö et al. 2005) suggests that the genus *Anodonta* is paraphyletic and that the depressed river mussel, *Pseudanodonta complanata*, is more closely related to *Anodonta cygnea* than is *A. anatina.* In contrast, Nagel et al. (1996) placed *A. cygnea* within a monophyletic *Anodonta* clade but it was the basal species within that clade. Nagel et al. (1996) suggested that *P. complanata* is the sister lineage to the Eurasian *Anodonta*. Furthermore, based on Nei's genetic distance values, Nagel et al. (1996) suggested that *A. cygnea* originated approximately 400–450,000 years ago in the middle Pleistocene. While there is clearly some taxonomic uncertainty regarding the *Anodonta-Pseudanodonta* complex, *A. cygnea* and *A. anatina* are not sister species and, therefore, they are not the most appropriate species to use for our comparative purposes.

Because of the highly divergent nature of the *H-orf* sequence obtained herein for *Anodonta cygnea* relative to the *F-orf* of *Anodonta anatina*, we plan to obtain samples from other European anodontine species including *Pseudanodonta complanata* and, subsequently, amplify and sequence their *F-orf*s to be able to accurately reconstruct the molecular evolutionary transition from the ancestral *F-orf* to the highly derived *H-orf* in *Anodonta cygnea.* Despite the challenges described herein, our methodology remains sound with the potential for significant gains in our understanding of breeding system transitions. In the future, we also will expand our survey of freshwater mussel populations living in recently formed "ox-bow" lakes in an attempt to examine these populations during the critical period of transition from dioecy to hermaphroditism. Analysis of DNA substitution and gene expression patterns for the various mitochondrial *orfs* will allow us to document the nature of the molecular changes that underlie the transitions in reproductive strategy in these species which is ultimately dependent upon the unusual system of doubly uniparental inheritance of mitochondrial DNA.

**Acknowledgements** Part of this work was initiated during a visiting fellowship to Gerhard Bauer's lab at Freiburg University for which D. Stewart is extremely grateful. This work was funded by a Discovery grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) to D. Stewart. D. Stewart's visiting fellowship was supported by a Harrison McCain Visitorship for Acadia Scholars Award (Acadia University) and by a Research Visit Grant from the Deutscher Akademischer Austausch Dienst (German Academic Exchange Service). S. Breton was financially supported by a CIHR post-doctoral fellowship. We thank M. Pfeiffer for collecting samples and Carola Sigle for technical assistance. D. Stewart also thanks Dr. J. Holzschuh, Freiburg University, for the use of his laboratory facilities while on sabbatical and K.–O. Nagel for discussions about hermaphroditism in the *Anodonta* species complex.

## References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Bauer G (1987) Reproductive strategy of the freshwater pearl mussel *Margaritifera margaritifera*. J Anim Ecol 56:691–704

Bengtsson BO (2009) Asex and evolution: a very large-scale overview. In: Schön I, Martens K, van Dijk P (eds) Lost sex: the evolutionary biology of parthenogenesis. Springer, New York, pp 1–21

Breton S, Doucet Beaupré H, Stewart DT, Hoeh WR, Blier PU (2007) The unusual system of doubly uniparental inheritance of mtDNA: Isn't one enough? Trends Genet 23:464–475

Breton S, Stewart DT, Shepardson S, Trdan RJ, Bogan AE, Chapman EG, Ruminas AJ, Piontkivska H, Hoeh WR (2011) Novel protein genes in animal mtDNA: a new sex determination system in freshwater mussels (Bivalvia: Unionoida)? Mol Biol Evol 28:1645–1659

Burt A, Trivers R (2006) Genes in conflict: the biology of selfish genetic elements. Belknap Press of Harvard University, Cambridge, MA

Charlesworth D, Mank JE (2010) The birds and the bees and the flowers and the trees: lessons from genetic mapping of sex determination in plants and animals. Genetics 186:9–31

Dalziel AC, Stewart DT (2002) Tissue-specific expression of male-transmitted mitochondrial DNA and its implications for rates of molecular evolution in *Mytilus* mussels (Bivalvia: Mytilidae). Genome 45:348–355

Davison A (2006) The ovotestis: an underdeveloped organ of evolution. BioEssays 28:642–650

Doucet-Beaupré H, Breton S, Chapman EG, Blier PU, Bogan AE, Stewart DT, Hoeh WR (2010) Mitochondrial phylogenomics of the Bivalvia: searching for the origin and mitogenomic correlates of doubly uniparental inheritance of mtDNA. BMC Evol Biol 10:50

Folmer O, Black M, Hoeh WR, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotech 3:294–299

Garrido-Ramos MA, Stewart DT, Sutherland BW, Zouros E (1998) The distribution of male-transmitted and female-transmitted mitochondrial DNA types in somatic tissues of blue mussels: Implications for the operation of doubly uniparental inheritance of mitochondrial DNA. Genome 41:818–824

Ghiselin MT (1969) The evolution of hermaphroditism among animals. Quarterly Review of Biology 44: 189–208

Heard WH (1975) Sexuality and other aspects of reproduction in *Anodonta* (Pelecypoda: Unionidae). Malacologia 15:81–103

Hoeh WR, Frazer KS, Naranjo-Garcia E, Trdan RJ (1995) A phylogenetic perspective on the evolution of simultaneous hermaphroditism in a freshwater mussel clade (Bivalvia: Unionidae: Utterbackia). Malacological Rev 28:25–42

Hoeh WR, Frazer K, Naranjo-Garcia E, Black MB, Berg DJ, Guttman SI (1998) Correlation between mating system and distribution of genetic variation in Utterbackia (Bivalvia: Unionidae). J Shellfish Res 17:1383–1393

Hoeh WR, Stewart DT, Guttman SI (2002) High fidelity of mitochondrial genome transmission under the doubly uniparental mode of inheritance in freshwater mussels (Bivalvia:Unionoidea). Evolution 56:2252–2261

Johnston MO, Das B, Hoeh WR (1998) Negative correlation between male allocation and rate of self-fertilization in a hermaphroditic animal. Proc Nat Acad Sci U.S.A 95:617–620

Källersjö M, von Proschwitz T, Lundberg S, Eldenäs P, Erséus C (2005) Evaluation of ITS rDNA as a complement to mitochondrial gene sequences for phylogenetic studies in freshwater mussels: an example using unionidae from north-western Europe. Zoolog Scr 34:415–424

Lane N (2005) Power, sex, suicide: mitochondria and the meaning of life. Oxford University Press, New York

Nagel K-O, Basino G, Alessandria B (1996) Population genetics of European Anodontinae (Bivalvia: Unionidae). J Molluscan Stud 62:343–357

Sano N, Obata M, Komaru A (2010) Mitochondrial DNA transmitted from sperm in the blue mussel *Mytilus galloprovincialis* showing doubly uniparental inheritance of mitochondria, quantified by real-time PCR. Zoolog Sci 27:611–614

Shepardson SP, Heard WH, Breton S, Hoeh WR (2012) Light and transmission electron microscopy of two spermatogenic pathways and unimorphic spermatozoa in *Venustaconcha ellipsiformis* (Conrad, 1836) (Bivalvia: Unionoida). Malacologia 55:263–284

Stewart DT, Breton S, Blier PU, Hoeh WR (2009) Masculinization events and doubly uniparental inheritance of mitochondrial DNA: a model for understanding the evolutionary dynamics of gender-associated mtDNA in mussels. In: Pontarotti P (ed) Evolutionary biology: concepts, modeling and application. Springer, Berlin, pp 163–174

Takebayahi N, Morrell PL (2001) Is self-fertilization an evolutionary dead end? Revisiting an old hypothesis with genetic theories and a macroevolutionary approach. Am J Bot 88:1143–1150

van der Schalie H (1970) Hermaphroditism among North American freshwater mussels. Malacologia 10:93–112

Walker JM, Bogan AE, Garo K, Soliman GN, Hoeh WR (2006a) Hermaphroditism in the Iridinidae (Bivalvia: Etherioidea). J Molluscan Stud 72:216–217

Walker JM, Curole JP, Wade DE, Chapman EG, Bogan AE, Watters GT, Hoeh WR (2006b) Taxonomic distribution and phylogenetic utility of gender-associated mitochondrial genomes in the Unionoida (Bivalvia). Malacologia 48:265–284

Walker JM, Bogan AE, Bonfiglio EA, Campbell DC, Christian AD, Curole JP, Harris JL, Wojtecki RJ, Hoeh WR (2007) Primers for amplifying the hypervariable, male-transmitted COII-COI junction region in amblemine freshwater mussels (Bivalvia: Unionoidea: Ambleminae). Mol Ecol Notes 7:489–491

Weisensee H (1916) Die Geschlechtsverhältnisse und der Geschlechtsapparat bei. Anodonta. Zeitschrift für Wissenschaftliche Zoologie 115:262–335

Zouros E (2012) Biparental inheritance through uniparental transmission: the doubly uniparental inheritance (DUI) of mitochondrial DNA. Evol Biol. doi:10.1007/s11692-012-9195-2

# Chapter 17
# Variable Lymphocyte Receptors in Jawless Vertebrates: Illuminating the Origin and Early Evolution of Adaptive Immunity

**Sabyasachi Das and Masayuki Hirano**

**Abstract** Phylogenetic studies of immunity have revealed that about 500 million years ago, two types of recombinatorial adaptive immune systems (AIS) arose in vertebrates. Jawed vertebrates diversify the repertoire of immunoglobulin-domain-based B and T cell antigen receptors mainly through the rearrangement of gene segments and somatic hypermutation, whereas an alternative AIS of jawless vertebrates is based on variable lymphocyte receptors (VLRs) which diversify through recombinatorial usage of leucine-rich repeat (LRR) units. None of the fundamental recognition elements of jawed vertebrates AIS have been found in jawless vertebrates. Despite differences in molecular architecture, the parallel 'two-arms' of the AIS evolved within the context of preexisting innate immunity and maintained over a long period of time in jawed and jawless vertebrates, respectively, as a consequence of powerful and enduring evolutionary selection pressure by pathogens and other factors.

## 17.1 Introduction

In order to survive in the competitive 'struggle for existence' in an environment, organisms must be able to protect themselves from pathogens. This requirement of self-defense in the ongoing struggle for survival led to the evolution of complex immune systems. Biologists have found that simplest multicellular life forms like sponges (phylum Porifera) have many of the elements used by vertebrates for immune response and pathogen defense. These ancient defense strategies defend against infection by potential pathogens in a relatively non-specific manner, collectively known as innate immunity, which can be found in representative species

S. Das (✉) · M. Hirano
Emory Vaccine Center, and Department of Pathology and Laboratory Medicine, Emory University, 1462 Clifton Road NE, DSB 405, Atlanta, GA 30322, USA
e-mail: sdas8@emory.edu

P. Pontarotti (ed.), *Evolutionary Biology: Exobiology and Evolutionary Mechanisms*,
DOI: 10.1007/978-3-642-38212-3_17, 

at almost every level of the evolutionary tree of life. In addition to the innate immunity, another layer of complexity in immune defenses emerged during animal evolution with the appearance of adaptive immunity in vertebrates around 500 million years ago (Boehm et al. 2012; Cooper and Alder 2006). The adaptive immune system (AIS) allows specific recognition and mounting of a protective response against numerous pathogens. Adaptive immunity is mediated through various genetic and cellular processes that create appropriate somatic variants of antigen-binding receptors under evolutionary pressure by pathogens and other factors (Flajnik and Kasahara 2010). The core elements of AIS are now mechanistically understood, such as compartmental differentiation of lymphocytes, the generation of immune recognition diversity and the supporting cellular complexity that selects and expands cell populations expressing suitable antigen-binding receptor variants and immunological memory. Immunologists and evolutionary biologists long believed that these general features of the adaptive immunity are exclusive to the jawed vertebrates. However, the recent discovery of a lymphoid cell-based system of adaptive immunity in jawless vertebrates (lamprey and hagfish) was surprisingly similar to the AIS in jawed vertebrates (Pancer et al. 2004, 2005; Rogozin et al. 2007; Guo et al. 2009). We are within reach of important breakthroughs in our understanding of how AIS evolved in the context of well-developed innate immunity and how these molecularly disparate systems are related to the evolutionary acquisition of immunological complexity.

## 17.2 Immune Response Molecules in Invertebrates

Common elements deployed for innate immune defense in invertebrates may provide insight into how and when our complex AIS evolved. Two protein families, which contain either the leucine-rich-repeat (LRR) motifs or the immunoglobulin superfamily (IgSF) domains, are widely involved for immune defense. Leucine-rich repeat containing proteins are consisted of multiples of 20–30 amino acid units to form horseshoe-like solenoid structures in which the concave surface is formed by parallel $\beta$ sheets and the convex surface by an array of helices (Buchanan and Gay 1996). The Toll-like receptors (TLRs) are well-defined examples of LRR-containing proteins which function as pattern-recognition receptors (PRRs) that constitute key components of innate immune systems throughout the animal kingdom (Hoffmann et al. 1999).

Members of the IgSF also serve important immune defense functions in invertebrates. IgSF members with important roles in innate immunity include the fibrinogen-related proteins (FREPs) in snails (Zhang et al. 2004), down syndrome cell adhesion molecule (Dscam) in insects (Watson et al. 2005), variable region-containing chitin-binding proteins (VCBPs) in amphioxus and sea squirt (Cannon et al. 2002). These molecules undergo repertoire diversification by alternative splicing or even somatic mutation to generate potential antigen recognition capacity. Both LRRs motifs and IgSF domains were readily available for co-option

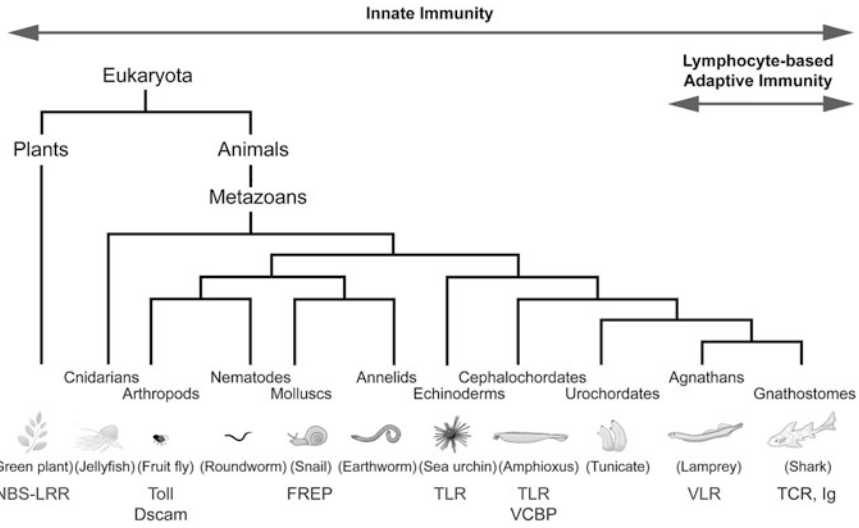



**Fig. 17.1** Overview of the evolution of innate and adaptive immune systems. The stages at which different families of immune molecules emerged are shown in the species phylogeny. The figure is modified from Hirano et al. (2011)

to provide the molecular basis for use in the somatic diversification of variable lymphocyte receptor (VLR) in jawless vertebrates or immunoglobulin (Ig)/T cell receptor (TCR) in jawed vertebrates. The hypothetical evolutionary scheme of innate and AIS is shown in Fig. 17.1.

## 17.3 Brief Overview of the AIS of Jawed Vertebrates

The most genetically distant organisms in which the Ig-domain-based AIS, as characterized in humans, is found are the cartilaginous fish. The well-defined mechanisms of the Ig-based AIS are mentioned briefly here for comparative purposes. The two major lineages of clonally diverse lymphocytes that specifically recognize and respond to antigens are named T and B lymphocytes, because they are generated in the thymus and the bone marrow (or the avian bursa of Fabricius), respectively (Cooper et al. 1965; Greaves et al. 1968). During their early developmental stages, the T and B lymphocyte progenitors rearrange different sets of variable (*V*), diversity (*D*), and/or joining (*J*) gene segments to generate the antigen-binding regions of the TCRs and B cell receptors (BCRs) (Tonegawa 1983). The recombination activating genes (RAG1/RAG2) encode enzymes that mediate *V*(*D*)*J* rearrangement (Schatz et al. 1989). The antigen-binding regions of the different *V*(*D*)*J* combinations are diversified further through splicing variability and the enzymatic addition of nucleotides in the joints created during *V*(*D*)*J* segment assembly (Dudley et al. 2005). This random nature of diversification

inevitably results in the generation of Ig/TCR repertoire expansion. The membrane bound and secreted antibodies made by B lineage cells typically recognize exposed determinants (epitopes) of intact molecules, including surface protein and carbohydrate moieties of invasive microbes (Klein and Horejsi 1997). In contrast, the TCRs recognize peptide fragments of antigens presented by accessory cells within cell surface molecules encoded by the major histocompatibility complex (*MHC*) class I and class II genes (Klein and Horejsi 1997). Therefore, TCRs typically recognize antigens that have been partially digested within the antigen presenting cells, primarily dendritic cells and phagocytic cells.

The acquisition of a mechanism for gene rearrangement to produce clonally diverse *Igs* and *TCRs* was critical for adaptive immunity in jawed vertebrates. Discovery of multiple *V*, *D*, and *J* gene segments with specific recombination signal sequences (RSSs) provided insight into the recombinatorial system employed in *Ig* and *TCR* loci to generate clonal diversity (Hedrick et al. 1984; Tonegawa 1983; Yanagi et al. 1984). The RAG1/RAG2 proteins were found to recognize the RSSs flanking the *V(D)J* gene segments to initiate the double-stranded DNA breaks and the recruitment of other proteins required for recombination (Jung and Alt 2004; Oettinger et al. 1990). The RAG1 and RAG2 proteins form a transposase that can excise DNA containing the RSSs and reinsert it elsewhere, thus supporting the theory that *RAG*1/*RAG*2 genes were once components of a transposable element (Agrawal et al. 1998; Hiom and Gellert 1997).

## 17.4 VLR-Based Alternative Adaptive Immune System in Jawless Vertebrates

An alternative AIS that uses Variable lymphocyte receptor as antigen receptors has been discovered in extant jawless vertebrates (lampreys and hagfish) only recently (Pancer et al. 2004, 2005; Rogozin et al. 2007). Jawless vertebrates mount specific responses to pathogens, elicit allograft recognition, and display other general immune-type responses that are characteristic of cellular immunity, but they do not use a *V(D)J* recombination-mediated form of adaptive immunity as found jawed vertebrates (Hirano et al. 2011; McCurley et al. 2012). The VLR-based alternative AIS of jawless vertebrates displays an anticipatory receptor repertoire complexity comparable to that of the Ig-domain-based AIS of jawed vertebrates.

The thymus-derived T lymphocytes and bone marrow-derived B lymphocytes are the cellular pillars of adaptive immunity in the jawed vertebrates. Cells with similar morphological features and molecular machinery of lymphocytes in jawed vertebrates are also found in lampreys and hagfish (Mayer et al. 2002; Nagata et al. 2002; Najakshin et al. 1999; Uinuk-Ool et al. 2002). These findings along with earlier observations that lampreys and hagfish produce specific agglutinins following immunization with bacteria and foreign red blood cells suggested that jawless vertebrates could have an AIS. However, from transcriptome analysis no

*MHC*, *TCR*, *BCR*, and *RAG* orthologues genes were found in jawless vertebrates, and this failure added to skepticism about the presence of adaptive immunity in agnathans. This view was dramatically changed when the *VLR* genes were identified as the key elements for AIS in lampreys and hagfish (Pancer et al. 2004, 2005).

### *17.4.1 VLR Discovery and Diversity Generation in Jawless Vertebrates*

Since the transcriptome analysis of lymphocyte-like cells of naïve lamprey did not reveal evidence for equivalent AIS of jawed vertebrates, lamprey larvae were stimulated by an antigen and mitogen mixture to survey the transcriptome of activated lamprey lymphocytes. The objective was to catch the lamprey lymphocytes in the act of an immune response. Large lymphoblastoid cells in blood were then sorted by their light scatter characteristics and used for the construction of a cDNA library (Pancer et al. 2004). Still no orthologs of *Ig*, *TCR*, and *MHC* genes were detected, but this experiment revealed a large number of transcripts for uniquely diverse Leucine-rich repeat proteins, which were named VLRs because of their lymphocyte-restricted expression and sequence diversity. Each VLR transcript was found to encode a conserved signal peptide (SP) followed by highly variable LRR modules: a 27–38 residue N-terminal LRR (LRRNT), the first 18-residue LRR (LRR1), several 24-residue variable LRRs (LRRV), one 13-residue connecting peptide LRR (LRRCP), and a 48–65 residue C-terminal LRR (LRRCT).

After the discovery of the first lamprey *VLR* gene (now known as *VLRB*), two hagfish *VLR* genes, *VLRA* and *VLRB*, were identified in an expressed sequence tags (EST) database of hagfish leukocyte transcripts (Pancer et al. 2005). The *VLRA* gene in lamprey was identified in a subsequent search of the draft sequence database of the sea lamprey genome (Rogozin et al. 2007). The sequence homology indicates that the hagfish *VLRA* and *VLRB* genes are homologous to the lamprey *VLRA* and *VLRB* genes, respectively. Another *VLR* gene, designated *VLRC*, has been identified recently through an analysis of the sea lamprey EST database (Kasamatsu et al. 2010). The predicted VLRC structure is very similar to that of lamprey VLRA and VLRB, except that VLRC lacks the thumb-like protrusion encoded in the LRRCT inserts of VLRA and VLRB that are important for antigen recognition. Phylogenetically, the *VLRC* gene is close to *VLRA* genes of lamprey and hagfish (Kasamatsu et al. 2010). The discovery of VLRC raises interesting questions about the function of VLRC-expressing lymphocytes, their antigen-binding potential, and their role in pathogen responses. Whether or not hagfish possess a *VLRC* homolog is presently unknown.

All of the germline *VLR* genes are incomplete, in that they have coding sequences only for the leader sequence, incomplete amino- and carboxy-terminal

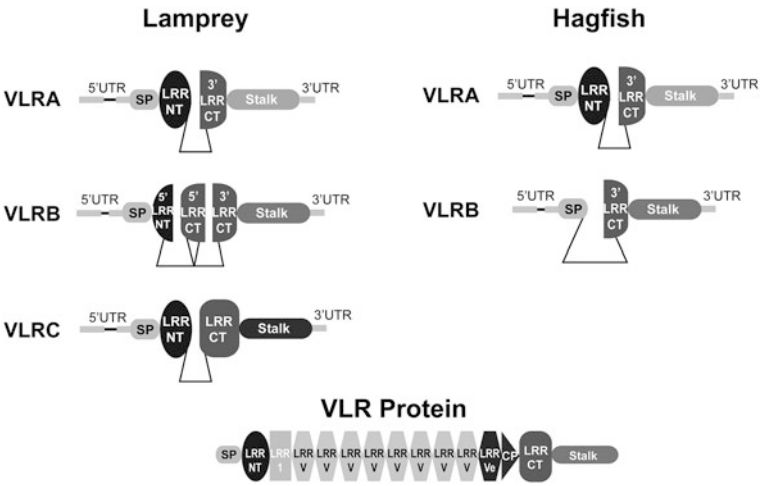



**Fig. 17.2** Organization of *VLR* genes and assembled VLR protein. Germline *VLR* genes in lamprey and hagfish are shown in *upper panel*. The incomplete germline *VLR* genes contain regions encoding for portions of the *LRRNT* and *LRRCT* separated by noncoding intervening sequences and for the invariant stalk region. Assembled VLR protein. Mature VLR protein consist of a signal peptide (*SP*), an LRRNT, an LRR1, up to eight LRRV modules, a connecting peptide (*CP*), an LRRCT, and an invariant stalk region. The figure is modified from Hirano et al. (2011)

*LRR* subunits (*LRRNT* and *LRRCT*), and for the stalk region (Fig. 17.2) (Hirano et al. 2011). There are two exons; the first exon encodes only a portion of the 5′ untranslated region while the second exon encodes the rest of the 5′ untranslated region, a sig, a 5′ portion of the *LRRNT*, a 3′ portion of the *LRRCT*, and the stalk region. For hagfish *VLRA* and *VLRB* and for lamprey *VLRA*, the 5′ *LRRNT* sequence is separated from the 3′ *LRRCT* sequence by a short noncoding intervening sequence that does not contain canonical splice donor and acceptor sites. The lamprey *VLRB* gene is more complex in that it has a *5′ LRRNT* coding sequence present between two intervening sequences (Pancer et al. 2004). Each germline *VLR* gene is flanked by hundreds of different LRR module-encoding genomic donor cassettes, which are used as randomly selected templates to add the *LRR* cassettes needed for production of a mature *VLR* gene (Fig. 17.3). From a genomic point of view, the incomplete germline *VLR* gene resembles the constant gene segment, whereas the clusters of different types of donor genomic cassettes represent the functional correlates of the clusters of variable, diversity, and joining gene segments of *Ig* or *TCR* loci in jawed vertebrates (Das et al. 2013).

A gene conversion-like mechanism has been postulated for the complex *VLR* gene assembly process (Cooper and Alder 2006; Nagawa et al. 2007) in which the intervening sequence is replaced in a stepwise manner by random selection of flanking *LRR* cassettes to serve as templates for adding the necessary sequences to

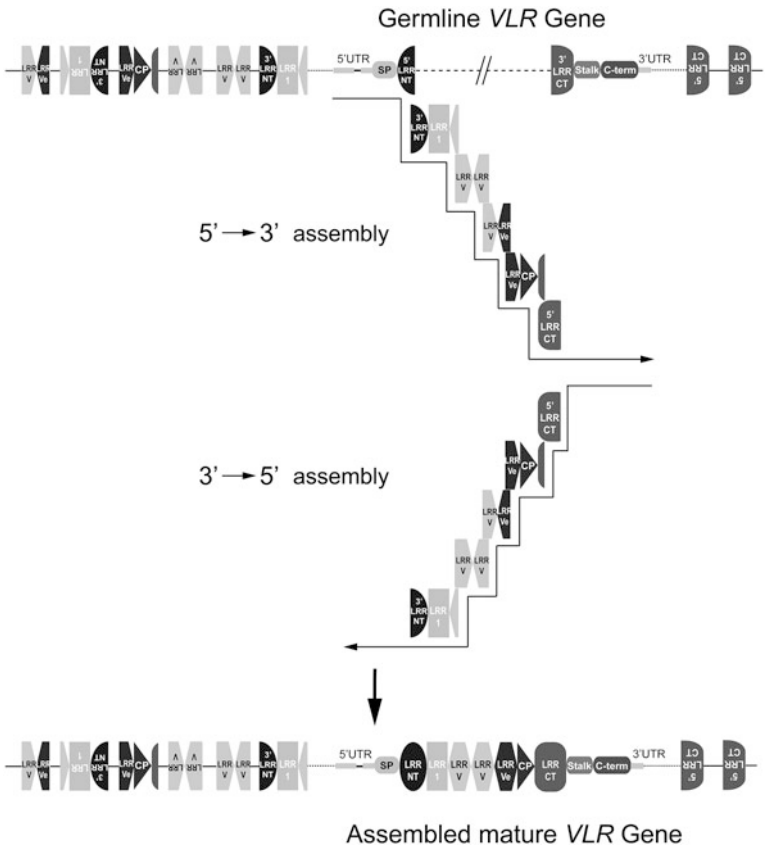



**Fig. 17.3** Mechanism of *VLR* gene assembly. The germline VLR genes are flanked by hundreds of donor *LRR* cassettes. The noncoding intervening sequence between portions of the *LRRNT* and *LRRCT* is replaced by the step-wise incorporation of genomic donor cassettes based on a short region of sequence homology (∼6–30 nt. The figure is modified from Hirano et al. (2011)

complete a *VLR* gene (Fig. 17.3). However, recent study indicates preferential usage of certain donor cassettes during development of the VLR repertoire (Das et al. 2013). The assembly process can be initiated at either the 5′ *LRRNT* or the 3′ *LRRCT* ends (Alder et al. 2005; Nagawa et al. 2007; Das et al. 2013) using short stretches of nucleotide homology (6–30 bp) between donor sequences and acceptor sequences (Nagawa et al. 2007; Das et al. 2013). *In silico* analysis of the diversity of *VLRB* gene sequences suggests a potential repertoire of >1,014 distinct receptors, i.e., a magnitude comparable to that of the theoretical repertoire diversity of our AIS (Alder et al. 2005).

## 17.5 Ancient T-like and B-like Lymphocyte Populations in Jawless Vertebrates

The lymphocytes expressing the VLRA or VLRB in lamprey are remarkably similar to T and B lymphocytes in jawed vertebrates (Guo et al. 2009). The lamprey VLRB-expressing lymphocytes resemble B cells of jawed vertebrates in that they respond to immunization with pathogens or foreign erythrocytes with proliferation, lymphoblastoid transformation and differentiation into plasmacytes that secrete VLRB antibodies specific for protein or carbohydrate epitopes (Alder et al. 2008; Herrin et al. 2008). On the other hand, the VLRA-expressing lymphocytes also respond to immunization, but they do not secrete the VLRA proteins either before or after immunization (Hirano et al. 2011). VLRB cells can bind to native antigens directly; however, it is still not clear whether VLRA cells could recognize native or processed antigens.

A limited transcriptome analysis indicates that VLRA and VLRB cells have different gene expression profiles. $VLRB^+$ lymphocytes preferentially express orthologues of genes that are preferentially expressed by B cells in jawed vertebrates (Guo et al. 2009): the hematopoietic progenitor homing receptor CXCR4, two components of the BCR-mediated signaling cascades, Syk and the B cell adaptor protein (BCAP), the chemotactic inflammatory cytokine IL-8, the IL-17 receptor, and the TLR orthologues TLR2abc, TLR7, and TLR10, the ligation of which is important for B cell activation. By contrast, $VLRA^+$ lymphocytes preferentially express genes orthologues to those typically expressed by the T cells in jawed vertebrates: GATA1/2, c-Rel, aryl hydrocarbon receptor (AHR), and BCL11b transcriptional factors used for T cell differentiation, the CCR9 chemokine receptor involved in homing of lymphocyte progenitors to the thymus, the Notch1 T cell fate-determining molecule, the CD45 tyrosine phosphatase receptor protein that is essential for T cell development, and proinflammatory cytokine genes interleukin-17 (IL-17) (Guo et al. 2009). All these evidence indicate that characteristics of $VLRA^+$ and $VLRB^+$ lymphocytes resemble those of mammalian T and B cells, respectively. Therefore, it can be postulated that the T-like and B-like cells were evolved before the divergence of jawless and jawed vertebrates (Fig. 17.4).

## 17.6 Conclusion

The discovery of VLR-type antigen receptors in lamprey and hagfish has provided unparalleled insight into the origins of adaptive immunity at the early stages of vertebrate evolution. Both VLRs in jawless vertebrates and BCRs/TCRs in jawed vertebrates rely on combinatorial diversity to generate a vast repertoire of immune receptors; however, the genes of Variable lymphocyte receptors show no structural

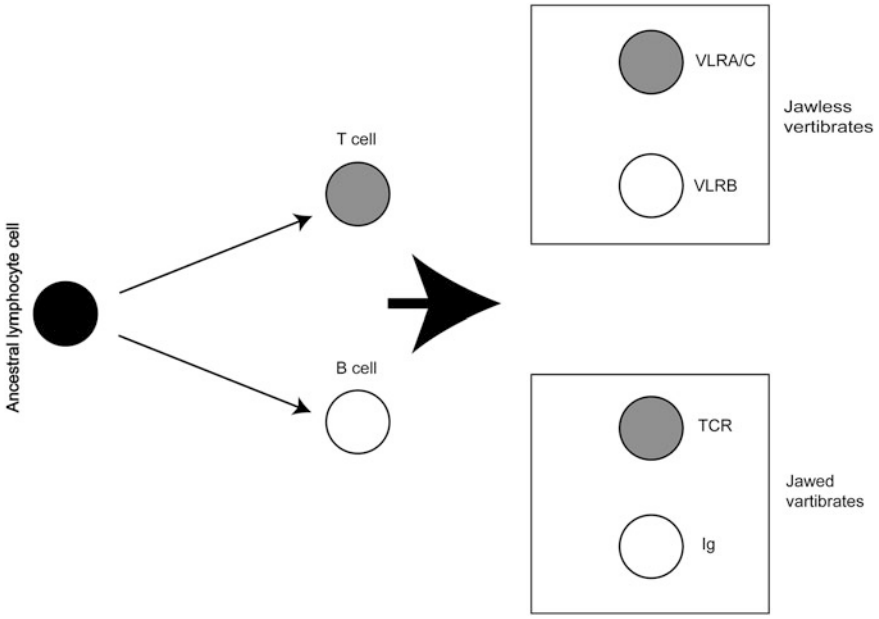



**Fig. 17.4** Schematic representation of the emergence and evolution of B-type and T-type lymphocytes in jawless and jawed vertebrates

similarity to those of BCRs or TCRs. *VLR* genes probably exist only in jawless vertebrates, whereas *Ig*, *TCR,* and *MHC* genes are possibly restricted in jawed vertebrates. Recent studies on the immune system of jawless vertebrates indicate that the preliminary separation of T and B lineage cells occurred earlier than the emergence of an anticipatory receptor system in the common ancestor of jawed and jawless vertebrates. Considering the monophyletic relationship of lamprey and hagfish (Takezaki et al. 2003; Heimberg et al. 2010), it is possible that the VLR-based antigen receptors in jawless vertebrates and Ig-domain-based antigen receptors in jawed vertebrates developed independently. In an alternative scenario, it is also possible that if the common ancestor possessed both VLR- and Ig-domain-based anticipatory receptors; after separation, the jawed and jawless vertebrates lost their VLR-based and Ig-domain-based anticipatory receptors, respectively. Taking into account the striking similarities in cellular and molecular basis of immune function, studies of the alternative AIS in jawless vertebrates may yield insight into several aspects of adaptive immunity in jawed vertebrates that remain elusive. Furthermore, the parallelism of lymphocyte lineages in jawed and jawless vertebrates may also guide in future studies on the origins of the different hematopoietic lineages in non-vertebrate species.

## References

Agrawal A, Eastman QM, Schatz DG (1998) Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. Nature 394:744–751

Alder MN, Rogozin IB, Iyer LM, Glazko GV, Cooper MD, Pancer Z (2005) Diversity and function of adaptive immune receptors in a jawless vertebrate. Science 310:1970–1973

Alder MN, Herrin BR, Sadlonova A, Stockard CR, Grizzle WE, Gartland LA, Gartland GL, Boydston JA, Turnbough CL Jr, Cooper MD (2008) Antibody responses of variable lymphocyte receptors in the lamprey. Nat Immunol 9:319–327

Boehm T, McCurley N, Sutoh Y, Schorpp M, Kasahara M, Cooper MD (2012) VLR-based adaptive immunity. Annu Rev Immunol 30:203–220

Buchanan SG, Gay NJ (1996) Structural and functional diversity in the leucine-rich repeat family of proteins. Prog Biophys Mol Biol 65:1–44

Cannon JP, Haire RN, Litman GW (2002) Identification of diversified genes that contain immunoglobulin-like variable regions in a protochordate. Nat Immunol 3:1200–1207

Cooper MD, Alder MN (2006) The evolution of adaptive immune systems. Cell 124:815–822

Cooper MD, Peterson RD, Good RA (1965) Delineation of the thymic and bursal lymphoid systems in the chicken. Nature 205:143–146

Das S, Hirano M, Aghaallaei N, Bajoghli B, Boehm T, Cooper MD (2013) Organization of lamprey variable lymphocyte receptor C locus and repertoire development. Proc Natl Acad Sci USA 110:6043–6048

Dudley DD, Chaudhuri J, Bassing CH, Alt FW (2005) Mechanism and control of V(D)J recombination versus class switch recombination: similarities and differences. Adv Immunol 86:43–112

Flajnik MF, Kasahara M (2010) Origin and evolution of the adaptive immune system: genetic events and selective pressures. Nat Rev Genet 11:47–59

Greaves MF, Roitt IM, Rose ME (1968) Effect of bursectomy and thymectomy on the responses of chicken peripheral blood lymphocytes to phytohaemagglutinin. Nature 220:293–295

Guo P, Hirano M, Herrin BR, Li J, Yu C, Sadlonova A, Cooper MD (2009) Dual nature of the adaptive immune system in lampreys. Nature 459:796–801

Hedrick SM, Cohen DI, Nielsen EA, Davis MM (1984) Isolation of cDNA clones encoding T cell-specific membrane-associated proteins. Nature 308:149–153

Heimberg AM, Cowper-Sal-lari R, Semon M, Donoghue PC, Peterson KJ (2010) microRNAs reveal the interrelationships of hagfish, lampreys, and gnathostomes and the nature of the ancestral vertebrate. Proc Natl Acad Sci USA 107:19379–19383

Herrin BR, Alder MN, Roux KH, Sina C, Ehrhardt GR, Boydston JA, Turnbough CL Jr, Cooper MD (2008) Structure and specificity of lamprey monoclonal antibodies. Proc Natl Acad Sci USA 105:2040–2045

Hiom K, Gellert M (1997) A stable RAG1-RAG2-DNA complex that is active in V(D)J cleavage. Cell 88:65–72

Hirano M, Das S, Guo P, Cooper MD (2011) The evolution of adaptive immunity in vertebrates. Adv Immunol 109:125–157

Hoffmann JA, Kafatos FC, Janeway CA, Ezekowitz RA (1999) Phylogenetic perspectives in innate immunity. Science 284:1313–1318

Jung D, Alt FW (2004) Unraveling V(D)J recombination; insights into gene regulation. Cell 116:299–311

Kasamatsu J, Sutoh Y, Fugo K, Otsuka N, Iwabuchi K, Kasahara M (2010) Identification of a third variable lymphocyte receptor in the lamprey. Proc Natl Acad Sci USA 107:14304–14308

Klein J, Horejsi V (1997) Immunology. Blackwell Science, Oxford

Mayer WE, Uinuk-Ool T, Tichy H, Gartland LA, Klein J, Cooper MD (2002) Isolation and characterization of lymphocyte-like cells from a lamprey. Proc Natl Acad Sci USA 99:14350–14355

McCurley N, Hirano M, Das S, Cooper MD (2012) Immune related genes underpin the evolution of adaptive immunity in jawless vertebrates. Curr Genomics 13:86–94

Nagata T, Suzuki T, Ohta Y, Flajnik MF, Kasahara M (2002) The leukocyte common antigen (CD45) of the Pacific hagfish, Eptatretus stoutii: implications for the primordial function of CD45. Immunogenetics 54:286–291

Nagawa F, Kishishita N, Shimizu K, Hirose S, Miyoshi M, Nezu J, Nishimura T, Nishizumi H, Takahashi Y, Hashimoto S et al (2007) Antigen-receptor genes of the agnathan lamprey are assembled by a process involving copy choice. Nat Immunol 8:206–213

Najakshin AM, Mechetina LV, Alabyev BY, Taranin AV (1999) Identification of an IL-8 homolog in lamprey (*Lampetra fluviatilis*): early evolutionary divergence of chemokines. Eur J Immunol 29:375–382

Oettinger MA, Schatz DG, Gorka C, Baltimore D (1990) RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. Science 248:1517–1523

Pancer Z, Amemiya CT, Ehrhardt GR, Ceitlin J, Gartland GL, Cooper MD (2004) Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. Nature 430:174–180

Pancer Z, Saha NR, Kasamatsu J, Suzuki T, Amemiya CT, Kasahara M, Cooper MD (2005) Variable lymphocyte receptors in hagfish. Proc Natl Acad Sci USA 102:9224–9229

Rogozin IB, Iyer LM, Liang L, Glazko GV, Liston VG, Pavlov YI, Aravind L, Pancer Z (2007) Evolution and diversification of lamprey antigen receptors: evidence for involvement of an AID-APOBEC family cytosine deaminase. Nat Immunol 8:647–656

Schatz DG, Oettinger MA, Baltimore D (1989) The V(D)J recombination activating gene, RAG-1. Cell 59:1035–1048

Takezaki N, Figueroa F, Zaleska-Rutczynska Z, Klein J (2003) Molecular phylogeny of early vertebrates: monophyly of the agnathans as revealed by sequences of 35 genes. Mol Biol Evol 20:287–292

Tonegawa S (1983) Somatic generation of antibody diversity. Nature 302:575–581

Uinuk-Ool T, Mayer WE, Sato A, Dongak R, Cooper MD, Klein J (2002) Lamprey lymphocyte-like cells express homologs of genes involved in immunologically relevant activities of mammalian lymphocytes. Proc Natl Acad Sci USA 99:14356–14361

Watson FL, Puttmann-Holgado R, Thomas F, Lamar DL, Hughes M, Kondo M, Rebel VI, Schmucker D (2005) Extensive diversity of Ig-superfamily proteins in the immune system of insects. Science 309:1874–1878

Yanagi Y, Yoshikai Y, Leggett K, Clark SP, Aleksander I, Mak TW (1984) A human T cell-specific cDNA clone encodes a protein having extensive homology to immunoglobulin chains. Nature 308:145–149

Zhang SM, Adema CM, Kepler TB, Loker ES (2004) Diversification of Ig superfamily genes in an invertebrate. Science 305:251–254

# Chapter 18
# An Evo-Devo Perspective on Hybrid Infertility and Speciation

**Priscilla Ambrosi, Sebastian Chahda, Emma Yang, Rui Sousa-Neves and Claudia M. Mizutani**

**Abstract** Understanding how new species arise and become reproductively isolated lineages is a fundamental problem of evolutionary biology and can provide us insights into the mechanisms of fertility, genomic incompatibility, and behavioral choice. Here we focus on hybrid infertility, a reproductive barrier that prevents gene flow among different species, with a special emphasis on hybrid female sterility among *Drosophila* species. We briefly review germline development in flies, describe classical interspecific transplantation experiments of germline cells, and discuss how hybrid infertility genes may be uncovered using current genomics and developmental biology tools.

## 18.1 Introduction

The simplest view of the process of speciation can be described as the separation of two branching populations of organisms that descend from a common ancestor. Within time, each population accumulates genetic changes such that even if they can still recognize each other, court and mate, their hybrid progeny is incompatible and displays lethality and/or infertility. At this point, gene flow between the two populations is blocked and species are in separate evolutionary paths.

P. Ambrosi · S. Chahda · E. Yang · C. M. Mizutani (✉)
Department of Biology, Case Western Reserve University, Cleveland, OH 44106, USA
e-mail: claudia.mizutani@case.edu

P. Ambrosi
Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

R. Sousa-Neves · C. M. Mizutani
Department of Genetics and Genome Sciences, Case Western Reserve University, Cleveland, OH 44106, USA

P. Pontarotti (ed.), *Evolutionary Biology: Exobiology and Evolutionary Mechanisms*,
DOI: 10.1007/978-3-642-38212-3_18, 

The Dobzhansky-Muller model makes the fundamental prediction that a minimum of two modified genes produce a deleterious epistatic interaction that is sufficient to generate hybrid incompatibility (Dobzhansky 1937; Muller 1940, 1942). Although species that diverged long back may have more than just two pairs of genes that are functionally incompatible in the hybrid, this model nonetheless indicates that speciation is a genetically tractable problem and opens the exciting possibility of identifying "speciation genes" and study this process at the molecular and cellular levels. Indeed, individual genes responsible for hybrid incompatibilities among related species have already been identified (Watanabe 1979; Hutter et al. 1990; Perez et al. 1993; Brideau et al. 2006).

Here we focus on hybrid infertility, a hybrid incompatibility that functions as a post-zygotic barrier among species. Hybrid infertility has been studied for several decades, in particular employing quantitative genetics methods. In parallel to these studies, impressive progress in the field of developmental biology has uncovered in great detail how germline cells develop and function in model organisms. Bringing these two fields together, evolution and development, may allow testing candidate speciation genes and uncover general mechanisms that lead to hybrid infertility.

We first introduce a group of sibling species of the melanogaster sub-group that diverged recently and can serve as a model to study hybrid infertility. Next, we briefly describe the development of germline cells in *D. melanogaster* females, and discuss how normal development may be affected in female hybrids and how these defects can be determined and tracked. Finally, we discuss classical work using interspecific transplantation experiments of germline cells and possible future avenues for the study of hybrid infertility in the genomic era.

## 18.2 Hybrid Incompatibility Among *Drosophila* Species

Hybrid incompatibility can be studied in species that underwent a recent speciation event and have not fully developed a prezygotic barrier, so that hybrids can still be generated and studied. *Drosophila* species are particularly suitable for this analysis since many related species can be investigated in the laboratory and in nature, where they form hybrid zones. The study of *Drosophila* hybrids begins with the fortuitous discovery of a *D. melanogaster* culture composed exclusively by sterile individuals of one sex in the beginning of the twentieth century. Upon close examination, Alfred Sturtevant and Calvin Bridges would find that these cultures were the result of the hybridization of *D. melanogaster* with another species, which Sturtevant named as *D. simulans*, due to its extreme resemblance to *D. melanogaster* (Sturtevant 1919). Soon, Sturtevant would uncover an asymmetric pattern of hybrid incompatibility that is X chromosome-specific. Crosses between *D. melanogaster* females and *D. simulans* males yield a male progeny that is lethal and adult females that are sterile. In the reverse cross using *D. simulans* as the maternal species, hybrid females normally die and the adult hybrid males are sterile (Sturtevant 1920). Because the hybrid progeny is sterile in either direction

of the cross, Sturtevant was unable to pursue his ultimate goal of mapping the genes involved in these incompatibilities. A historical account of Sturtevant's experiments and alternative methods that ensued to investigate the problem of speciation is reviewed in (Barbash 2010). Those alternative methods included a genetic trick of using triploid *D. melanogaster* flies to introgress chromosomal segments of *D. simulans* into *D. melanogaster* genome (Muller and Pontecorvo 1940), and turning to other species pairs with incomplete reproductive isolation, such as in *D. pseudoobscura* and *D. persimilis* that can yield fertile F1 female hybrids (Dobzhansky 1936). These and other experiments lent support that hybrid incompatibilities arise from interacting alleles rather than gross chromosomal incompatibilities.

Further advancement in the study of hybrid incompatibilities is owed to two important sets of discoveries. The first were the findings of strain variants of *D. simulans* and *D. melanogaster* from natural populations that rescue hybrid lethality or fertility in female F1 hybrids, which led to the molecular identification of key genes involved in hybrid incompatibilities, and provided tools for screening interacting genes [reviewed in (Barbash 2010; Watanabe 1979; Brideau et al. 2006; Presgraves and Stephan 2007; Barbash and Ashburner 2003)]. The second set of discoveries was the identification in the late 1970s–1990s of other species belonging to the *melanogaster* subgroup with much shorter divergence spans than 5.4 mya that separate *D. melanogaster* from *D. simulans*. Among those newly discovered species are *D. sechellia* and *D. mauritiana*, which together with *D. simulans* and *D. melanogaster*, form the *melanogaster* complex (Fig. 18.1) (Lachaise et al. 1986). *D. sechellia* is endemic to the Seychelles islands and feeds exclusively on the Morinda fruit, and *D. mauritiana* is endemic to the Mauritius island. Both species have an estimated divergence of only 0.5 mya from *D. simulans*. Due to their recent divergence, *D. simulans, D. sechellia* and *D. mauritiana* have an incomplete reproductive isolation and yield fertile hybrid females. Additional species of the *melanogaster* subgroup were also found. The *yakuba* complex is formed by *D. santomea*, from São Tomé e Príncipe island, *D. teisseire* and *D. yakuba*. Finally, the more divergent *erecta* complex is formed by *D. erecta* and *D. orena* (reviewed in (David et al. 2007)).

The discovery of these new species complexes greatly increase our understanding of phylogenetic relationships of *D. melanogaster* and sister species and facilitate the study of speciation since some of them allow backcrosses that can be used for mapping purposes. Furthermore, the genome sequencing of many of these species allowed comparisons of their genomes. Nevertheless, in spite of these advances, the genes involved in female sterility in hybrids remain largely unknown, with the exception of *Hmr* (Barbash and Ashburner 2003). Previous work indicates that genes responsible for male sterility evolve faster than genes responsible for either hybrid lethality or female sterility. In addition, male sterility involves a large number of genes (Wu and Palopoli 1994; Hollocher and Wu 1996; True et al. 1996; Sawamura et al. 2000). However, despite the fact that there are fewer genes involved in female hybrid sterility than in male sterility, the progress in identifying these genes has been slow.

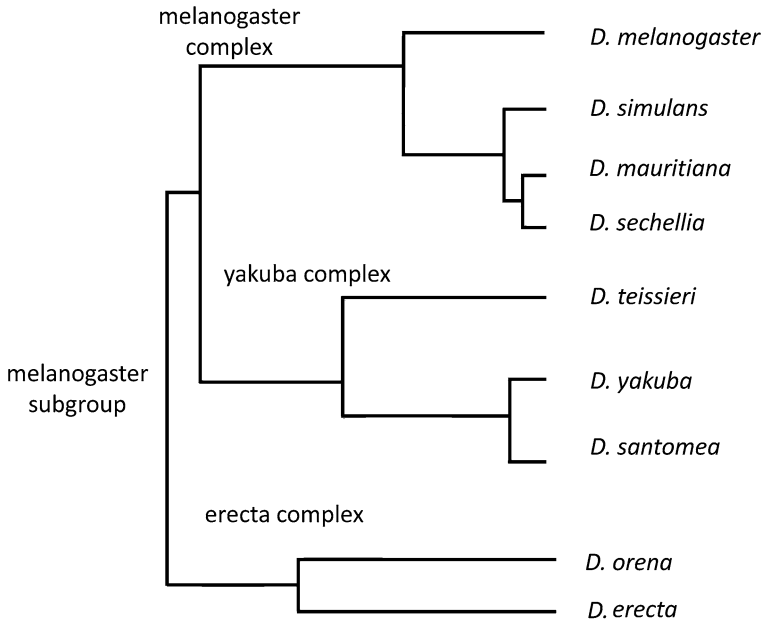



**Fig. 18.1** Phylogenetic tree of melanogaster subgroup species. Three complexes of the melanogaster subgroup are depicted: *melanogaster*, *yakuba,* and *erecta* complexes. In the melanogaster complex, an estimated 5.4 mya separate *D. melanogaster* from *D. simulans*, and 0.5 mya separate *D. simulans* from *D. sechellia*. This phylogenetic distance is reflected in the pattern of infertility that arises in hybrid females from these sibling species

## 18.3 Developmental Biology of Germline Cells

Hybrid females between *D. melanogaster* and any of its sister species, *D. simulans*, *D. sechellia* or *D. mauritiana*, have atrophied agametic ovaries (Sturtevant 1920; Santamaria 1977; Lachaise et al. 1986; Hollocher et al. 2000). Identifying which developmental process is impaired in hybrids that causes the agametic phenotype is crucial for understanding the underlying factors of reproductive isolating mechanisms and the developmental basis of speciation. In this section, we provide an overview of *Drosophila* germline development and highlight mechanisms that could explain the sterility phenotype in hybrid females.

Germline precursor cells are formed early in development and must escape somatic differentiation, while maintaining an "immortal" potential to generate viable fertile progeny over generations (Santos and Lehmann 2004). After the first syncytial nuclear divisions following fertilization, some nuclei migrate to the posterior pole of the embryo and are separated from those destined to become somatic cells. The posterior region of the embryo contains the germ plasm, a maternally deposited material rich in proteins and RNAs that induces the formation of pole cells, which later develop into primordial germline cells (PGCs)

(Fig. 18.2a) (Illmensee and Mahowald 1974; Underwood et al. 1980). Genetic screenings identified several genes involved in pole plasm assembly and function, including *vasa* (Rongo and Lehmann 1996; Williamson and Lehmann 1996). *vasa* encodes an RNA helicase and is a highly conserved germline-specific cell marker expressed from embryonic to adult stages in several species (Lasko and Ashburner 1988; Doren et al. 1998; Sano et al. 2002; Sheng et al. 2009). During gastrulation, the PGCs are tightly packed and carried inside the midgut pocket by the germ band extension (Fig. 18.2b). From there, individual PGCs split into two groups and undergo a transepithelial migration through the midgut wall toward the developing somatic gonads (Fig. 18.2c, d) (Jaglarz and Howard 1994). *wunen* and *wunen2* are key genes for this concerted migration, and their expression in both somatic and germ cells is suggested to underlie a competition for lipid phosphate required for germ cell migration and survival (Renault et al. 2004).

At the end of embryogenesis, PGCs re-enter mitosis and proliferate, which offsets the large loss of pole cells targeted to apoptosis at earlier stages or that differentiate into somatic cells (Yamada et al. 2008). PGCs then interact with the gonadal somatic tissue, maturing into germline stem cells (GSCs). The PGC to GSC transition coincides with the establishment of the GSC somatic niche, a microenvironment that ensures survival and maintenance of stem cell identity during asymmetric divisions (Sheng et al. 2009). The male niche is located at the testis tip (the hub), while the female niche is located at the tip of the ovarian

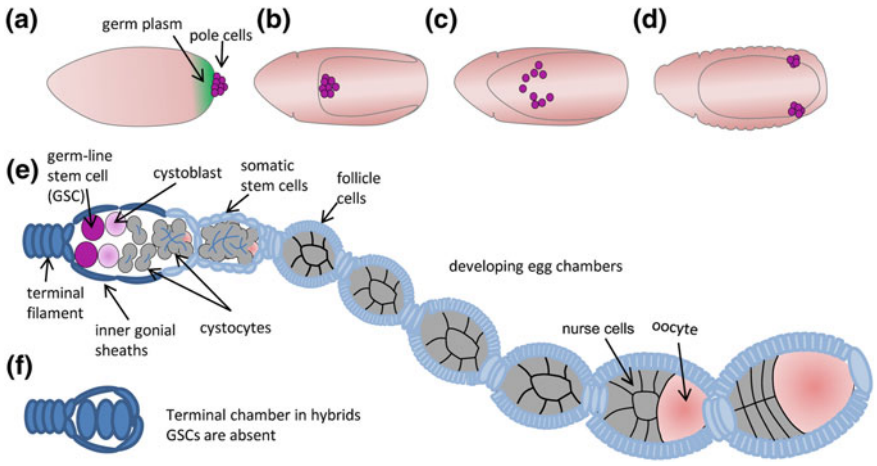



**Fig. 18.2** Germline development in wildtype and infertile hybrid females. **a–d** Wildtype embryonic stages. **a** Specification of pole cells (*purple*) by germ plasm (*green*) in the posterior pole of a preblastoderm embryo. **b** Pole cell location in the midgut after germ band elongation. **c**, **d** Migration from the midgut to the prospective gonads. Wildtype and *melanogaster/simulans* hybrid germline development are similar up to this point. **e** Ovariole from an adult wildtype and **f** hybrid ovary. Note the absence of GSCs in the hybrid germarium and lack of developing egg chambers

germarium and is composed of terminal filaments and cap cells (Fig. 18.2e) (Gilboa and Lehmann 2004).

The adult *D. melanogaster* female has two ovaries with about 18 ovarioles each. The number of ovarioles defines the rate of egg production and is a species-specific trait among *Drosophila* species. In particular, *D. sechellia* has a drastic reduction in ovariole numbers with the lowest rate of egg production compared to its sister species (Lachaise et al. 1986). Each ovariole consists of one germarium housing two to three GSCs, and a string of developing egg chambers (Fig. 18.2e). Upon division, one GSC remains attached to the niche, while the other GSC daughter cell differentiates into a cystoblast. The cystoblast undergoes a series of divisions with incomplete cytokinesis, forming interconnected cysts that are later enveloped by somatic follicle cells. A key signaling factor released by the niche is Dpp (Decapentaplegic; BMP-4 homolog), which regulates self-renewal and GSC differentiation (Xie and Spradling 1998; Casanueva and Ferguson 2004). Dpp signaling is stronger in the proliferative pool of GSC, but weaker in the differentiating pool that expresses *bag-of-marbles* (*bam*), a gene that regulates GSC differentiation (McKearin and Ohlstein 1995; McKearin and Spradling 1990). In the egg chamber, only one out of the 16 cells in the cyst becomes the oocyte, while the others become nurse cells. Nurse and follicle cells assist the egg development, contributing with maternally deposited RNAs and proteins that provide positional information of the main axes of the embryo (Williamson and Lehmann 1996).

### *18.3.1 Germline Development in Sterile Hybrid Females*

From the description above, there are several crucial steps in which the development of germline cells could go awry in infertile hybrid females. The rich arsenal of markers and transgenic constructs developed in *D. melanogaster* can be transferred to hybrid embryos in order to investigate the critical developmental stage when the phenotype of sterility first appears. For instance, Vasa can be used as a marker to visualize pole cell formation in hybrid embryos and track PGCs during migration in vivo. Hybrid embryos generated by crosses between either *D. simulans* or *D. sechellia* males and *D. melanogaster* females have similar numbers of *vasa*$^{+}$ pole cells in comparison to pure species (Mizutani et al. unpublished data). Therefore, the initial germline development is normal in hybrids, indicating that the germ plasm from a single species can sustain the normal formation of pole cells with hybrid genome content. Live-imaging of PGCs tagged with a fusion protein Vasa-GFP (Sano et al. 2002, 2005) also shows a correct migration pattern in hybrid embryos, resulting in similar numbers of PGCs reaching the developing gonads (Mizutani et al. unpublished data). These findings rule out several genes involved in pole cell formation and migration as candidates for interspecific sterility genes.

However, between larval stages L1 and L2, *vasa* expression is lost in hybrids (Mizutani et al. unpublished data). It is possible that the PGCs fail to transition into GSCs in hybrids, or that the pool of germline cells cannot self-renew and is lost

due to uncontrolled divisions and somatic differentiation. Another possible explanation for loss of *vasa* expression would be if germline cells are targeted for apoptosis at a higher rate than normal, due to a hybrid genomic instability. In either case, the loss of *vasa* expression in the hybrid larvae explains the agametic ovary phenotype, which only display a rudimentary germarium containing terminal filaments and terminal chambers that do not develop further and lack expression of *vasa* and other markers (Fig. 18.2f) (Hollocher et al. 2000; Mizutani et al. unpublished observation). The loss of cross-talk between soma and germline cells in the forming niche may also explain the atrophied ovary in hybrids.

As mentioned earlier, using a combination of certain strains from natural populations, *Hmr* mutation, and low rearing temperature, it has been possible to rescue the agametic phenotype of *melanogaster/simulans* hybrid females. These rescued hybrids sustain a complete development of germline cells that produce progeny, but a large percentage of embryos display additional phenotypes of egg chorion malformation, defective embryonic polarity, and aberrant nuclei divisions (Hollocher et al. 2000; Barbash and Ashburner 2003), indicating that additional genes are responsible for fully rescuing normal egg development. For instance, proper chromosomal segregation during meiosis and suppression of transposons are essential to germline development to maintain the fidelity and accuracy of genetic transmission. In the case of transposon silencing, Piwi-interacting RNAs (piRNAs) play a key role in protecting genetic information in the germline [reviewed in (Khurana and Theurkauf 2010)], and this pathway was recently shown to be disrupted in hybrids between D. melanogaster and D. simulans (Kelleher et al. 2012). Transposon silencing by piRNAs can occur post-transcriptionally with cleavage of mature transposon transcripts (Saito et al. 2006; Gunawardane et al. 2007; Nishida et al. 2007), or transcriptionally with Piwi proteins binding to the nuclear heterochromatin protein HP1a and mediating transcriptional silencing through heterochromatin assembly (Pal-Bhadra et al. 2004; Brower-Toland et al. 2007). It will be interesting to test if genes involved in meiotic segregation and transposon silencing are divergent in *D. melanogaster* in comparison to its sibling species, and if the replacement of these genes by sequences from a single species would increase the yield of normal embryos in hybrid females with rescued fertility.

## 18.4 Transplantation Experiments

Transplantation experiments of germline cells provide further insights into the role of soma-germline interactions in hybrid incompatibility. When genetically marked pole cells from *D. melanogaster* are transplanted into *melanogaster/simulans* or *melanogaster/mauritiana* hybrid eggs, which normally produce sterile females, it was possible to retrieve fertile hybrid females (Santamaria 1977). This finding shows that the germline from pure species can develop normally in a hybrid somatic environment. Thus, it is likely that the terminal filaments which are still

formed in hybrid females actually represent a functional GSC niche, since it can effectively mature germline cells from a pure species in these transplantation experiments. In contrast, hybrid pole cells transplanted into *D. melanogaster* host are unable to develop and produce progeny. This result suggests that sterility arises from a cell autonomous incompatibility, not through a germline/soma incompatibility.

The use of interspecific transplantation of pole cells also proved as an useful tool to create hybrids between more distantly related species in the *melanogaster* subgroup that cannot normally be produced by matings, and to investigate interspecific germline development. Sanchez and Santamaria transplanted pole cells from each species of the *melanogaster* subgroup (Fig. 18.1, except for *D. santomea*) into *D. melanogaster* host embryos that lack pole cells due to the *oskar* mutation (Sanchez and Santamaria 1997). These authors found that the somatic tissue of *D. melanogaster* is able to nurture transplanted pole cells into functional oocytes from all species, except the most divergent species *D. erecta* and *D. orena*. They conclude that this failure can be explained by the more distant separation time between *D. melanogaster* and species of the *erecta* complex, which is expected to have modified gene complexes responsible for normal soma-germline interactions.

### *18.4.1 Ancestral Alleles of D. simulans and D. sechellia Should Include Female Infertility Genes in Hybrids with D. melanogaster*

The work described above provides invaluable information about the biology of germline cells and indispensable data regarding the mechanisms by which cross-species incompatibilities may arise. However, the identity of the incompatible pairs of genes still remains largely unknown. One way of addressing this issue is to use the hybrid phenotypes between *D. melanogaster*, *D. simulans,* and *D. sechellia* and test whether there are specific patterns of gene divergence consistent with these phenotypes (Fig. 18.3a). In this regard, since hybrid females between *D. melanogaster* and either *D. simulans* and *D. sechellia* are invariably sterile and hybrid females between *D. simulans* and *D.sechellia* are fertile, it is likely that the genes in question should be those that diverge the least in *D. simulans* and *D. sechellia* but diverge the most in *D. melanogaster*. We previously identified this class of genes as ancestral alleles of the lineage of *D. simulans* and *D. sechellia* (Fig. 18.3b) (Sousa-Neves and Rosas 2010). Among ancestral alleles, a small subset is exclusively expressed in the germline from embryonic to adult stages. These candidate genes are predicted to encode proteins with conserved domains present in other key proteins that regulate germline specification and development. Among the ancestral alleles is the aforementioned gene *bag-of-marbles*, which displays divergent sequences between *D. melanogaster* and *D. simulans* and is

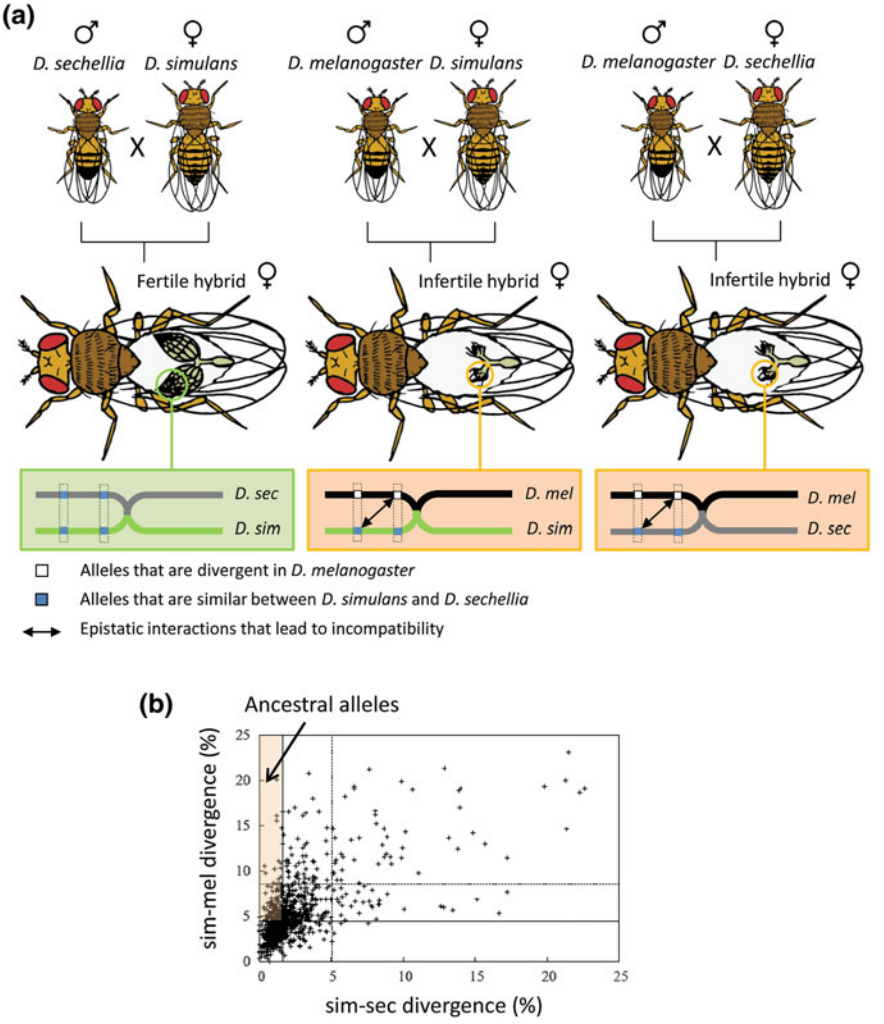



**Fig. 18.3** Female hybrid progeny from interspecific crosses among *D. melanogaster* sibling species and identification of genes involved in hybrid sterility. **a** Hybrid females between the more closely related species *D. simulans* and *D. sechellia* are fertile with fully developed ovaries, while hybrids between these species and *D. melanogaster* are infertile with atrophied ovaries. Boxes display a hypothetical hybrid chromosome with either conserved alleles that are compatible (*simulans/sechellia* hybrid), or diverged alleles with deleterious epistatic interactions causing sterility (*arrows*). **b** Scatter-plot displaying pair-wise comparison of X- chromosome homologous genes between *D. simulans* and *D. sechellia*, and *D. simulans* and *D. melanogaster*. *Highlighted box* shows ancestral alleles on this chromosome that are similar between *simulans* and *sechellia* but divergent from *melanogaster*, expected to include genes responsible for female hybrid infertility [Adapted from Sousa-Neves and Rosas (2010)]

suggested to have evolved through positive selection (Civetta et al. 2006). Preliminary data suggests that removal of these candidate genes in *D. melanogaster* affect germline development and produce progenies of skewed proportions of fewer males than females (Mizutani et al. unpublished data). We hypothesize that the distortions in sex ratio could be due to abnormal meiosis or de-repression of transposition, since both cases should lead to a deficit in the heterogametic sex. However, in order to show that these genes are indeed sterility genes responsible for the infertility barrier among *D. melanogaster* sibling species, it would be necessary to remove their epistatic interactions and rescue fertility in hybrid females.

## 18.5 Conclusion

The use of *Drosophila* hybrids to investigate genetic incompatibilities and speciation was pioneered by Sturtevant and continues to this day (excellently reviewed by (Barbash 2010)). *D. melanogaster*, *D. sechellia,* and *D. simulans* are of particular and renewed interest in this quest due to the growing availability of genetic markers, genome sequences, and studies about their biology, but also due to the fact that crosses between them yield either fertile or infertile female progeny. These advantages allow the design of experiments that combine genomic, genetic, and reverse genetic approaches to identify sterility genes and mechanisms that cannot be assessed in single species. By identifying the precise stage in development when hybrid germline cells become aberrant and how this proceeds at the cellular and molecular level, it is likely that we will have a better understanding of mechanisms that generate sterility among species and their role in evolution.

## References

Barbash DA (2010) Ninety years of Drosophila melanogaster hybrids. Genetics 186:1–8

Barbash DA, Ashburner M (2003) A novel system of fertility rescue in Drosophila hybrids reveals a link between hybrid lethality and female sterility. Genetics 163:217–226

Brideau NJ, Flores HA, Wang J, Maheshwari S, Wang X, Barbash DA (2006) Two Dobzhansky-Muller genes interact to cause hybrid lethality in Drosophila. Science 314:1292–1295 New York, N.Y

Brower-Toland B, Findley SD, Jiang L, Liu L, Yin H, Dus M, Zhou P, Elgin SCR, Lin H (2007) Drosophila PIWI associates with chromatin and interacts directly with HP1a. Genes Dev 21:2300–2311

Casanueva MO, Ferguson EL (2004) Germline stem cell number in the Drosophila ovary is regulated by redundant mechanisms that control Dpp signaling. Development 131:1881–1890 Cambridge, England

Civetta A, Rajakumar SA, Brouwers B, Bacik JP (2006) Rapid evolution and gene-specific patterns of selection for three genes of spermatogenesis in Drosophila. Mol Biol Evol 23:655–662

David JR, Lemeunier F, Tsacas L, Yassin A (2007) The historical discovery of the nine species in the Drosophila melanogaster species subgroup. Genetics 177:1969–1973
Dobzhansky T (1937) Genetics and the origins of Species. Columbia University Press, New York
Dobzhansky TH (1936) Studies on hybrid sterility. II. Localization of sterility factors in Drosophila pseudoobscura hybrids. Genetics 21:113–135
Doren M, Williamson AL, Lehmann R (1998) Regulation of zygotic gene expression in Drosophila primordial germ cells. Curr Biol 8:243–246
Gilboa L, Lehmann R (2004) Repression of primordial germ cell differentiation parallels germ line stem cell maintenance. Curr Biol 14:981–986 Elsevier
Gunawardane LS, Saito K, Nishida KM, Miyoshi K, Kawamura Y, Nagami T, Siomi H, Siomi MC (2007) A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in Drosophila. Science 315:1587–1590
Hollocher H, Agopian K, Waterbury J, O'Neill RW, Davis AW (2000) Characterization of defects in adult germline development and oogenesis of sterile and rescued female hybrids in crosses between Drosophila simulans and Drosophila melanogaster. J Exp Zool 288:205–218
Hollocher H, Wu CI (1996) The genetics of reproductive isolation in the Drosophila simulans clade: X vs. autosomal effects and male vs. female effects. Genetics 143:1243–1255
Hutter P, Roote J, Ashburner M (1990) A genetic basis for the inviability of hybrids between sibling species of Drosophila. Genetics 124:909–920
Illmensee K, Mahowald AP (1974) Transplantation of posterior polar plasm in Drosophila. Induction of germ cells at the anterior pole of the egg. Proc Natl Acad Sci USA 71:1016–1020
Jaglarz M, Howard K (1994) Primordial germ cell migration in Drosophila melanogaster is controlled by somatic tissue. Development 120:83–89
Khurana JS, Theurkauf W (2010) piRNAs, transposon silencing, and Drosophila germline development. J Cell Biol 191:905–913
Kelleher ES, Edelman NB, Barbash DA (2012) Drosophila interspecific hybrids phenocopy piRNA-pathway mutants. PLoS Biol 10(11):e1001428. doi: 10.1371/journal.pbio.1001428
Lachaise D, David JR, Lemeunier F, Tsacas L, Ashburner M (1986) The reproductive relationships of Drosophila sechellia with D. mauritiana, D. simulans, and D. melanogaster from the Afrotropical region. Evolution 40:262–271 Society for the Study of Evolution
Lasko PF, Ashburner M (1988) The product of the Drosophila gene vasa is very similar to eukaryotic initiation factor-4A. Nature 335:611–617
McKearin DM, Spradling AC (1990) Bag-of-marbles: a Drosophila gene required to initiate both male and female gametogenesis. Genes Dev 4:2242–2251
McKearin DM, Ohlstein B (1995) A role for the Drosophila bag-of-marbles protein in the differentiation of cystoblasts from germline stem cells. Development 121:2837–2847
Muller H (1940) Bearing of the Drosophila work on systematics. In: Huxley J (ed) The new systematics. Clarendon, Oxford, pp 185–268
Muller H (1942) Isolating mechanisms, evolution, and temperature. Biol Symp 6:71–125
Muller HJ, Pontecorvo G (1940) Recombinants between Drosophila species the F1 hybrids of which are sterile. Nature 146:199–200
Nishida KM, Saito K, Mori T, Kawamura Y, Nagami-Okada T, Inagaki S, Siomi H, Siomi MC (2007) Gene silencing mechanisms mediated by Aubergine piRNA complexes in Drosophila male gonad. RNA 13:1911–1922
Pal-Bhadra M, Leibovitch BA, Gandhi SG, Rao M, Bhadra U, Birchler JA, Elgin SCR (2004) Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. Science 303:669–672
Perez DE, Wu NA, Johnson CI, Wu M (1993) Genetics of reproductive isolation in the Drosophila simulans clade: DNA marker-assisted mapping and characterization of a male-hybrid sterility gene, Odysseus (Ods)
Presgraves DC, Stephan W (2007) Pervasive adaptive evolution among interactors of the Drosophila hybrid inviability gene, Nup96. Mol Biol Evol 24:306–314
Renault AD, Sigal YJ, Morris AJ, Lehmann R (2004) Soma-germ line competition for lipid phosphate uptake regulates germ cell migration and survival. Science 305:1963

Rongo C, Lehmann R (1996) Regulated synthesis, transport and assembly of the Drosophila germ plasm. Trends Genet 12:102–109

Saito K, Nishida KM, Mori T, Kawamura Y, Miyoshi K, Nagami T, Siomi H, Siomi MC (2006) Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. Genes Dev 20:2214–2222

Sanchez L, Santamaria P (1997) Reproductive isolation and morphogenetic evolution in Drosophila analyzed by breakage of ethological barriers. Genetics 147:231–242

Sano H, Nakamura A, Kobayashi S (2002) Identification of a transcriptional regulatory region for germline-specific expression of vasa gene in Drosophila melanogaster. Mech Dev 112:129–139

Sano H, Renault AD, Lehmann R (2005) Control of lateral migration and germ cell elimination by the Drosophila melanogaster lipid phosphate phosphatases Wunen and Wunen 2. J Cell Biol 171:675–683

Santamaria P (1977) On the causes of sterility in some interspecific hybrids from the Melanogaster subgroup of Drosophila. Dev Genes Evol 310:305–310

Santos AC, Lehmann R (2004) Germ cell specification and migration in Drosophila and beyond. Curr Biol 14:R578–R589

Sawamura K, Davis AW, Wu CI (2000) Genetic analysis of speciation by means of introgression into Drosophila melanogaster. Proc Natl Acad Sci USA 97:2652–2655

Sheng XR, Posenau T, Gumulak-Smith JJ, Matunis E, Van Doren M, Wawersik M (2009) Jak-STAT regulation of male germline stem cell establishment during Drosophila embryogenesis. Dev Bio 334:335–344

Sousa-Neves R, Rosas A (2010) An analysis of genetic changes during the divergence of Drosophila species. PLoS ONE 5:e10485

Sturtevant AH (1919) A new species closely resembling Drosophila melanogaster. Psyche 26:153–156

Sturtevant AH (1920) Genetic studies on Drosophila simulans. I. Introduction. Hybrids with Drosophila melanogaster. Genetics 5:488–500

True JR, Weir BS, Laurie CC (1996) A genome-wide survey of hybrid incompatibility factors by the introgression of marked segments of Drosophila mauritiana chromosomes into Drosophila simulans. Genetics 142:819–837

Underwood E, Caulton J, Allis C, Mahowald A (1980) Developmental fate of pole cells in Drosophila melanogaster. Dev Biol 77:303–314

Watanabe TK (1979) A gene that rescues the lethal hybrids between Drosophila melanogaster and Drosophila simulans. Jpn J, Genet 54

Williamson A, Lehmann R (1996) Germ cell development in Drosophila. Annu Rev Cell Dev Bio 12:365–391

Wu CI, Palopoli MF (1994) Genetics of postmating reproductive isolation in animals. Annu Rev Genet 28:283–308

Xie T, Spradling A (1998) Decapentaplegic is essential for the maintenance and division of germline stem cells in the Drosophila ovary. Cell 94:251–260

Yamada Y, Davis KD, Coffman CR (2008) Programmed cell death of primordial germ cells in Drosophila is regulated by p53 and the outsiders monocarboxylate transporter. Development 135:207–216

# Chapter 19
# Genetic and Molecular Dissection of Animal Decision-Making: A New Frontier for Genetic Analysis

**Youngmin Chu, Joseph Schinaman and Rui Sousa-Neves**

**Abstract** In 1859, Darwin and Wallace presented extensive evidence that evolution occurs through natural selection. Twelve years later he would provide evidence that females select males and that this sexual selection changes the evolutionary paths of populations. Since then, biologists have shown that female mate choice is widespread in the animal kingdom, which strongly suggests that this may be an ancient behavior that ultimately drives the evolution of novel nervous systems and behaviors. More recently, female mate choice began to emerge as a central paradigm to probe and understand animal decision-making. The development of new genetic systems and the sequencing of genomes allow us to identify genes and mechanisms responsible for animals to make choices. Here we discuss the genetics basis of behavior and decision-making, the use of genetic models to study female mate choice, and how different disciplines and strategies may help us better understand animal decision-making.

## 19.1 The Genetic Basis of Behavior

Most of us are well aware that related individuals of the same species often share similar gestures, facial expressions, gazes, walking gaits, and temperaments. Abnormal behaviors such as schizophrenia and aggression, among others, also seem to appear more frequently in related individuals than in individuals of different pedigrees. Furthermore, some behaviors such as facial expressions appear

Y. Chu · J. Schinaman
Department of Biology, Case Western Reserve University, Cleveland, OH 44106, USA

R. Sousa-Neves (✉)
Department of Genetics and Genome Sciences, School of Medicine, Case Western Reserve University, Cleveland, OH 44106, USA
e-mail: rui.sousaneves@case.edu

P. Pontarotti (ed.), *Evolutionary Biology: Exobiology and Evolutionary Mechanisms*,
DOI: 10.1007/978-3-642-38212-3_19, 

more similar in species that are more closely related (e.g., humans and chimpanzees) than those that are more distantly related (e.g., chimpanzees and dogs).

Together, these simple observations suggest that there is a strong relationship between genetics and behavior. However, many studies that aim to determine how individual genes contribute to these traits are often confronted with results that suggest a complex inheritance, possibly highly influenced by external factors. The reason for such conclusions usually stems from the difficulty of linking a given allele to a specific behavior. In other words, it is often difficult to predict whether the presence or absence of a given allele influences the way an organism will behave. To understand this difficulty, perhaps it would be helpful to briefly mention how genes operate.

Genetics and developmental genetics have shown that proteins encoded by genes act sequentially to modify the way other proteins perform a given task. These studies have firmly established that the loss of genes that participate in the same biological pathway lead to the same abnormal phenotype or its exact opposite. Thus, if behavior has a strong genetic component, it would not be surprising if a given behavior could be modified by different alleles of the same gene as well as by alleles of different genes that act in the same pathway.

Although it is generally assumed that behavior is a complex trait encoded by polygenes, the strongest evidence in favor of polygenic inheritance usually relies on the lack of correlation between a behavioral phenotype and a simple genotype (i.e., a mutation in a single gene). However, the lack of evidence for correlation does not serve as evidence that there is no correlation especially if we consider that mutations in two or more simple genotypes (i.e., single genes) can lead to the same phenotype.

### *19.1.1 Simple Trait Versus Complex Trait*

Simple traits are those in which there is clear and strong relationship between genotype and phenotype. The equivalence between genotype and phenotype allows us to follow traits by Mendelian proportions and to establish linkage groups. Ultimately, the determination of a linkage group allows us to pinpoint which chromosome carries the trait and to determine genetic distances from this given trait to any existing markers on its chromosome. In contrast, inheritance of complex traits usually often implies that minute allelic variations in a series of genes lead to a quantitative variation of phenotypes. Due to the quantitative nature of some complex inheritance, and also the fact that minute allelic variations are usually assumed to be spread along chromosomes, it is not possible to establish linkage groups or locate the position of individual genes by standard genetic mapping.

Although the variation of complex traits is usually assumed to be continuous, it can also be discrete. Continuous variations imply that the contribution of several different *loci* produce a continuous range of phenotypes (e.g., height). In contrast, discrete variations of complex traits imply that several genes must act in concert to achieve a final phenotype that tolerates little variation. In this case, one might

expect that the disruption of any of the components of such genetic system lead to the collapse of the phenotype.

In summary, it is important to keep in mind these issues before designing and interpreting experiments that aim to determine how genetics influences behavior.

### *19.1.2 Genetics and the Genetic Basis of Animal Decision-Making*

In the same way that genetics is an integral part of behaviors we can extend this view to a particular behavior such as decision-making. Here again, the evidence is based on the inheritability of the ability to make similar decisions in different animal groups. Although little is known about the mechanisms that allow animals to make decisions, it is clear that most animals evaluate events happening in time and space and are capable to come to a decision. Decision-making is by no means a privilege of humans. For instance, crustaceans such as crayfish can establish hierarchical relationships, which indicate that individuals can critically evaluate the risk of challenging an opponent that is larger and stronger (Bovbjerg 1953). Similar abilities to make decisions can be observed in worms, fish, mammals, and birds (Esch et al. 2002; Fischhoff et al. 2007; Siebeck et al. 2009).

Due to the fact that humans make more complex decisions than most animals, we often attribute to decision-making an evolutionary hierarchical order that places humans on the top branch, just like naturalists did in the past while classifying animals and early philosophers did while describing the universe. However, before accepting that we are in the top branch of decision-making, we perhaps should consider the possibility that decision-making might rely on circuits with same basic architecture that so happen to be connected to circuits with a higher or lower storage and processing capacity. This view is in part supported by studies that implicate the mammalian basal ganglia in decision-making and the fact that this structure has a homologous structure in birds, which diverged from mammals approximately 310 million years ago (Benton 1993, 2004).

The fact that behaviors are physiological responses at first seems to indicate that they cannot be understood in terms of their genetic components, or that these components are too distant from the physiological response to be relevant. In the early years of developmental genetics, development also seemed to be the result of cell interactions believed to be too complex to be subject to genetic analyses. However, systematic genetic analyses demonstrated that cell decisions can be broken down into simple decisions which ultimately require genes (Nusslein-Volhard and Wieschaus 1980).

If we accept for the time being that decision-making can be dissected genetically, then what should we expect to find, continuous variation, or discrete variation? *A priori*, the fact that most decisions rely on finite mutually exclusive possibilities or scenarios (e.g., attack versus remain passive) suggest that those

might be governed by discrete variables (i.e., there is no state between attacking and remaining passive).

The points highlighted above suggest that there are significant reasonable doubts to believe that decision-making might be divorced from genetics. If that is so, what do we need to do to test whether or not animal decision-making can be genetically dissected in its discrete components? Before addressing this question, we will briefly discuss about female mate choice, a behavior that requires decision-making.

## 19.2 Female Evaluation of Courtship is an Ancient Process of Animal Choice and an Ideal Platform to Decode Mechanisms of Decision-Making

By looking at the diversity in nature, we cannot help but to ask: "How are animals with different forms and behaviors generated?" Wallace and Darwin presented extensive evidence that this variation was due to a process of natural selection and the survival of the fittest. In other words, somehow, every generation of animals are endowed with variations that can be selected to adapt them to particular environments. The fittest, in his view, are those most adapted to a given environment, and as a consequence, leave progeny and thrive.

The development of the concept of natural selection was heavily influenced by animal breeders who had successfully generated new forms of animals from existing variants in nature. While natural selection and the survival of the fittest *per se* could explain the diversification of life forms, it seems that the process would be more effective and directional if animals were actively selecting particular genes from existing varieties like human breeders do. Soon, Darwin would find evidence that animals do select traits (i.e., genes) in a most unexpected way: by sexual selection.

Before Darwin, the chirps and colors of birds were generally believed to be the expression of an idyllic and peaceful world. However, Darwin noted that these vocalizations were frequently used to call females or scare away opponents. He noticed that the bright colors, songs, and dances that males often display serve for the purpose of attracting females. If males are performers and females are judges of these performances, in polygynous species it is likely that the best performer will leave the largest offspring. Furthermore, if males with different attributes are equally capable of enticing females, it is likely that over time these different attributes are combined in a single individual. This could easily explain why male courtships often include a variety of displays.

The view above illustrates that sexual selection is a mechanism by which a nervous system evolves to avoid random mating encounters and gains control over the evolutionary process. Not surprisingly, this becomes a recipe for a behavioral diversity that has the potential to set up individuals of a single population in different evolutionary paths. Since this control tends to enhance the ability of

females to discriminate their mates, it also prevents matings with species that are closely related which result in sterile progenies and inviability of one sex (Sturtevant 1920). Based on the number of existing species that exhibit sexual selection, it seems reasonable to conclude that the ability to make mate choices was very successful. Together, these facts suggest that female mating choice might be an ancient process of decision-making. Furthermore, the fact that this type of decision is binary (i.e., results in either acceptance or rejection of male courtship) makes it an attractive model to determine how animals make decisions.

### *19.2.1 Drosophila Melanogaster and Sibling Species as Models of Decision-Making*

The courtship of Drosophilids involves a series of complex steps performed by males, which are shown in Fig. 19.1a. A courting male initially orient toward a female and begins to track her (Spieth 1974). The initial female response is to stroll away from the male. The male then taps the female abdomen and encircles her few times, making displays with the wing and vibrating it to generate a song. During encircling, the male alternates the vibrating wing that produces the song to ensure that it is always closest to the female's organs that detect sound and odors, the arista, and antenna. It is not unusual to observe females rejecting male courtship by curling or raising the abdomen, or by kicking the male and running away (Connolly and Cook 1973). However, under normal conditions, the repetition of the courtship steps removes this rejection. Females then reduce their locomotor activity and open their vaginal plates, which are necessary steps for mounting and copulation. The entire process can take from few seconds up to 30 min (Fig. 19.1b). Most females readily accept a male, and those that fail to mate within one hour are very unlikely to ever mate. Thus, acceptance and rejection are mutually exclusive and generates clear outcomes. Experimental quantification of this behavior is simple, as it can be monitored by video recordings and scoring progenies. It is also possible to distinguish whether the females are attractive or unattractive to males by quantifying the intensity of male courtship (Fig. 19.1c).

Using acceptance and rejection as readout of female decision and systematically ablating olfaction, hearing and audition, Bastock, Ewing and others noted that the removal of individual senses increases the time required for females to accept courting males (Bastock 1956; Ewing 1964). These findings led them to propose that female acceptance is the result of sensory summation. Since then, a great effort has been made to identify genes and mechanisms that allow females to decode male courtship. Due to historical reasons, the first step of this quest was to ask whether genes that determine maleness and femaleness modify the ability of females to select males. The answer was that sex determination genes can indeed modify female mate choice. Another interesting findings of this quest was the

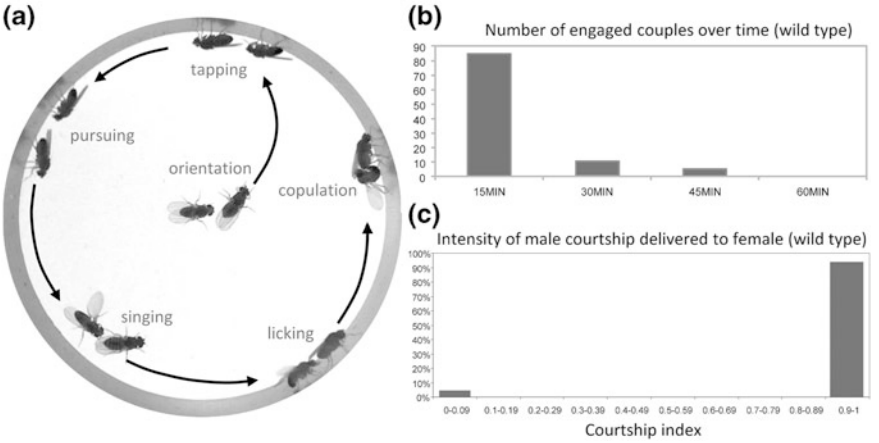



**Fig. 19.1** Courtship steps and quantitative responses **a** The *arena* shows discrete behaviors during the courtship of *Drosophila melanogaster*. **b** The *graph* shows the quantification of the stereotyped acceptance profile of wild type females courted by wild type males over time (*Source of the data*: Chu and Sousa-Neves 2013). Note that within 30 min all couples copulate. The presence of copulating couples in 45 min is due to the fact that couples stay engaged for approximately 20 min. **c** The *graph* quantifies the amount of courtship delivered to females (*Source of data*: Schinaman and Sousa-Neves unpublished). In the X-axis are indicated the highest Courtship Index (CI) which is the amount of time that a male courts a female within 10 min or the amount of time it courts until copulation occurs. CIs can also serve as measurements of female sex appeal

identification of *fruitless*, a gene that renders females behaving like males and unable to interpret male courtship (Gill 1963; Hall 1978). Studies on *fruitless* demonstrated that male and female circuits may be as different as their external appearance, but that is far from being the whole story.

In addition to the role of sex determination genes in establishing differences between genders and consequently the decisions that they make, there is evidence that mutations in other genes either impair or enhance the ability of females to accept males. One of such examples is the Sex Peptide (SP) and its correlate receptor the Sex Peptide Receptor (SPR). SP is produced by males and transferred during insemination where it binds to SPR and causes females to be unwilling to mate again (Chapman et al. 2003; Yapici et al. 2008). Other genes still have incompletely understood functions in female decision-making and do not necessarily fit in the sex dichotomizing view of "male specific genes" versus "female specific genes". Among those are *spinster*, *dissatisfaction*, *chaste*, and others (Finley et al. 1997; Keiko et al. 1997; Juni and Yamamoto 2009). These genes probably represent different pieces of a larger picture of female decision-making. Some of these genes are expressed in both sexes and mutations in them do not render females behaving like males. However, they do play a role in decision-making in females. One recently isolated gene implicated in female behavior is *datilógrafo* (*dati*), a gene we located on the *Drosophila* fourth chromosome, which

encodes a zinc finger transcription factor broadly expressed in the brain. *dati* mutant females do not exhibit male behaviors but reject males. The rejection of courting males by *dati*$^{1}$ mutant females is very strong and involves powerful kicks with hind legs, running away from males, and raising, and curling the abdomen to avoid intromission.

### 19.2.2 *Analysis of Somatic Mosaics Mutant for* dati$^{1}$ *to Determine Brain Centers Required for Female Decision-Making*

One way of addressing the role of *dati* in decision-making is to remove its product in the brain of females that are otherwise wild type (Fig. 19.2), This method is commonly used for genes located in most chromosomal regions of the fly, but a new tool, named as FYT system had to be developed for the fourth chromosome (Sousa-Neves and Schinaman 2012). By using this system, it is possible to systematically generate a collection of randomly distributed mosaics mutant for *dati*$^{1}$ in the brain and ventral nerve cord and test whether they lead to acceptance or rejection of male courtship. For instance, the number of times that a given region is hit by a *dati*$^{1}$ clone can be recorded and referenced to a behavioral outcome of acceptance or rejection (Fig. 19.2). The expectation is that by compiling the position of clones in the brain, it should be possible to pinpoint with great precision the regions in the brain that are unambiguously required for female decision-making.

### 19.2.3 *Do Circuits Required for Decision-Making Rely on Stimulation, Inhibition or Stimulation Plus Inhibition?*

Bastock and Manning performed a series of ablation experiments to quantify the contribution of each sensory modality to mating success. While the absence of light (i.e., no visual information) had no discernible effect on courtship acceptance in *D. melanogaster*, it did cause a dramatic effect in *D. simulans* (Manning 1960). Clipping the courting males' wings (i.e., disrupting the courtship song and wing displays), or ablating the females' antennae (i.e., removing olfaction and audition) strongly reduced mating in both species. However, even when all treatments were performed, matings still occurred, albeit at a low rate. Therefore, sound and smell contribute to courtship acceptance, but neither is absolutely required. Despite the fact that taste was not tested, these observations suggest that decisions can progress through separate sensory paths and that adding different sensory modalities improves the efficiency of the process. In other words, different stimulations

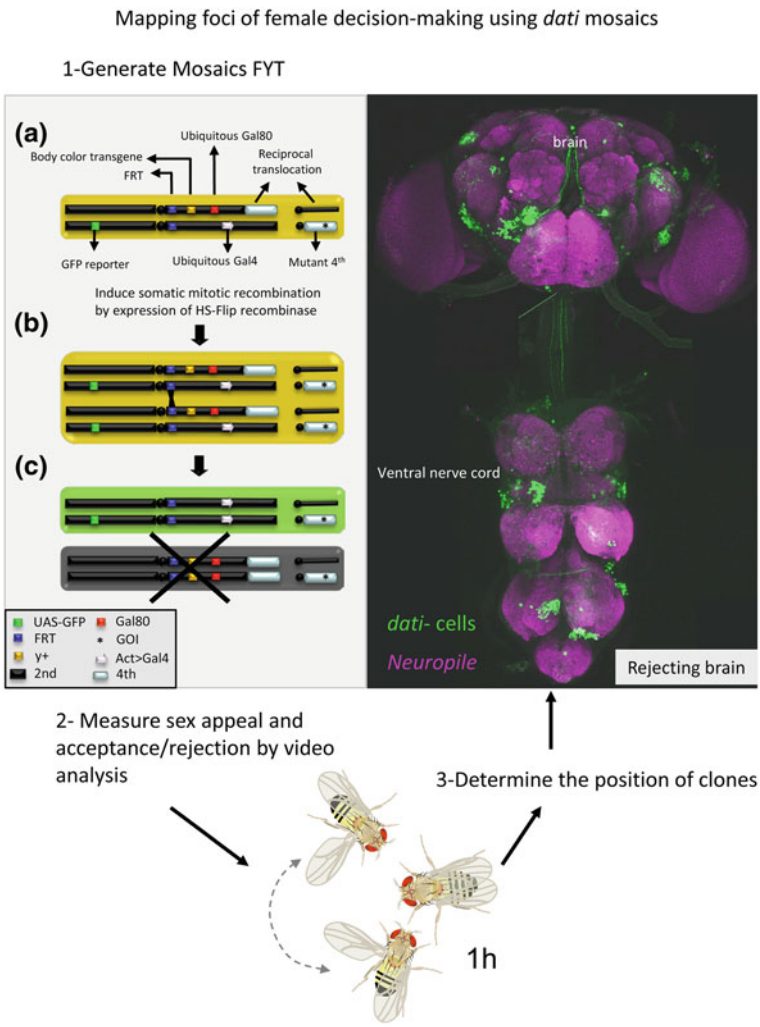



**Fig. 19.2** Mapping brain foci where *dati* is required. This analysis can be done in *three* steps. In the first the clones are generated. **a** The elements of the FYT system present on the second and fourth chromosomes that allow the generation of somatic clones with recessive mutations of the fourth chromosome. **b** Upon induction of heat inducible source of Flp- recombinase in a dividing cell, a mitotic recombination takes places. **c** This recombination simultaneously removes the fourth chromosome appended on the second, exposes a recessive mutation *dati*$^{1}$ located on the fourth chromosome and labels these cells and its descendants with *green florescent protein* (GFP). In *step 2* the female sex appeal is measured using courtship indexes (CIs) and acceptance or rejection recorded. In *step 3*, the clones (In *green*) are visualized by confocal microscopy and their position as well as status as accepting or rejecting tabulated

should converge to the same "processing" unit. This view is supported by anatomical studies that reveal a high degree of convergence of different sensory inputs in the brain (Kamikouchi et al. 2006; Otsuna and Ito 2006; Tanaka et al. 2012).

Courtship appears to cause a progressive and higher female engagement and less sensitivity to other stimuli (food, light air currents, and so on) as if the response to other stimuli is reduced. This view is supported by studies that have shown that male courtship effectively reduces the locomotor activity of females (Tompkins et al. 1982).

These pieces of information suggest that a first step in decision-making might be to stimulate as many neurons that make part of the circuit as possible, and to then implement secondary routines that shut down or attenuate other sensory paths. This view suggests that decision-making could involve a massive stimulation followed by an inhibition that eliminates the stimulation of neurons dedicated to alternative responses or scenarios. This possibility could be tested by knocking down the expression of genes required in female acceptance, such as *dati*, in inhibitory or excitatory circuits to determine what kind of informational content becomes disrupted. Are *dati* mutants stimulated but unable to implement secondary inhibitory responses or are they not appropriately stimulated (e.g., disruption of convergence and dispersion)? So far, it appears that the removal of *dati* in inhibitory GABAergic neurons does not interfere with acceptance. Interestingly, the loss of *dati* in cholinergic neurons leads to a dramatic increase in rejection (Giesey et al. 2013).

### *19.2.4 Variation of Mating Choice Across Species and Mating Asymmetry of Closely Related Drosophila Species*

So far, we have been discussing female decision-making in the context of a single species. However, if we sample the choice that females make in different species, we usually find that closely related species of many different animals exhibit a puzzling mating asymmetry (Wirtz 1999). That is the case of three species of *Drosophila*: *D. melanogaster*, *D. simulans*, and *D. sechellia*. In these closely related species, females either have the ability to mate with males of all three species (e.g., *D. melanogaster* females), or selectively accept only two of the three (e.g., *D. simulans* females), or still accept only one of the three (e.g., *D. sechellia* females) (Fig. 19.3). Although not impossible, it is unlikely that these variations might be due to sex determination. It is more plausible that during evolution, either the genes required for sensory acquisition or information processing in females might have been modified across these species. For example, one could envision that changes in auditory sensory perception might lead to an inability to detect some frequencies conveyed by the courtship song and as a result, the song of a

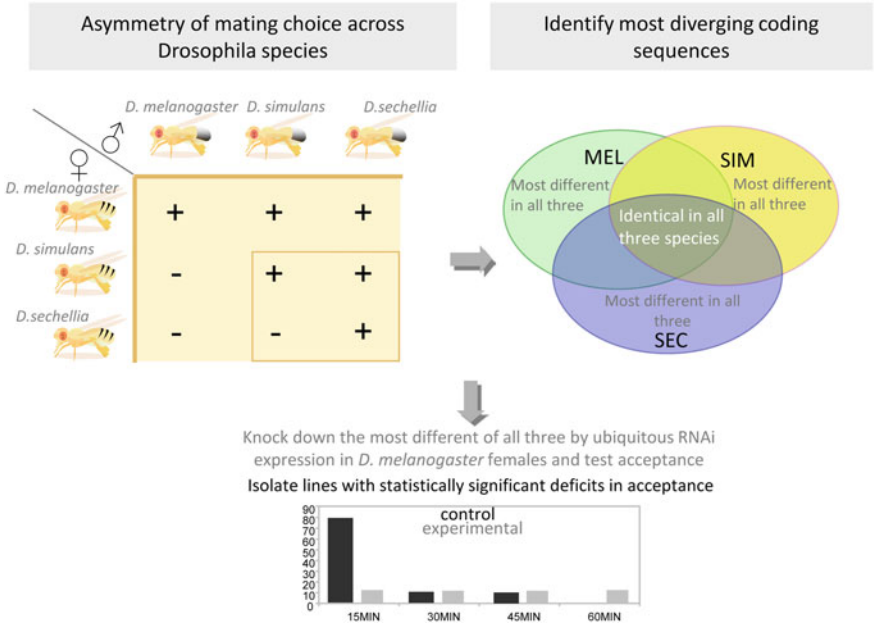



**Fig. 19.3** Searching and validating genes required female mate choice. On the *left* above are indicated the profiles of mating choice of *D. melanogaster, D. simulans* and *D. sechellia*. Females of all species prefer conspecific males, but in the absence of such males they may or may not mate with heterospecific males. + signs indicate that mating happens and − signs indicate that rarely if ever such matings happen. This profile suggests that different genes in each species confer a preference for specific males. To test this hypothesis, the most divergent of the *three species* can be identified (diagram to the *right*, above) and their products removed by RNA interference in *D. melanogaster*. This is done by using a source of Gal4 combined to an UAS with an RNAi for the gene of interest and searching for knockdowns that exhibit statistically significant deficits in acceptance. The graph below shows the percentage of the mating of wild type females over time (*black bars*) and a possible result when the removal of a gene reduces mating

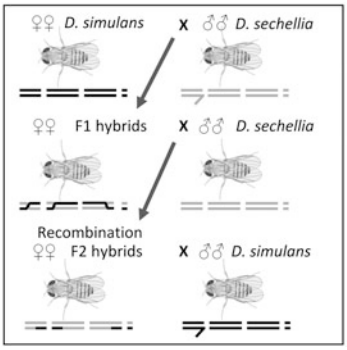
related species would be either annoying or meaningless. An annoying song could serve as a simple mechanism to discriminate a foreign species. Variations that blind or distort the perception of vision, olfaction, and taste are also possible. Although some genes required in different sensory modalities were modified in these species (Guo and Kim 2007; Sousa-Neves and Rosas 2010), it remains unknown if any of these genes is responsible for the different choices that these species make. Variants of genes that are involved in processing in the brain would be also of special interest since they could provide insights into the mechanism that allow different organisms to respond negatively or positively to the same stimulus and thus result in a different behavioral choice. However, so far, the evidence for such genes is still elusive.

## 19.2.5 Isolating Candidate Genes for Female Mating Choice Across Species

### 19.2.5.1 Application of Genome-Wide Comparisons Among Related Species

There are different ways of searching for genes involved in female mating choice across species. Many behavioral studies rely on testing known genes from one species in a foreign species. An obvious limitation in these studies is that no new genes involved in behavior can be discovered and tested. Other studies rely heavily on correlations between behavior and a genotype. In the beginning of this chapter, we discussed some of the problems of this strategy. How then do we go about identifying these genes? One approach is to use bioinformatics tools to search for particular gene variations in the genome of species that exhibit different choices (e.g., *D. simulans* and *D. sechellia*, see Fig. 19.3). Alternatively, one can create fertile hybrids between such two species and separate genetically the components of this choice (Figs. 19.4 and 19.5).

Following the first approach, pair wise comparisons of the coding sequences of *D. melanogaster, D. simulans*, and *D. sechellia* revealed a notable distribution in the nucleotide divergence of genes (Sousa-Neves and Rosas 2010). First, most genes in these species are nearly identical (i.e., they do not vary more than 5 % at




| Introgressions | Inheritance | Matches Mendelian proportion | Number of independently segregating genes |
|---|---|---|---|
| *D. simulans* strain 1 X *D. sechellia* strain 1 | dominant | 12.5% | 3 genes |
| *D. simulans* strain 2 X *D. sechellia* strain 1 | dominant | 25.0% | 2 genes |
| *D. simulans* strain 2 X *D. sechellia* strain 2 | dominant | 25.0% | 2 genes |

**Fig. 19.4** Searching and validating genes involved in species mating preference. This approach relies on creating F1 hybrids between *D. simulans* and *D. sechellia* with marked chromosomes, backcrossing them to *D. sechellia* males and testing which F2 females genotypes choose *D. simulans* males. By analyzing the number of females that produce progeny in two different strains of *D. simulans* and two different strains of *D. sechellia*, it is possible to determine whether the frequency of females that mated is consistent with a segregation of one, two or three independently segregating factors. In the case illustrated, two different strains of *D. simulans* have a different number of dominant mate choice genes. To test in which chromosomes these determinants are localized, we can analyze the frequency that the *D. simulans* chromosome appear. In this case, any chromosome of *D. simulans* that appears at a frequency greater than 25 % (i.e., maximum number of non-recombined chromosomes) will be a carrier of the gene

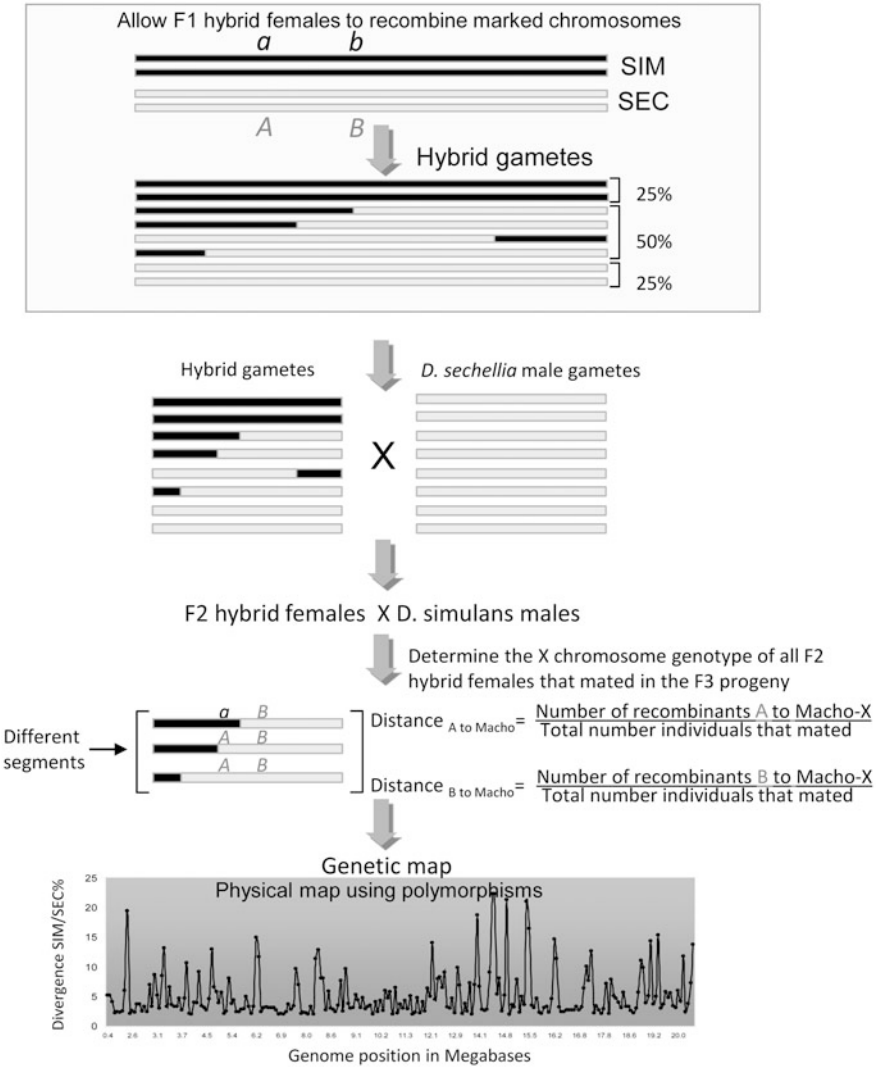



**Fig. 19.5** Mapping genes required for mate choice in F2 hybrids. This mapping tests whether there is a fixed distance between the gene that promotes matings between *D. simulans* females and *D. simulans* males. The *D. simulans* chromosomes carrying recessive markers are allowed to recombine with *D. sechellia* and the recombinant and non recombinant gametes fertilized by a *D. sechellia* sperm like depicted in Fig. 19.4. The adult F2 females that emerge are tested for their ability to mate with *D. simulans* males and the genotypes of those that mate is determined. From the frequency of recessive markers in the F3 we can either have a fixed position between the markers and the determinant or no distance. This approach reveals that there is a fixed distance between markers on the X chromosome and the mate choice determinant, which rules out polygenic inheritance. Using this genetic distance it is possible to determine the approximate position of the Mate choice determinant and test molecularly this inference using polymorphisms in coding sequences that distinguish the two species (*graph at the bottom*)

the nucleotide level). Second, some genes are very divergent between *D. simulans* and *D. sechellia* in regard to *D. melanogaster*. Third, a group of genes is divergent in all three species. This approach does not distinguish whether or not these genes are under positive selection, which would be a complex and ambiguous analysis since it depends on prior knowledge of all protein domain functions. However, the linear versus nonlinear distribution pattern seen above effectively singled out genes that evolve much faster than others.

The latter group of genes evolving faster in all three species is of particular interest since each female of these species prefers males of its own type and these preferences are likely to be encoded by species-specific gene variants. These fast-evolving genes can be tested for specific functions in female mate choice by knocking them out in *D. melanogaster* with RNA interference (Fig. 19.3). Such screening identified a number of genes that when knocked down, significantly reduce female acceptance for conspecific males (Chu and Sousa-Neves 2013). It will be interesting to find out whether these genes function at the level of sensory acquisition (i.e., olfaction, vision, taste, and sound detection), neural processing in the brain, or hormonal signaling.

#### 19.2.5.2 Interspecific Chromosomal Replacement to Locate Genes Required for Female Mating Choice

Another way to identify genes required for female mating choice is through genetic mapping in interspecific hybrids (Fig. 19.4). As mentioned before, *D. simulans* females mate with either *D. simulans* or *D. sechellia* males, while *D. sechellia* females only mate with *D. sechellia* males. Since fertile hybrid females can be obtained and the chromosomes of these species are completely syntenic, it is possible to recombine segments of *D. simulans* onto an animal that is otherwise *D. sechellia* and select the chromosomal segments responsible for the choice for *D. simulans* males. It has been shown before that F1 hybrid females accept males from either species, indicating that the preference for *D. simulans* males by *D. simulans* females is dominant over the rejection of *D. simulans* males by *D. sechellia* females (Coyne 1992) (Chu et al. 2013). This observation allows for the design an experiment in which F1 hybrid females carrying marked chromosomes are backcrossed to *D. sechellia* males, and the F2 hybrid female progeny tested for their ability to mate with *D. simulans* males (Fig. 19.4). These F2 hybrid females carry half of their genome from *D. sechellia* and different pieces of the *D. simulans* chromosomes. By analyzing the frequency of F2 females that mated with *D. simulans* males and the frequency of recessive markers that appear in the F3 generation, it was possible to rule out that mate choice relies on multiple genes scattered across the genome. Instead, this choice is associated to only two functionally redundant genes on the X chromosome and one or two autosomal independently segregating units (depending on the strains used), named as *Mate choice* (*Macho*) factors (Fig. 19.4). The Macho factors linked to the X chromosome were

mapped to single positions (Chu et al. 2013) using the approach outlined in Fig. 19.5. In sum, female choice does not appear to be polygenic, but governed by discrete genetic units.

### *19.2.6 Summary*

Behaviors have been traditionally viewed as a complex series of physiological responses to the environment where genetics plays a relatively small and usually unclear part. In this view, the genetic components of behavior are usually assumed to be polygenic. The study of Drosophila courtship gives us a platform to probe both the genetic and neural basis of decision-making. The ability to discriminate heterospecific males and poor-performing conspecific can be quantified and allows for the identification of several mutants that depart from normal wild-type behavior. Together, studies in *D. melanogaster* and its sibling species *D. simulans* and *D. sechellia* suggest that female choice can be broken down in discrete simple traits. They also allow the separation of variables such as sex appeal from female choice. The expectation is that both approaches combined may allow future generalizations regarding the evolutionary process of behavioral differentiation and the molecular mechanism behind them.

## References

Bastock M (1956) A gene mutation which changes a behavior pattern. Evolution 10:421–439
Benton MJ (1993) Fossil record 2, Springer
Benton MJ (2004) Vertebrate palaeontology, Wiley
Bovbjerg RV (1953) Dominance order in the crayfish orconectes virilis (Hagen). Physiol Zool 26:173–178
Chapman T, Bangham J, Vinti G, Seifried B, Lung O, Wolfner MF, Smith HK, Partridge L (2003) The sex peptide of Drosophila melanogaster: female post-mating responses analyzed by using RNA interference. PNAS 100:9923–9928
Chu Y and Sousa-Neves R (2013) An RNAi screening for genes involved in female mate choice in *Drosophila melanogaster*. 54th Annual Drosophila Research Conference, 631A
Chu Y, Yang E, Schinaman JM, Chahda S, Sousa-Neves R (2013) Genetic Analysis of Mate Discrimination in *Drosophila simulans*. Evolution DOI: 10.1111/evo.12115
Connolly K, Cook R (1973) rejection responses by female drosophila melanogaster: their ontogeny, causality and effects upon the behaviour of the courting male. Behaviour 44:142–166
Coyne JA (1992) Genetics of sexual isolation in females of the Drosophila simulans species complex. Genet Res 60:25–31
Esch T, Mesce KA, Kristan WB (2002) Evidence for sequential decision making in the medicinal leech. J Neurosci 22:11045–11054
Ewing AW (1964) The influence of wing area on the courting behavior of Drosophila melanogaster. Anim Behav 12:316–320

Finley KD, Taylor BJ, Milstein M, McKeown M (1997) Dissatisfaction, a gene involved in sex-specific behavior and neural development of Drosophila melanogaster. Proc Natl Acad Sci USA 94(3):913–918

Fischhoff IR, Sundaresan SR, Cordingley J, Larkin HM, Sellier M-J, Rubenstein DI (2007) Social relationships and reproductive state influence leadership roles in movements of plains zebra, Equus burchellii. Anim Behav 73:825–831

Giesey R, Schinaman J, Mizutani C and Sousa-Neves R (2013) Role of dati in the neuronal circuitry of female acceptance of male courtship in *Drosophila melanogaster*. 54th Annual Drosophila Research Conference, 1014A

Gill K (1963) A mutation causing abnormal courtship and mating behavior in males of Drosophila melanogaster. Am Zool 3:507

Guo S, Kim J (2007) Molecular evolution of drosophila odorant receptor genes. Mol Biol Evol 24:1198–1207

Hall JC (1978) Courtship among males due to a male-sterile mutation in Drosophila melanogaster. Behav Genet 8:125–141

Juni N, Yamamoto D (2009) Genetic analysis of chaste, a new mutation of Drosophila melanogaster characterized by extremely low female sexual receptivity. J Neurogenet 23:329–340

Kamikouchi A, Shimada T, Ito K (2006) Comprehensive classification of the auditory sensory projections in the brain of the fruit fly Drosophila melanogaster. J Comp Neurol 499:317–356

Keiko S, Naoto J, Daisuke Y (1997) Enhanced mate refusal in female drosophila induced by a mutation in the spinster locus. Appl Entomol Zool 32:235–243

Manning A (1960) The sexual behaviour of two sibling drosophila species. Behaviour 15:123–145

Nusslein-Volhard C, Wieschaus E (1980) Mutations affecting segment number and polarity in Drosophila. Nature 287:795–801

Otsuna H, Ito K (2006) Systematic analysis of the visual projection neurons of Drosophila melanogaster, I Lobula-specific pathways. J Comp Neurol 497:928–958

Siebeck UE, Litherland L, Wallis GM (2009) Shape learning and discrimination in reef fish. J Exp Biol 212:2113–2119

Sousa-Neves R, Rosas A (2010) An analysis of genetic changes during the Divergence of Drosophila species. PLoS ONE 5:e10485

Sousa-Neves R, Schinaman JM (2012) A novel genetic tool for clonal analysis of fourth chromosome mutations. Fly 6(1):49–56

Spieth HT (1974) Courtship behavior in Drosophila. Annu Rev Entomol 19:385–405

Sturtevant AH (1920) Genetic studies on Drosophila Simulans, I Introduction. Hybrids with Drosophila Melanogaster. Genetics 5:488–500

Tanaka NK, Endo K, Ito K (2012) Organization of antennal lobe-associated neurons in adult Drosophila melanogaster brain. J Comp Neurol 520:4067–4130

Tompkins L, Gross AC, Hall JC, Gailey DA, Siegel RW (1982) The role of female movement in the sexual behavior of Drosophila melanogaster. Behav Genet 12:295–307

Wirtz P (1999) Mother species-father species: unidirectional hybridization in animals with female choice. Anim Behav 58, 1–12

Yapici N, Kim Y-J, Ribeiro C, Dickson BJ (2008) A receptor that mediates the post-mating switch in Drosophila reproductive behaviour. Nature 451:33–37

# Index



P. Pontarotti (ed.), *Evolutionary Biology: Exobiology and Evolutionary Mechanisms*,
DOI: 10.1007/978-3-642-38212-3, © Springer-Verlag Berlin Heidelberg 2013