HPLS (2017) 39:33
DOI 10.1007/s40656-017-0160-3

ORIGINAL PAPER

Hierarchy, determinism, and specificity in theories of development and evolution

Ute Deichmann[1]

Published online: 16 October 2017

Abstract The concepts of hierarchical organization, genetic determinism and biological specificity (for example of species, biologically relevant macromolecules, or genes) have played a crucial role in biology as a modern experimental science since its beginnings in the nineteenth century. The idea of genetic information (specificity) and genetic determination was at the basis of molecular biology that developed in the 1940s with macromolecules, viruses and prokaryotes as major objects of research often labelled "reductionist". However, the concepts have been marginalized or rejected in some of the research that in the late 1960s began to focus additionally on the molecularization of complex biological structures and functions using systems approaches. This paper challenges the view that 'molecular reductionism' has been successfully replaced by holism and a focus on the collective behaviour of cellular entities. It argues instead that there are more fertile replacements for molecular 'reductionism', in which genomics, embryology, biochemistry, and computer science intertwine and result in research that is as exact and causally predictive as earlier molecular biology.

Keywords Mechanistic systems biology · Molecular reductionism · Holism · DST · Regulatory Genome · Big-data genomics

When I kill a fly, I don't think and must not think which organization is destroyed (Goethe (1959 [1817], 802), translation UD).

Every living thing is not single, but multiple; even insofar as it appears to us as an individual it remains nonetheless an association of living self-sufficient beings, which though alike in idea or plan, can in their manifestations be identical or similar, unlike or dissimilar. The less developed the creature is, the more alike or similar are these parts and the more they resemble the whole. The more highly developed the creature becomes, the more dissimilar become the parts. The more alike the parts are, the less they are

✉ Ute Deichmann
uted@post.bgu.ac.il

[1] Jacques Loeb Centre for the History and Philosophy of the Life Sciences, Ben-Gurion University of the Negev, P.O. Box 653, 84105 Beer Sheva, Israel

subordinated. Subordination of parts points to a more highly developed creature (Goethe [1817], English translation quoted from Reynolds (2008), 126).

1 Introduction

The term *hierarchy*, an ancient word originally relating to the religious sphere, has had a diversity of meanings throughout history (Verdier 2006). It is understood here as a systemic organization into levels that are subordinated by relationships of power or control or by an order between elements that are classified in different nested categories, as in the Linnaean taxonomic system. Hierarchies are at the core of many complex systems such as biological systems (Pumain 2006, 1–3). According to Michel Morange, in order to understand "the logic of life, we need to understand its structural hierarchy" (Morange 2001, 160). Hierarchies also play an important role in the models of how living systems are organized through evolution and how they function (Pavé 2006), as well as in the causal analysis of development through Gene Regulatory Networks (GRNs; see Peter and Davidson 2015).

Biological or genetic *determinism* here is not related to human, mental, or intellectual abilities, nor to claims about the determinative role of individual biology for social and other achievements. It refers instead to the idea that basic characteristics, such as body plans of species, genera and higher taxonomic ranks are determined by hereditary factors. This idea became increasingly subject to examination when biochemistry, cell theory, Mendelian genetics and molecular genetics provided the necessary experimental tools for the analysis of such determinisms. Already around 1900 it was widely believed that "differences in the constitution of proteins determine the species specificity" (Loeb 1916, p. 65). According to Raphael Falk, "biology became more and more established along the [nineteenth] century as a determinist science like physics and chemistry, reducible to its composing elements" (Falk 2014, 186). The fact that biology was never completely deterministic because of the occurrence of random events such as the undirected nature of mutations or the separation and recombination of maternal and paternal chromosomes did not contradict this aim of biology as an exact science. It should be added that determinism is not necessarily connected with reductionism, as the determination of early embryonic development by a complex system of regulatory genomic genes shows.

According to Michel Morange, the existence of an organizational and structural organic hierarchy ranging from protein machines via cells to populations has undermined genetic determination, because the "action of gene products is only indirectly expressed" (Morange 2001, 159). He does not, however, call into the question the causal role of genes; as he states, there is a "precise causal chain linking the product of a gene to the actions of that gene within the organism" (ibid.). Here I do not deal with the determination of single traits by particular genes, but the hereditary determination of animals' body plans through the control of their early development by gene networks and their products.

The concept of *biological specificity* holds that individual organisms, species, and higher entities in the hierarchy of taxonomic ranks (genera, orders, classes etc.) are

special and different from other entities of the same rank. This specificity expresses itself in body structures and proteins that are specific to organisms, species, and so on and is now explained by the existence of specific information encoded in the genome. Since the late nineteenth century, biological specificity has been regarded as a basic characteristic of life, as expressed by chemist Linus Pauling:

> Biological specificity is the set of characteristics of living organisms or constituents of living organisms of being special or doing something special. Each animal or plant species is special. It differs in some way from all other species. … Biological specificity is the major problem about understanding life (Quoted from Marinacci 1995, p. 96).

Even though some ideas of *hierarchy* related to the phenomena of life (for example in Aristotle's *scala naturae*; for details see Roth 2011), hereditary *determination* of traits, and constant differences between groups of organisms (*specificity*) have existed since ancient times, contrasting notions of often inexplicable change and fluidity were likewise widespread in natural history until the eighteenth century, in some areas of research until the nineteenth century. I will show in this article that the beginning of biology as a modern experimental science was closely related to the introduction, in part revival, of the concepts of hierarchical organization, genetic determinism and biological specificity (or genetic information), into basic areas of research, concepts which have remained fruitful in much of basic research until the present time (see e.g. Jacob 1973; Mazumdar 1995; Pavé 2006; Deichmann 2007b).

I will argue moreover that since the late twentieth century, the importance of some or all of these concepts has been neglected or rejected in research related to complex characteristics such as development that attempts to overcome what has been called 'molecular reductionism' by invoking principles of fluidity, holism, and the collective behaviour of cellular entities. More importantly, the increasing number of large genomic sequencing projects has given rise to a new kind of scientific practice and reasoning that largely excludes causal analysis.

In what follows, I will first provide a historical sketch of the establishment of hierarchy, determinism, and specificity as central biological concepts. I will then examine the fertility of these concepts in recent research on development, genetics and evolution, and their neglect or outright rejection in other research. The epistemological implications will also be explored.

2 The establishment of hierarchy, determinism, and specificity as central biological concepts: a short survey

It is widely believed that the European world in medieval times was an unchanging world regarding society as well as nature. However, as philosopher Amundson (2007 [2005], 35–40) has shown, this assumption does not hold true, at least with regard to ideas about the living nature. It is not well known that species fixism was not an ancient Christian belief, but became widely accepted for the first time among naturalists and theologians only during the eighteenth century after it had been

established by the botanist and taxonomist Carl Linnaeus and colleagues. Before that, naturalists, theologians and common people held a large variety of "transmutational beliefs" (ibid., 35, 36). To mention just a few: there was a wide acceptance of the idea of the spontaneous generation of small organisms such as eels, frogs, insects, worms, and even mice from inanimate matter or decaying organisms. Organisms were also believed to change during their lifetime, such as from "worm" to insect in insect metamorphosis. In various cases hybridization was thought to be the cause of transmutations across generations; thus the giraffe was thought to have arisen from a pairing of a leopard with a camel. It was generally believed that plants could change their species. According to Francis Bacon, not only might one species change into another, but it was also a "matter of chance what the transmutation would be" (Ibid., 36). That the climate would permanently modify plant species, e.g. rye into cornflower, was common belief.

Given the prevalence of spontaneous generation and transmutations, the idea of species fixism was an important innovation that became a crucial requisite for any theory of evolution because it made the construction of a Natural System possible. In the words of Amundson (ibid., 39) "evolution theory could no more have been discovered by a prefixist transmutationist than the Bohr atom could have been discovered by an alchemist". The idea of species specificity and the hierarchical classification of species in artificial or natural systems were therefore pre-requisites for any theories of organic evolution.

In other areas of natural history, notions of frequent changes lasted even longer. Thus spontaneous generation continued to be upheld for micro-organisms until the late nineteenth century, long after having been abandoned regarding higher organisms. It was correlated with the idea of microbial polymorphism i.e. their occurrence in a multiplicity of forms. These notions of change had nothing to do with, and did not generate, modern ideas of bacterial genetics and horizontal gene transfer. On the contrary, the concept of the individuality and specificity of bacteria, that is, of bacterial species which replaced the notion of constantly changing bacterial types, was a pre-requisite for studies of the causes of both stability and change of bacteria, and of bacterial genetics (Deichmann 2007a).

Until the advent of Mendelian genetics and the recognition of chromosomes as causal agents of heredity and development in early twentieth century, a variety of factors were believed to influence hereditary traits, among them environmental factors which were thought to act in a directed way on germ cells or on the embryo. Male sperm allegedly also influenced a female to such an extent that her future offspring with another male showed resemblances with the first male, a belief shared also by Darwin (Darwin 1868, II, chapter 27). In line with widely-held beliefs of his time, Darwin concluded that "variability is not a principle co-ordinate with life or reproduction, but results from special causes, generally from changed conditions acting during successive generations" (ibid.).

With the increasing acceptance of August Weismann's germ plasm theory (1893), which stated that only the hereditary material of the germ cells and not that of the soma cells is transmitted, and of the chromosome theory of heredity, biological determinism as genetic determinism became widely accepted. But nineteenth century hypotheses by Weismann, Theodor Boveri, Edward Wilson and

other biologists that development, too, is causally related to chromosomes, was superseded, in the first half of the twentieth century, by a non-causal, phenomenological explanation of experimental results regarding development, in which genes no longer played a role. Developmental genetics originated only in the 1960s (apart from work in the 1930s on gene "hormones", see e.g. Müller-Wille and Rheinberger (2012, 153ff.).

Nineteenth and early twentieth century biochemistry devoted itself to the study of molecules believed to be characteristic of life, such as proteins and enzymes. Emil Fischer's lock and key model of enzyme action, followed by Paul Ehrlich's proposition of the side-chain (today: receptor) theory of antibodies, highlighted the importance of specificity. During the subsequent period of biocolloidy, unspecific fluidity and vagueness superseded specificity (Florkin 1972, 271–275; Deichmann 2007b). The demonstration of the existence of biologically relevant macromolecules with clear-cut physical and chemical properties in the 1930s, followed by the elucidation of the DNA double helix structure and the transformation of biological specificity into genetic information in the 1950s, initiated the first phase of molecular biology. With its exclusion of the flow of information from environment or proteins to DNA, and its focus on genetic information and gene regulation, this phase of molecular biology epitomized the importance of genetic determinism and specificity at the molecular level of life phenomena. This molecularization of central biological concepts focused on the macromolecular description of the gene, its replication, expression and regulation in prokaryotes. Phenomena such as heredity and mutation were explained by events in underlying simpler parts, e.g. genes or macromolecules; viruses and micro-organisms were used as models for higher organisms.

Starting in the 1960s research began to search for molecular explanations for phenomena that occurred only in higher organisms, such as development, behaviour or certain diseases. Prokaryotes were no longer primary objects of molecular biological research, which now dealt with a large number of different eukaryotic organisms including humans. From the beginning of the 2000s, the availability of new techniques made systems approaches possible. For some scientists the study of complex systems was just another step in genomics research, with post-genomics techniques allowing simultaneous characterization of many or all of an organism's genes and proteins and their interactions. Others, however, disappointed with the Human Genome Project that left many questions unanswered, tried to develop new visions in order to find explanations. For them, the study of complexity required getting rid of the reductionism of molecular biology by returning to a holistic biology. The principles of hierarchy, genetic determinism and specificity became marginalized in these new approaches. The next section will survey these approaches and analyze the importance that the concepts of hierarchy, determinism, and specificity have within each of these approaches.

3 Holistic, non-mechanistic responses to "molecular reductionism"

Carl Woese, an early proponent of a new holistic biology, some years ago launched a heavy attack on reductionism, which he seems to have equated with biological engineering and application: "The pinnacle of fundamentalist reductionism in biology was reached with the Watson–Crick structure of DNA… Biology today is little more than an engineering discipline… It must choose between two paths: either continue on its current track, in which case it will become mired in the present, in application, or break free of reductionist hegemony, reintegrate itself, and press forward once more as a fundamental science. The latter course means an emphasis on holistic, 'nonlinear,' emergent biology" (Woese 2004, 185).

While many may agree with his criticism, his suggested remedy is not shared by most molecular biologists. Woese's outstanding research, for example on the genetic code, and his discovery of archaebacteria, followed conventional molecular biological methodology. In his later philosophical generalizations, however, he denounced this conventional approach: "The molecular reductionism that dominated twentieth-century biology will be superseded by an interdisciplinary approach that embraces collective phenomena" (Goldenfeld and Woese 2007, 369). The authors use the existence of horizontal gene transfer between bacteria ("gene-swapping collectives") to demand a revision to concepts such as organism, species and evolution itself (ibid). In his vision of "twenty-first century biology", Woese obviously gave up the aim, shared by many of his contemporaries, to use post-genomic tools to tackle phenomena of complexity with molecular mechanistic research. Instead, he connected with descriptive non-reductionist nineteenth-century morphology, morphogenesis, and evolution (Woese 2004, 176).

Woese's vision of a descriptive biology that embraces collective phenomena and a "continuity of energy flux and informational transfer" from the genome, through cells and the environment (Goldenfeld and Woese 2007, 369), called into question the boundaries between organism and environment and marginalized molecular causation and specificity. This vision shows parallels with postmodernist reasoning, its blurring of boundaries, its marginalization of causation and predilection for phenomenological descriptions.

Philosopher Bruno Latour, who calls himself non-modernist or a-modernist, has most vigorously expressed and propagated a notion of collectivity in which hierarchies, specificities, even the reality of nature as a distinct entity have no place (Latour 1993, 2004). According to him, society perceives the 'scientific culture' as being characterized by certainty, straightness, objectivity, and coldness, in contrast to the 'culture of research', characterized by uncertainty, warmth, and emotions, a distinction of which he does not seem to be critical. Latour suggests the notion of 'collective experiment' in order to capture the new spirit of the times based on the new deal between research and society (Latour 1998). The notions of 'research' and collectivity presented by Latour have become increasingly influential. They do not pay respect to the fact that both historically and at the present time, science is full of examples of scientists who have been the opposite of unemotional and detached, but

who at the same time clearly aimed at achieving certainty and objectivity in their research.

Collectivity has become a fashionable topic in the philosophy of biology and is adapted also by some scientists. This is illustrated, for example, by titles such as "Rethinking Immunity: Moving from the Autonomous Individual to the Ecological Collective" or "Computing the State of the Body: Collective Dialogue in Autoimmunity and Tumor Immunity" at the 29th Annual International Workshop in the History and Philosophy of Science "Landscapes of Collectivity in the Life Sciences" at Tel Aviv University in June 2015.[1]

Postmodernist tendencies such as the blurring of boundaries, in particular between environment and organisms, and between causal hierarchies, are also apparent in other anti-mechanistic holistic studies. Developmental Systems Theory (DST), not a clear-cut theory, but a movement of philosophers and biologists with a holistic approach to development, is a case in point. While some scholars associated with this movement such as evolutionary biologist and psychologist Russell D. Gray hold that accepting DST leads to a different kind of scientific research, others, such as psychologist and philosopher of science Susan Oyama, have presented DST as a general and abstract "way of seeing" the biological world and the investigation of it (Godfrey-Smith 2001, 283). The main representatives of DST stated that their intention was to address in a new way the many unresolved "vexed questions" of the genetic determinism or evolutionary causation of complex biological traits such as human behaviour and their underlying oppositions such as nature or nurture, genes or environment, biology or culture. They define DST as "an attempt to do biology without these dichotomies" (Oyama et al. 2001, 1).

DST is fundamentally opposed to reductionism of all kinds and is based on the idea that all biological processes (including evolution and development) operate by continually assembling new structures, which transcend the structures from which they arose and have their own systematic characteristics, information, functions and laws. The rejection of genetic determinism of any hereditary trait or development is central. According to Godfrey-Smith (2001, 283), the informational gene is the "'preformationists' last stand for DST", that is something that has to be overcome. DST advocates multiple interacting causes, none of which more relevant than the other; the environment being equally important as genes in bringing about development. In this logic, the attribution of different causal roles to genes and environmental factors in development or evolution becomes obsolete. They are replaced by an environment—organism system changing over time. Development (as well as evolution) is brought about by construction through the interaction of many different factors: "The life cycle of an organism is developmentally constructed, not programmed or preformed. It comes into being through interactions of the organism and its surroundings as well as interactions within the organism" (Oyama et al. 2001, 4). Some of the founders of DST go so far as to call into question the distinction between the organism and environment in models of evolution: "Perhaps the most radical departure [from evolutionary theory] is that the

[1] See the programme of the workshop available at http://www.vanleer.org.il/sites/files/LandscapesCollectivity_3.pdf (accessed 20/09/2017).

separation of organism and environment is called into question" (Griffiths and Gray 1994, 300). In later publications they propose the use of the symbol "OE" in models of evolution and, completely disregarding the fact that the development of most organisms proceeds the same in different environments, claim that "there is no distinction between organism and environment" (Griffiths and Gray 2001, 207; see also Griffiths and Gray 2004).

What is at issue here is not the existence of multiple interacting causes during development—no developmental biologist will deny this—but the rejection of the notion held by most developmental and evolutionary biologists of the existence some kind of guiding genetic program for development with the genome as primary cause (for the various historical meanings of the term "genetic program" see Peluffo 2015). This rejection, and the claim that none of the causes that bring about development is more relevant than the other, is reminiscent of Aristotle's dispute with materialistic philosophers about the *Pangenesis* hypothesis of Hippocrates, i e. the idea that particles from every part of the body mix and form the generative elements. According to Aristotle development was a gradual process of increasing complexity from initially homogeneous material, an idea for which William Harvey in the seventeenth century introduced the term *epigenesis*, with *genesis* (gr.) meaning origin, and *epi* on or after. Aristotle and his followers considered it inconceivable that the material particles could form an organized body without what he called an immaterial principle of form (which today could be called a principle of organization and which since the early twentieth century has been widely accepted to reside in, to use a modern term, genomic information). The promoters of DST rarely ask how it is that development results in a functioning organization and, moreover, that development of individuals of a species always results in the same body plan, independently of the environment in which it takes place.

A hundred years ago, the awareness of these facts had led to the rise of neo-vitalism on the one hand, and the idea of genomic control of development on the other. Neo-vitalist Hans Driesch's view that it was inconceivable "how a self-determining system can increase its own initial complexity by interaction of its chemical and physical components;" was also shared by anti-vitalists such as August Weismann (quoted from Wilson [1896] 1928, 1110). While Driesch invoked an immaterial principle—entelechy—as a guiding principle, Theodor Boveri and Edmund B. Wilson accepted chromosomes as "an original preformation", envisaging that the apparent cytoplasmic epigenesis was based on the transmission of nuclear preformation i.e. chromosomes. This way they were able to explain the apparent epigenesis in development in a mechanistic way (Deichmann 2014).

In contrast, DST cannot answer in any plausible or convincing way the question of what causes the early development of a species to always proceed in the same way; in other words, what is the cause for the stability of body plan within species, genus, or higher taxonomic ranks over evolutionary times, a stability without which the question of evolutionary change would be obsolete. DST proponents suggest an "alternative explanation of transgenerational stability of form", in which "species-typical traits are constructed by a structured set of species-typical developmental resources in a self-organizing process that does not need a central source of information. Some of these developmental resources are genetic, others, from the

cytoplasmic machinery of the zygote to the social events required for human psychological development, are non genetic" (Griffiths and Gray 1994, 283). However, the question remains of how the cytoplasmic machinery of the zygote or social events could possibly impact the robust development of the body plan in a particular species, or how self-organizing processes without a source of information can lead to a predictable outcome any more than they could 100 years ago when such an idea was ridiculed by Hans Driesch.

The DST proponents made it clear that they have revolutionary aims: "What we need is the 'stake-in-the-heart move'—a way of thinking about development that does not rely on a distinction between privileged, essential causes and merely supporting and interfering causes" (Oyama et al. 2001, 1). The rejection of the principles of hierarchy and genetic determinism and the emphasis on blurring boundaries is reminiscent of the postmodern principles of collectivity and the relativization of causes: attributing equal importance to the many causes that bring about development and denying the distinction between causal factors such as environment or genes result in the marginalization of causes altogether. Similarly, the above principles are weakened in interpretations of epigenetics as having relativized the importance of genes, as is frequently done in the arguments of DST proponents. These interpretations disregard the fact that genes are not turned on and off by epigenetics, but by DNA specific proteins (transcription factors or repressors). They direct the enzymes that bring about "epigenetic changes" i.e. DNA methylation or various kinds of histone modification to the correct places on the genome. That means the epigenetic changes are controlled by the genome (Deichmann 2016).

4 Big-data genomics

The Human Genome Project and other genomics projects, i.e. the sequencing of genomic DNA on an organism-wide scale and the creation of accessible databases thereof, have not only provided huge amounts of new data for researchers and thus contributed to transforming biology. They have also begun to create a new type of science that has been characterized as being data-driven. This "data driven" science is sometimes contrasted with traditional "hypothesis-driven science", characterized as small-scale, often narrowly focused and best exemplified by molecular biology as its most prominent and successful enterprise (Aebersold et al. 2000).

The generation of large amounts of genomic data, independently of any hypothesis on function, and disregarding the conventional methods of hypothesis formation, prediction and experimental testing, has aroused strong debate about the fruitfulness of hypothesis-driven research in the era of data-driven research and on the potential dangers for science of the latter (see e.g. Weinberg 2010; Golub 2010). While a detailed analysis of this debate would exceed the scope of this paper, I will deal with some problems that arise when description, the search for patterns, and statistical correlations replace, and not only amend, the prominence of hypothesis, prediction, and experimental testing. To be sure, observation, description, and induction are also, unlike Popper's dictum of falsifiability, the starting point in most

"hypothesis-driven" research. But description alone does not provide explanations for mechanisms or generate causal explanations. Whereas the search for mechanisms and causes is central in "hypothesis-driven research", data-driven research does not go beyond establishing correlations and computer-generated hypotheses.

An example for big-data research is the work of the ENCCODE (Encyclopedia of DNA Elements) Consortium, a public research project launched and funded by the US National Human Genome Research Institute in September 2003 as a follow-up to the Human Genome Project. In 2012 it comprised 32 institutes and 442 consortium members; 1649 experiments were conducted on 147 cell-types, and the main paper has nearly 450 authors (Birney 2012). ENCODE aims at building a comprehensive parts list of functional elements in the human genome, including elements that act at the protein and RNA levels. Among the methods used to identify functional elements are comparative genomics, sequencing of a diverse number of RNA sources, integrative bioinformation methods, human curation, and immunoprecipitation of proteins that interact with DNA and RNA (see the project's website at https://www.encodeproject.org/).

While the databases provided by ENCODE are appreciated by researchers, approaches and interpretations of ENCODE as a scientific enterprise have been heavily disputed; their methods were strongly criticized and some of their results rejected as non-valid. In particular ENCODE's claim that they were able to ascribe functions for 80% of the genome (ENCODE 2012) has been widely rejected by renowned evolutionary biologists and genome researchers (e.g. Graur et al. 2015; Niu and Jiang 2013; Doolittle 2013; Morange 2014). According to Morange (2014), the authors of the various ENCODE publications did not distinguish between the different meanings of the term "function". An evolutionary function can be attributed to a DNA sequence when its modification has an effect on the fitness of the organism. A biochemical function can be concluded from the observation that a sequence binds a transcription factor or is transcribed into RNA. Some sequences only bear the mark of a potential function, for example, if they can bind a regulatory protein or initiate transcription without any evidence that this really happens in reality. Morange relates these problems to the fact that the ENCODE authors did not take into consideration that biological systems are noisy (for example transcription factors can interact with many non-functional sites), and that many observed biochemical activities were detrimental for the organism (for example those of transposons present in the genome).

The lack of relation to biological functions as they exist in reality was also a focus of the criticism of Davidson (2016) who, in addition, pointed to a shift in the underlying epistemology: A large gene regulatory network (GRN) that ENCODE predicted for *Drosophila* development (based on ChIP data obtained with 76 transcription factors for which antibodies were available) performed about the same as a totally randomized GRN; the ratio of correct prediction of the GRN to those of the randomized GRN was 1.04 (modENCODE 2010). According to Davidson, this example illustrates that "in the world of ENCODE genomics, the analysis is to be published on the basis that the measurements […] were made and analyzed by sophisticated mathematical statistics; whether the result has any power of predictability is not relevant" (Davidson 2016, 168). Concerning the topic of this

article, while specificity, the differences between DNA base sequences, plays an important role with ENCODE and other data-driven research, genetic determinism of early development cannot be established because there is no causal testing by experimental perturbation and prediction. This exclusion also prevents the establishment of hierarchies, not only of genes in the genome, but also of different kinds of functions, with those of an evolutionary importance in the center.

Many scientists suggest that hypothesis-driven and data-driven research should not be mutually exclusive but should complement each other, as is the case in epidemiology: hypotheses are the *result* of an epidemiological study, not the starting point (except for the hypothesis that there will be meaningful associations between environmental and genetic conditions). According to Kell and Oliver (2003), the criticism of not starting with a hypothesis may be directed to epidemiology as to gene-expression patterning, and they suggest that hypothesis-driven and data-driven research are not mutually exclusive but complement each other.

This is also the opinion of Hans Lehrach, a pioneer of genomics, who has been working for many years on the development of a truly personalized medicine. He became known for his phrase, early on, that hypothesis-driven research is out. In a personal communication, he made it clear (relating to Popper) that "we are not really talking about data versus hypotheses, since in all cases I can imagine we will need both, data and hypotheses. You can therefore either start with a hypothesis, and then generate the data to disprove it (that's all you can ever do with hypotheses anyway) or you first generate (maybe at higher cost) the data allowing the cost-effective testing of many hypotheses and then test many hypotheses on the available data (often gaining the enormous advantage that your power to disprove hypotheses is much higher)… In most real projects we keep alternating between data generating hypotheses generating the need to generate new data generating new hypotheses, etc. in a multistage process". Lehrach is critical of big-data projects that are not combined with "scientific evaluation that is based on more than genomic information" (personal communication to the author, 9 February 2016).

Similarly, cancer researcher Todd Golub emphasized the importance of a data-first approach though he appreciates the centrality of hypothesis-driven research for the elucidation of mechanisms (Golub 2010). Generally, most big-data projects are related to medical research, not to basic biological questions which require the elucidation of mechanisms. Researchers often do not proceed to prediction and experimental testing of the computer-generated hypotheses.

5 Mechanistic systems biology in development and evolution

The basic questions of molecular embryology are now successfully tackled not in holistic approaches, but in mechanistic systems research. Interestingly, in this research the concepts of hierarchy, genetic determination, and specificity are again of the greatest importance. This is demonstrated by the work of Eric Davidson and his collaborators, who integrated computer-generated big data into a systems approach that is based on experiments and aims at elucidating mechanisms and causal relationships. Davidson, a world-leading researcher in molecular

embryology, used genomic sequencing to solve fundamental mechanisms of experimental embryology; his work was not related to medical applications. He showed that at least in sea urchins early development is regulated by the genome, to him a logical necessity and requirement for evolution, because this genomic regulatory program for development "insures that within each species the outcome is extremely reproducible" and largely independent of changes in the environment. Davidson was the founder of the concept of developmental Gene Regulatory Networks (GRNs) executing the cascade of molecular mechanisms that transform an egg cell into a complex creature. These GRNs consist of regulatory genes, which encode transcription factors, and signalling genes, which encode ligands and receptors for intercellular communication, as well as the sequences that control the expression of each of these genes. Together, these components—elements of coding and non-coding DNA sequence—constitute the 'regulatory genome'. These interactions depend causally on the DNA sequences that determine which transcription factors control each gene. While in some developmental systems, for example the mammalian immune system, signalling is used in a less deterministic way (see the contribution by Ellen Rothenberg in this collection), Davidson was able to treat signalling deterministically because of the "canonical embryonic cleavage planes of the embryo that placed one cell, with its GRN state, predictably next to another cell with a different GRN state, and this was the same in every embryo". In addition, the sea urchin embryo did not grow in overall volume during the period that was studied, so that there were no problems of "scaling gene expression responses across a growing tissue" (Personal communication to the author by Ellen Rothenberg, 5 September 2016).

Davidson established the network parts—the specific sets of regulatory genes that are expressed in particular parts of the embryo such as its future endoderm or mesoderm—experimentally over many decades. He and his collaborators systematically examined the cell-type specific gene expression patterns before moving on from the "gene-by-gene characterization of the sea urchin embryo to full comprehensiveness" (Rothenberg 2016, 512). This systems approach was made possible when sequencing data of the whole sea-urchin genome was available. The Davidson lab succeeded in nearly completing the GRN of the (endoderm and mesoderm) sea-urchin development up to gastrulation. It included all regulatory genes, their target genes and all interactions between them. Among other things, it was based on systems-level perturbations for which secondary and tertiary effects had to be considered because of the functional interactions within the system, as well as of the effects of multiple inputs at each node of the system (Davidson 2016).

In order to test the proposed network, Davidson and his collaborators transformed it into a computational engine that according to the network structure would generate predictions of when and where in the embryo every regulatory gene should be expressed or not expressed (Peter et al. 2012). The computational predictions were then compared to direct experimental observation in order to find out whether the network sufficed to explain the spatial and temporal patterns of gene expression or how much of these patterns it explains. In their first complete computational model of the early sea urchin embryo network, Davidson and his collaborators showed that computation and observation (in normal and genetically manipulated

sea-urchins) agree in most cases (ibid.). Their results thus made it clear that a complex feature such as early development "can indeed be experimentally accessed and solved at a system-wide causal level" (Davidson 2016, 180). GRNs are studied now in different organisms; while other researchers extend the work to the cellular networks that are activated by GRNs.

The research by Davidson and others on developmental gene regulatory networks underlines the fundamental importance of all of the three concepts in question: *Hierarchy*, together with its logic processing functions—the rules encoded in GRNs that determine how regulatory genes cooperate within a given developmental context—and its structural organization in the form of gene regulatory networks are considered the main characteristics of the genomic control system (Peter and Davidson 2015, ix). *Specificity* remains vital, regarding the species specificity of GRNs as well as the specificity of particular genes, as made clear by Davidson by stating that, in these networks, "there are [now] literally scores of genes for which detailed experimental analyses have demonstrated sharply modular cis-regulatory elements, such that given, non-overlapping regions of the genomic DNA each control a specific subcomponent of the overall expression pattern" (Davidson 2006, 33). The theory of the *genomic determination* of the development of the body plan and the *hierarchical organization* of the GRNs also have clear implications for evolution (Erwin and Davidson 2009): They are pre-requisites not only for the explanation of evolutionary changes of body plans, but also of the stability of animal forms—species, genera and so on—over long evolutionary times.

Douglas Erwin and Davidson put forward the hypothesis that the nature of the evolutionary alterations depends on the position of the change within the hierarchy of a GRN. When the genomic regulatory DNA sequences change in evolution, the GRN structure and the developmental process change as well, resulting in great or small changes in the outcome of development (that is, morphology). In the hierarchical structure of developmental GRNs, the portions controlling initial stages of development are at the top of the hierarchy, those controlling intermediate processes are in the middle, and the portions responsible for the detailed functions of cell differentiation at the end of the hierarchy.

This hierarchical organization impacts strongly on evolution, because different portions differ in evolutionary stability or lability. While the regulatory interactions that operate at the initial phases of pattern formation of a developmental part are highly conserved and change only very slowly because most changes are eliminated by natural selection, the parts at the periphery of a GRN have a high lability and minor changes occur frequently. According to Erwin, it is possible to link the appearance of evolutionary novelties, which he understands as individuations of new characters, not just changes in old ones (see Erwin's contribution in this collection), to the structure of developmental gene regulatory networks. This means that small networks of transcription factors within larger GRNs would be responsible for these novel characters. Such subnetworks of developmentally significant genes have been identified for several characters ranking from feathers to heart formation (Erwin 2015).

The concepts of hierarchy and genetic determination employed in systems research on developmental GRNs and evolution make possible an explanation not

only for large and small evolutionary changes, but also, for the first time, for phenotypic stability over long evolutionary times. An example is the stability of the body plans of most animal phyla since around 500 million years. Erwin and Davidson assume that the conserved structures of the GRN portions that control initial developmental stages might be "responsible for the phenotypic stability of animal body plans that has persisted at least since the early Cambrian period i.e. 520 million years ago" (Erwin and Davidson 2009, 142).

6 Summary: the fruitfulness of the concepts of hierarchy, genetic determination and biological specificity

Since the late twentieth century, various approaches have emerged that deal with complex biological phenomena such as development, evolution or diseases, and that aim at overcoming the "reductionism" of classical molecular biology. Holistic approaches, which reject the conventional concepts of a hierarchy of causes and genetic determinism (for early development), instead proposing a collectivity of various causes that act on par, were shown to be unable, in principle, to explain most basic characteristics of development and evolution. In particular they cannot explain the species-specificity of development and the stability of phenotypical forms over long periods of time. Big-data projects for example in genomics produce enormously valuable DNA sequencing and other data. But the functional genomics conducted in these projects is prone to produce questionable results. The exclusion of experimental perturbation and testing from their methods, renders the elucidation of mechanisms nearly impossible, if these projects do not involve cooperation with experimenting scientists. Likewise, different genomic functions, such as evolutionary relevant genomic activities and pure biochemical noise cannot be distinguished.

The most far-reaching approach concerning understanding the mechanisms of development and evolution was achieved in interdisciplinary experimental systems research by molecular embryologists and computer scientists who combine the usage of large-scale genomics data and mathematical modelling with experimentation i.e. perturbation, predictions and testing of causal hypotheses.

This research shows that development can be experimentally accessed at a system-wide level, and mechanisms and causes can be elucidated. It shows, moreover, that the old concepts of hierarchy, determinism and specificity, marginalized or rejected by other research dealing with complexity, are crucially important as guiding principles. Hierarchical gene-regulatory networks are at the center of research that aims at understanding early animal development and its species specificity in a causal-mechanistic way, and that also suggests explanations for unresolved problems of evolution such as the phenotypical stability of certain characteristics over evolutionary times. Research on developmental GRNs has also laid a theoretical ground for a new experimental approach in evolution, namely 'synthetic experimental evolution', which aims at experimentally reproducing evolutionary pathways including the generation of new species (Erwin and Davidson 2009).

References

Aebersold, R., Hood, L. E., & Watts, J. D. (2000). Equipping scientists for the new biology. *Nature Biotechnology, 18,* 359.

Amundson R. ([2005] 2007). *The changing role of the embryo in evolutionary thought*. Cambridge: Cambridge University Press.

Birney, E. (2012). Lessons for big-data projects. *Nature, 489,* 49–51.

Darwin, C. R. (1868). *The variation of animals and plants under domestication*. London: John Murray.

Davidson, E. H. (2006). The regulatory genome: Gene regulatory networks in development and evolution. Burlington, MA: Academic Press.

Davidson, E. H. (2016). Genomics, "Discovery Science", systems biology, and causal explanation. What really works? *Perspectives in Biology and Medicine, 58,* 165–181.

Deichmann, U. (2007a). Empiricism and the discreteness of nature: Ferdinand Cohn (1828–1998), the founder of microbiology. In U. Charpa & U. Deichmann (Eds.), *Jews and sciences in German contexts: Case studies from the 19th and 20th centuries* (pp. 39–50). Tübingen: Mohr-Siebeck.

Deichmann, U. (2007b). "Molecular" versus "colloidal": Controversies in biology and biochemistry, 1900–1940. *Bulletin for the History of Chemistry, 32,* 105–118.

Deichmann, U. (2014). The Concept of the causal role of chromosomes and genes in heredity and development. Opponents from Darwin to Lysenko. *Perspectives in Biology and Medicine, 57,* 57–77.

Deichmann, U. (2016). Epigenetics: The origins and evolution of a fashionable topic. *Developmental Biology, 416,* 249–254.

Doolittle, W. F. (2013). Is junk DNA bunk? A critique of ENCODE. *Proceedings of the National Academy of Sciences USA, 110*(14), 5294–5300.

ENCODE Project Consortium. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature, 489,* 57–74.

Erwin, D. (2015). Novelty and innovation in the history of life. *Current Biology, 25,* R930–R940.

Erwin, D., & Davidson, E. H. (2009). The evolution of hierarchical gene regulatory networks. *Nature Reviews Genetics, 10,* 141–148.

Falk, R. (2014). Biology comes of age. *Israel Journal of Ecology & Evolution, 59,* 186–188. doi:10.1080/15659801.2013.898403.

Florkin, M. (1972). *A history of biochemistry*. Amsterdam: Elsevier.

Godfrey-Smith, P. (2001). On the status and explanatory structure of developmental systems theory. In S. Oyama et al. (Eds.), *Cycles of contingency: Developmental systems and evolution* (pp. 283–297). Cambridge, MA: MIT Press.

Goethe, J. W. V. (1959 [1817]). *Schriften zur Morphologie I.* Stuttgart: Cotta-Verlag.

Goldenfeld, N., & Woese, C. (2007). Biology's next revolution. *Nature, 445,* 369.

Golub, T. (2010). Counterpoint: Data first. *Nature, 464,* 679.

Graur, D., et al. (2015). An evolutionary classification of genomic function. *Genome Biology and Evolution, 7*(3), 642–645.

Griffiths, P., & Gray, R. (1994). Developmental systems and evolutionary explanation. *Journal of Philosophy, 91,* 277–304.

Griffiths, P., & Gray, R. (2001). Darwinism and developmental systems. In S. Oyama et al. (Eds.), *Cycles of contingency* (pp. 195–218). Cambridge, MA: MIT Press.

Griffiths, P., & Gray, R. (2004). The developmental systems perspective: Organism-environment systems as units of development and evolution. In M. Pigliucci & K. Preston (Eds.), *Phenotypic integration: Studying the ecology and evolution of complex phenotypes* (pp. 409–430). Oxford: Oxford University Press.

Jacob, F. (1973). *The logic of life. A history of heredity*. Princeton: Princeton University Press.

Kell, D. B., & Oliver, S. G. (2003). Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. *BioEssays*, *26*, 99–105.

Latour, B. (1993). *We have never been modern*. Cambridge: Harvard University Press.

Latour, B. (1998). From the world of science to the world of research. *Science, 280,* 208–209.

Latour, B. (2004). Politics of nature. *How to bring the sciences into democracy* (C. Porter, Trans.). Cambridge: Harvard University Press.

Loeb, J. (1916). *The organism as a whole from a physicochemical viewpoint*. New York: Putnam's Sons.

Marinacci, B. (Ed.). (1995). *Linus Pauling in his own words: Selections from his writings, speeches and interviews*. New York: Simon and Schuster.
Mazumdar, P. (1995). *Species and specificity* (p. 1995). Cambridge: Cambridge University Press.
modENCODE Consortium. (2010). Identification of functional elements and regulatory circuits by Drosophila modENCODE. *Science, 330*(6012), 1787–1797.
Morange, M. (2001). *The misunderstood gene*. Cambridge: Harvard University Press.
Morange, M. (2014). Genome as a multipurpose structure built by evolution. *Perspectives in Biology and Medicine, 57,* 162–171.
Müller-Wille, S., & Rheinberger, H. J. (2012). *A cultural history of heredity*. Chicago: University of Chicago Press.
Niu, D. K., & Jiang, L. (2013). Can ENCODE tell us how much junk DNA we carry in our genome? *Biochemical and Biophysical Research Communications, 430*(4), 1340–1343.
Oyama, S., Griffiths, P. E., & Gray, R. D. (2001). *Cycles of contingency: Developmental systems and evolution*. Cambridge, MA: MIT Press.
Pavé, A. (2006). Hierarchical organization of biological and ecological systems. In D. Pumain (Ed.), *Hierarchy in natural and social sciences* (pp. 39–70). Dordrecht: Springer.
Peluffo, A. E. (2015). The "genetic program": Behind the genesis of an influential metaphor. *Genetics, 200,* 685–696.
Peter, I., & Davidson, E. H. (2015). *Genomic control process: Development and evolution*. Amsterdam: Academic Press.
Peter, I. S., Faure, E., & Davidson, E. H. (2012). Predictive computation of genomic logic processing functions in embryonic development. *PNAS, 109*(41), 16434–16442.
Pumain, D. (Ed.). (2006). *Hierarchy in natural and social sciences*. Dordrecht: Springer.
Reynolds, A. (2008). Ernst Haeckel and the theory of the cell state: Remarks on the history of a biopolitical metaphor. *History of Science, 46,* 123–152.
Roth, S. (2011). Evolution und Fortschritt. Zum Problem der Höherentwicklung in der organischen Evolution. In T. Schlicht (Ed.), *Zweck und Natur. Historische und systematische Untersuchungen zur Teleologie*. München: Wilhelm Funk Verlag.
Rothenberg, E. (2016). Eric Davidson: Steps to a gene regulatory network for development. *Developmental Biology, 412,* S7–S19.
Verdier, N. (2006). Hierarchy: A short history of a word in Western thought. In D. Pumain (Ed.), *Hierarchy in natural and social sciences* (pp. 13–38). Dordrecht: Springer.
Weinberg, R. (2010). Point: Hypotheses first. *Nature, 464,* 678.
Weismann, A. (1893). *The germ-plasm. A theory of heredity* (English translation in Charles Scribner's Sons). Chapter VI: The formation of germ cells—The continuity of the germ plasm. German original: Weismann, A. 1893. *Das Keimplasma, Eine Theorie der Vererbung*. Jena: G. Fischer.
Wilson, E. B. ([1896] 1928). *The cell in development and heredity*. New York: Macmillan.
Woese, C. (2004). A new biology for a new century. *Microbiology and Molecular Biology Reviews, 68*(2), 173–186.