Ana Lúcia Leitão · Francisco J. Enguita
Editors

Non-coding RNAs and Inter-kingdom Communication

Springer

Non-coding RNAs and Inter-kingdom Communication

Ana Lúcia Leitão · Francisco J. Enguita
Editors

Non-coding RNAs and Inter-kingdom Communication

Springer

Editors
Ana Lúcia Leitão
Universidade Nova de Lisboa
Caparica
Portugal

Francisco J. Enguita
Universidade de Lisboa
Lisbon
Portugal

ISBN 978-3-319-39494-7 ISBN 978-3-319-39496-1 (eBook)
DOI 10.1007/978-3-319-39496-1

Library of Congress Control Number: 2016941096

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Preface

More than 60 years ago, in 1954, James Watson and the Russian Physicist George Gamow founded the "RNA Tie club", a scientific club of selected gentlemen members with the main objective of sharing their ideas and findings not yet mature enough to be published in scientific journals. Brainstorming sessions gave rise to many seminal concepts in cell biology, which were further demonstrated by laboratory experiences. Gamow postulated the concept of a "genetic code" based on the existence of triplets of bases (codons) that will translate the DNA language into protein amino acids. Furthermore, in the last 1950s, another illustrated member of the club, Francis Crick, enunciated the "adaptor theory" based on an intermediate molecule (transfer-RNA) that would be the physical link between nucleic acids and proteins. These ideas were later described in an essential Crick's article, which is considered by many authors as the embryonic core of the concept of the RNA world. The RNA world stands on the idea that RNA molecules were the first precursors of the living cells, being the origin of the evolution of all biological macromolecules. Several arguments clearly support the theory that the primitive cells could be built over processing units based on RNA molecules, including the existence of catalytic RNA molecules. However, for long time the chemical and structural similarities between DNA and RNA, catalogued RNA as a short-life accessory player in cell physiology. Moreover, the paradigm of the reduced stability of RNA vs DNA prevented for a long time the consideration of RNA as a precursor molecule. In consequence, RNA must be placed at the root of the molecular tree of life.

The central dogma of biology holds that genetic information normally flows from DNA to RNA and to proteins. As a consequence it has been generally assumed that genes code for proteins, and that proteins fulfil not only most structural and catalytic but also most regulatory functions in cells. This is essentially true in prokaryotic organisms whose genomes are almost entirely composed of closely packed protein coding sequences. However, this is not the case in higher organisms in which proteomes and their coding sequences occupy only a tiny fraction of the genome. Around 97–98 % of the transcriptional output of the human genome is non-protein coding RNA (ncRNA). RNA is an ideal molecule to regulate biological networks due

its sequence information and structural plasticity. The intrinsic relevance of ncRNAs in the regulation of genomic output has been rapidly unveiling during the last decade. Nevertheless, functional elements in the primary sequence of the majority of ncRNAs that determine their regulatory role remain unknown. The dominance of ncRNAs in the genomic output of the higher organisms suggests that they are not simply occasional transcripts with idiosyncratic functions, but rather that they may constitute an extensive unrecognized regulatory network within higher organisms. The fact that noncoding RNAs constitute the majority of the transcription of the genomes of humans and other complex organisms suggests that a second tier of genetic output and a network of parallel RNA-mediated interactions has evolved in these organisms, which may enable the integration and coordination of sophisticated suites of gene expression required for differentiation and development.

Recent evidences also pointed out to the pivotal role of ncRNAs in the cell-to-cell communication phenomenon, suggesting a transversal role of ncRNA molecules as modulators in organism interaction. There is a common and central biological language represented by functional RNA molecules. Within a complex multi-cellular organism, cells are able to secrete ncRNAs that can travel using the circulating biofluids to reach distant targets where they will exert their regulatory actions functioning as slow genetic hormones. In pathological conditions such as cancer, experimental evidences suggested the use of ncRNAs by tumour cells to prepare their tissue niche before a metastatic colonization. Different organisms can also take advantage of ncRNAs for their functional associations. This is not only the case of the parasitic associations between infectious agents such as viruses, bacteria and fungi that can hijack the host defenses using ncRNAs, but also this phenomenon can be observed in mutualistic and symbiotic interactions.

The present book is a compilation of selected biological topics where ncRNAs are mediators of inter-kingdom communication. The book is divided in two main sections covering the role of ncRNAs in bacterial and viral interaction with different hosts and the modulatory effect of ncRNAs in the interactions between eukaryotic organisms.

This overall project would not be possible without the involvement and the passion of all the contributing authors, but also without the help and collaboration of the anonymous reviewers whose suggestions and criticisms were essential for the consecution of the work that you have now in your hands. Our most sincere acknowledgements to all of them… we are in debt with you.

Caparica, Portugal Ana Lúcia Leitão
Lisbon, Portugal Francisco J. Enguita

Contents

Part I
Non-coding RNAs in Bacterial and Viral Interactions with Different Hosts

Chapter 1
The Roles of MicroRNAs and PiRNAs in Virus-Host Interactions

Anh T. Tran

Abstract MicroRNAs (miRNAs) and Piwi-interacting RNAs (piRNAs) are two groups of small non-coding RNAs with different functional roles. miRNAs are post-transcriptional regulators of gene expression in a plethora of critical processes in multicellular eukaryotes. Therefore, it comes as no surprise that viral pathogens have evolved ways to subvert the miRNA network. It is increasingly evident that miRNAs have functional roles in viral replication as well as their potential employment by host cells to combat viral infection. A number of viruses are now known to encode for miRNAs, predominantly in DNA viruses such as herpesviruses. Although virus-encoded miRNAs have been reported in retroviruses such as HIV-1, their functional significance is under debate. This controversy also extends to RNA viruses and their ability to express miRNAs. Identification of target genes for some of these viral miRNAs suggests they may function in the regulation of lytic and latent viral replication and in restricting antiviral responses. Viruses have also evolved the ability to downregulate or upregulate the expression of specific cellular miRNAs to enhance their replication. I will also briefly review evidence that demonstrate the role of piRNAs in silencing transposable elements to maintain germline genome integrity. This chapter provides an overview of our current understanding of the complex relationship between viruses and cellular miRNA and piRNA machineries.

1.1 miRNA and piRNA Biogenesis and Function

1.1.1 miRNAs

miRNAs are noncoding RNAs ~21–23 nucleotide (nt) in length that post-transcriptionally regulate the expression of a plethora of eukaryotic genes.

A.T. Tran (✉)
Developmental and Stem Cell Biology, Peter Gilgan Centre for Research and Learning, The Hospital for Sick Children, 686 Bay Street, M5G 0A4, Toronto, ON 18-9410, Canada
e-mail: Anh_T.Tran@outlook.com

A.L. Leitão and F.J. Enguita (eds.), *Non-coding RNAs and Inter-kingdom Communication*, DOI 10.1007/978-3-319-39496-1_1

miRNAs are first transcribed as primary transcripts (pri-miRNAs) by RNA polymerase II (RNAP II) (Fig. 1.1). Many pri-miRNA are capped and polyadenylated, and all contain a stem-loop secondary structure. At the nucleus, this secondary structure is recognized by the RNAse type-III Drosha, in association with its co-factor DGCR8, which cleaves the pri-miRNA into its intermediate form (pre-miRNA) of about 70 nt in length with a 2 bp overhang at its 3′ end. The excised hairpin loop is then recognized by the nuclear export factor exportin 5, which facilitates the transport of the pre-miRNA from the nucleus to the cytoplasm where it is processed by another RNAse III Dicer, and its co-factor TRBP, into its mature duplex form. The "guide" strand of the mature miRNA (often the anti-sense strand) is subsequently recruited to the RNA-induced silencing complex (RISC) by its interaction with the Argonaute protein, whereas the "passenger" strand frequently gets degraded (Shukla et al. 2011; Van Wynsberghe et al. 2011; Erson-Bensan 2014). Although the mechanism in allocating the "guide" and "passenger" strand designation is not fully understood, it is thought that once a strand is selected and loaded onto RISC, the other ("star") strand is destroyed. The proteins present in RISC vary between species, but the core proteins include Dicer and the Argonaute protein family. Argonaute proteins are phylogenetically categorized into two clades based on sequence similarity: the Argonaute (Ago) clade

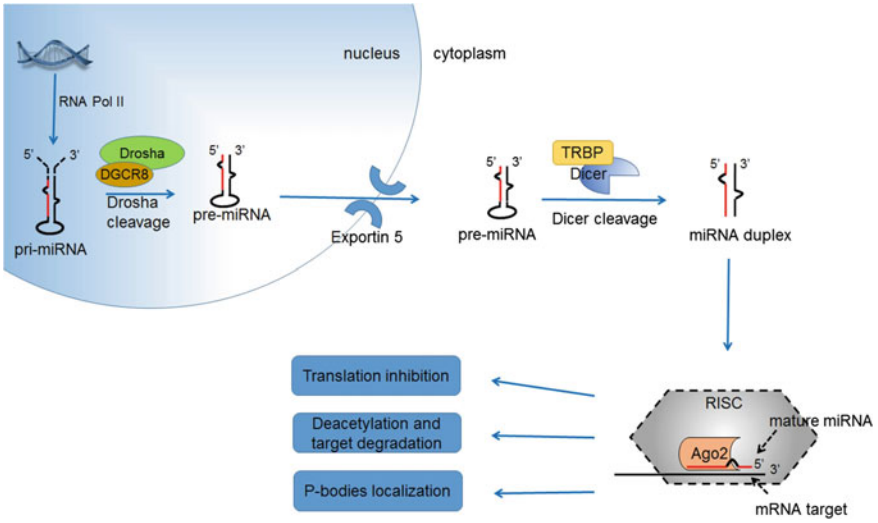


Fig. 1.1 The canonical miRNA biogenesis pathway. Pri-miRNA are transcribed from viral or cellular genomic DNA by RNA polymerase II. The transcript is processed by Drosha and its cofactor DCGR8 to produce pre-miRNA, which gets exported to the cytoplasm by cellular exportin-5 protein. In the cytoplasm, pre-miRNA is further processed into a miRNA duplex by Dicer and its cofactor TRBP. The mature strand of the miRNA duplex (~21–23 nt) is loaded into the miRISC complex. Depending on the degree by which the miRNA "seed" sequence binds to the target mRNA, the resultant mRNA target can be translationally inhibited, deacetylated and degraded, or sequestered to P-bodies for storage

and the P-element induced wimpy testis (Piwi) clade (Carmell et al. 2002). The Ago clade is associated with miRNA and short interfering RNA (siRNA) activities whereas the Piwi proteins are associated with a different sncRNA pathway to be discussed in the next subsection. There are 4 Argonaute proteins in human cells but so far only Argonaute 2 (Ago2) has been reported to contain endonuclease activity (Meister 2013). The miRNA acts as a guide to direct the RISC complex to the target mRNA via base complementarity between the miRNA 5′ seed region (nucleotides at positions 2–8) and the mRNA 3′ untranslated region (3′-UTR) (Bartel 2009; Agarwal et al. 2015). Perfect sequence complimentary between the miRNA and target mRNA may result in target cleavage by the endonucleolytic activity of RISC and significant mRNA destabilization. However, if sequence complimentarity is only partial then RISC binding induces translational inhibition. miRNA regulation of genes was initially known to occur mainly through translational repression, but subsequent observations revealed that miRNAs can also induce mRNA degradation as a result of deadenylation of target mRNAs (Bagga et al. 2005; Krutzfeldt et al. 2005; Valencia-Sanchez et al. 2006). A relatively recent study reported that miRNA mode of regulation may encompass both aspects with initial repression of target mRNAs followed by deadenylation and subsequent degradation (Djuranovic et al. 2012). Moreover, Bartel's group showed that mRNA destabilization may be a major consequence of mRNA repression by miRNA (Eichhorn et al. 2014).

Despite significant advances in our understanding of miRNA activity, the molecular mechanism by which miRNAs suppresses protein production of targeted mRNAs is not completely understood. It has been proposed that miRNA translocation of the targeted mRNA into cytoplasmic processing bodies (P-bodies) leads to induction of translational inhibition, deadenylation, and degradation of the target (Leung and Sharp 2013). P-bodies lack ribosomes and are reported to regulate mRNA turnover and degradation (Leung and Sharp 2013). They also may participate in miRNA regulation of gene expression based on evidence that miRNA-mRNA complexes and components of the miRISC complex such as DGCR8 and Ago localize to these cytoplasmic foci (Leung and Sharp 2013; Baril et al. 2015; Chen and Shyu 2013; Jakymiw et al. 2007; Nishi et al. 2015). It has been proposed that the translocation of miRNA-bound target mRNA complexes to P-bodies promote their catalytic function or for temporal storage, and that this translocation may be mediated by GW182, a component of P-bodies, through its binding to the Argonaute protein in miRISC. Deadenylases that reside within P-bodies can deadenylate targeted mRNAs, which are then decapped and degraded. P-bodies can also function as a temporary storage compartment where targeted mRNAs are held in stasis, spatially removed from the translational machinery (Nilsen 2007).

1.1.2 piRNAs

PiRNAs were first identified as small RNAs that specifically interact with Piwi proteins in mouse and rat germ cells (Aravin et al. 2006; Girard et al. 2006;

Lau et al. 2006). Subsequent studies revealed an extremely complex population of piRNAs that is highly enriched in the germline tissues of most metazoans examined to date (Lim and Kai 2015). Unlike miRNAs, piRNAs are transcribed by RNA polymerase II from intergenic loci called piRNA clusters as long continuous, single-stranded precursor transcripts which are processed by a Dicer-independent mechanism into ~24–31 nt with 2′–*O*–methyl modification sites at the 3′ end (Hirakata and Siomi 2015; Vagin et al. 2006) (Fig. 1.2), which may be targets for the murine methylase HENMT1 (Kirino and Mourelatos 2007). PiRNAs constitute the largest class of noncoding RNAs and have the greatest sequence diversity among known classes of cellular RNAs (Moazed 2009; Lim and Kai 2015). PiRNAs predominantly regulate transposon activities within the genome to preserve normal gametogenesis and reproduction because the expression and transposition of these transposable elements pose a high risk of destabilizing genome integrity. Piwi proteins and piRNAs are conserved in a broad range of metazoans. The absence of Piwi resulted in fertility defects in diverse animal species, indicating the

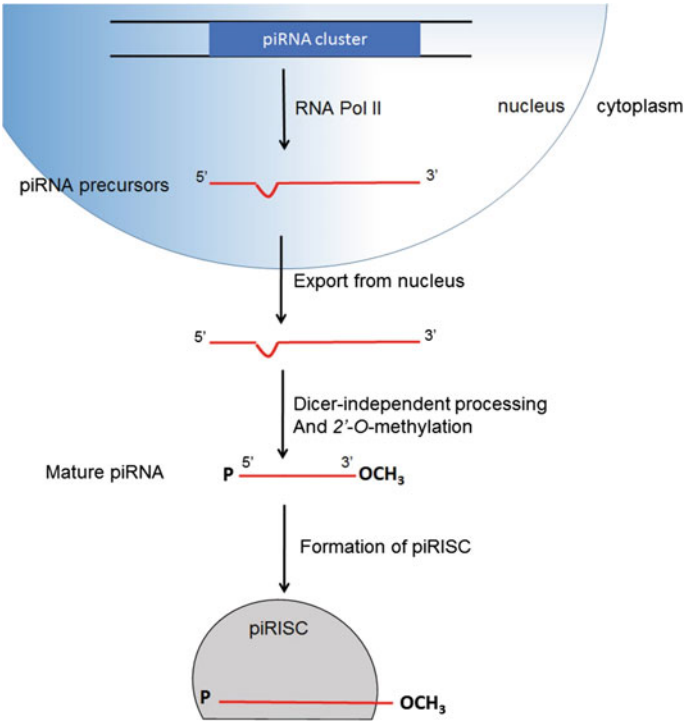


Fig. 1.2 piRNA biogenesis pathway. piRNA precursors are transcribed by RNA polymerase II from piRNA clusters. These precursors have a 5′ cap and 3′ poly-A tail. The precursor piRNAs are exported to the cytoplasm where they are further processed by an unidentified exonuclease at the 3′ end, which is methylated by the murine homolog HENMTI. The mature piRNA associates with Piwi proteins and gets recruited into the piRISC complex

Piwi/piRNA pathway has an important role in maintaining fertility (Carmell et al. 2007; Das et al. 2008; Houwing et al. 2007). Notably, piRNAs possess the ability to distinguish between "self" and "non-self" through a complex mechanism that effectively identifies non-self genes for regulation (Malone and Hannon 2009), the details of which won't be discussed in this chapter.

After assembly, the piRISC complex gets imported into the nucleus where it directs histone 3 lysine 9 (H3K9me3) methylation of target transposon loci to induce a heterochromatin state that transcriptionally silences transposons (Lim and Kai 2015; Le Thomas et al. 2013). In *Drosophila*, the nuclear protein Asterix/DmGTSF1 is required for piRISC to mediate the addition of this silencing histone marker (Ohtani et al. 2013). The precise mechanisms for how piRISC directs the deposition of H3K9me3 at targeted transposon loci have yet to be elucidated.

1.2 Current Methods Used for Viral MiRNA Identification

The very nature and function of miRNAs make them an attractive strategy for viruses to use to manipulate their host environments. Due to the limited size of most viral genomes, the low coding capacity needed to encode the small size of miRNAs, coupled with their non-immunogenic characteristics, makes them an attractive tool to incorporate into a virus' arsenal. Furthermore, a single miRNA has the potential to target numerous host and viral RNAs, which allows a virus to modulate the infection cycle with only limited virus-encoded factors. Viral miRNAs (vmiRNAs) are encoded by many viruses, but the large dsDNA herpesvirus family is the predominant group of viruses that have the most miRNAs characterized within their genomes. The biogenesis of vmiRNAs utilize the same cellular machinery involved in processing cellular miRNAs, and they undergo a similar cascade of steps from the transcription of pri-vmiRNAs transcribed in the nucleus to their subsequent maturation in the cytoplasm. VmiRNAs have been documented to modulate the host environment by targeting either viral or cellular mRNAs to facilitate different facets of the viral lifecycle such as latency.

The most commonly used method to identify vmiRNAs requires the isolation of total small RNAs from infected cells, reverse transcription into cDNA followed by sequencing. Computer algorithms such as TargetScan (Agarwal et al. 2015), miRanda (John et al. 2004; Betel et al. 2008), and RNAhybrid (Rehmsmeier et al. 2004), are also used to predict potential miRNA coding regions which are then verified by direct experimental assays (Bennasser et al. 2004; Pfeffer et al. 2005).

The identification of each miRNA target(s) is not a simple task because a single miRNA can potentially target multiple cellular mRNAs. Bioinformatics computation is used to query for miRNA-seed sequences in the 3′-UTR of potential target mRNAs (Kim and Nam 2006; Rajewsky 2006). Predicting miRNA targets is complicated by the variability in "seed" sequence complementarity between miRNA-mRNA, and that a single miRNA has the potential to regulate the

expression of up to 100 discrete mRNAs (Brennecke et al. 2005). Currently, the use of bioinformatics platforms to identify the entire complement of potential mRNA targets (the 'targetome') of a given miRNA results in long lists that very likely contain many false positives. Nevertheless, successful identifications of miRNA targets have been reported using this method and advances are continually being made in this area.

Messenger RNA microarrays have also been employed to identify targets of a given miRNA by measuring the change in global gene expression in the presence or absence of the miRNA. Differential expression of specific mRNAs in the presence or absence of a given miRNA suggests it may be a potential target of the miRNA, and bioinformatics tools also help predict a target site for the miRNA in the identified mRNA, providing stronger support that this is likely a real target.

Alternatively, target mRNAs can be recovered and sequenced through methods such as mRNA-protein crosslinking followed by immunoprecipitation (CLIP) with a miRISC component such as AGO2 or from P bodies by immunopurification (Easow et al. 2007). CLIP is a powerful tool for the global recovery of miRISC target sites, but the accurate identification of the compliment miRNA responsible for mediating the recruitment of the mRNA to miRISC remains a challenge. A common assay used to confirm that a miRNA targets an identified mRNA is the use of reporter constructs that contain a chimeric transcription with the 3′-UTR from the target mRNA. When the miRNA is overexpressed, its ability to silence the reporter transcript with the target 3′-UTR would appease one criterion supportive of specific targeting.

1.3 Herpesviruses

Herpesviruses are a group of DNA viruses whose infectious lifecycle encompasses both lytic and latent cycles. During latency, viral gene expression is limited to a few specialized genes that maintain the latency state. Life-long persistence in hosts infected with herpesvirus is closely associated with the virus' ability to evade immune detection and establish latency (Feldman and Tibbetts 2015; Frappier 2015). The first virally encoded miRNA identified arose from a cloning experiment in human B cells latently infected with the herpesvirus Epstein-Barr virus (EBV). This initial discovery spurred the prospect that other herpesviruses or large DNA viruses in general might also encode for vmiRNAs (Pfeffer et al. 2004). Indeed, vmiRNAs were recovered from cells infected with herpes simplex 1 (HSV-1), human cytomegalovirus (HCMV), and Kaposi's sarcoma herpesvirus (KSHV) (Feldman and Tibbetts 2015; Kincaid and Sullivan 2012; Pfeffer et al. 2005). Interestingly, most herpesvirus vmiRNAs identified to date are expressed during latency, and have been found to regulate both viral and cellular functions to allow the virus to evade immune detection and persist in the infected host. Nevertheless, certain herpesviruses such as HHV-6, HHV-7, and Varacella Zoster virus (VZV) do not appear to encode miRNAs. A group reported the inability to identify viral

miRNAs in cells latently infected with VZV, but this does not rule out the possibility that there may be miRNAs produced during VZV lytic infection (Umbach et al. 2009). Nevertheless, this observation is particularly interesting given that Varicelloviruses such as Bovine Herpesvirus 1 and Suid Herpesvirus 1 do encode miRNAs (Anselmo et al. 2011; Glazov et al. 2010). This raises the question as to what is different between viruses that do and do not encode miRNAs, the answer of which will be informative in understanding virus miRNA function. Notably, most vmiRNAs encoded by different herpesviruses are not conserved with each other or with host miRNAs, which suggests herpesvirus vmiRNA genes may undergo rapid evolution. However, Poxviruses, which are DNA viruses that replicate in the cytoplasm, do not appear to encode for vmiRNAs (Skalsky and Cullen 2010).

1.3.1 Herpes Simplex Virus

The most studied virally encoded miRNAs among the herpesvirus family are encoded by HSV-1, the prototypical alpha herpesvirus (Table 1.1). miR-H1 is a late gene product initially identified from cells lytically infected with HSV-1 and reported to downregulate an ND10 component alpha-thalassemia/mental retardation syndrome X-linked (ATRX) (Jurak et al. 2012). During HSV-1 latent infections in the sensory ganglia, the latency associated transcript (LAT) is expressed. LAT is encoded antisense to the immediate early (IE) gene *ICP0* in the long terminal repeat end of the unique long genome segment (Roizman and Whitley 2013). AlteRNAtive splicing gives rise to 3 isoforms of the LAT transcript, all of which show different expression patterns. LAT transcripts have not been observed to translate into any peptides, but studies have reported further processing of LAT in HSV-1 latently infected cells to produce 6 functional miRNAs designated as miR-H2, miR-H3, miR-4, miR-H5, miR-H7, and miR-H8 (Umbach et al. 2009). miR-H2 expression leads to a reduction of ICP0 protein level by translational inhibition, as *ICP0* mRNA level is not affected (Umbach et al. 2008). ICP0 is an E3 ubiquitin ligase that allows for a lytic mode of replication at low multiplicity of infection (MOI) (Roizman and Whitley 2013). This protein also facilitates the remodeling of ND10 or PML, which are repressive bodies in the nucleoplasm (Roizman and Whitley 2013). Despite being transcribed antisense to *ICP34.5* transcript, miR-H3 and miR-H4 do not appear to effect *ICP34.5* levels (Umbach et al. 2008). On the other hand, miR-H6, another HSV-1 encoded miRNA, inhibits translation of the viral transactivator for early viral gene expression, ICP4, via imperfect binding to the *ICP4* mRNA. The downregulation of ICP0 and ICP4 by miR-H2 and miR-H6 may inhibit entry into lytic replication and maintenance of an established latent state.

Although there is limited sequence homology between miRNAs expressed by the closely related herpes simplex virus 2 (HSV-2), its miRNAs, also expressed from LAT transcripts, also target ICP0 and ICP34.5 for downregulation, suggesting similar mechanisms for establishing and maintaining latency.

Table 1.1 Selected virus-encoded miRNAs and known mRNA targets

	Viral MIRNA	Viral targets	Cellular targets	Reference
Herpesviruses				
HSV-1	miR-H1	–	Alpha-thalassemia/mental retardation syndrome X-linked (ATRX)	Jurak et al. (2012)
	miR-H2	ICP0	–	Umbach et al. (2008)
	miR-H6	ICP4	–	Umbach et al. (2008)
HSV-2	miR-I, -II	ICP34.5	–	Tang et al. (2008, 2009)
	miR-III	ICP0	–	Tang et al. (2009)
HCMV	miR-UL112-1	IE1	IL-32, MICB	Murphy et al. (2008), Huang et al. (2013), Stern-Ginossar et al. (2007)
	miR-US25-2-3p	–	TIMP3	Esteso et al. (2014)
	miR-UL148	–	RANTES	Kim et al. (2012)
	miR-US4-1	–	ERAP1	Kim et al. (2012)
	miR-UL112-3p	–	TLR2	Landais et al. (2015)
	miR-US25-1	–	Cyclin E2	Grey et al. (2010)
EBV	miR-BART2	BALF5	–	Barth et al. (2008)
	miR-BART16, 17-5p, and 1-5p	–	LMP1	Lo et al. (2007)
	miR-BART5	–	PUMA	Choy et al. (2008)
	miR-BART15-3p	–	BRUCE	Choi et al. (2013)
KSHV	miR-K5, -K9, -K10	–	BCLAF1	Ziegelbauer et al. (2009)
	miR-K9	RTA	–	Bellare and Ganem (2009)
	miR-K1, -K3-3p, -K6-3p	–	THBS1	Narizhneva et al. (2005)
	miR-K1	–	p21	Gottwein and Cullen (2010)
Polyomaviruses				
SV40	miR-M1	T-Antigens	–	Sullivan et al. (2005)
BK	miR-B1	T-Antigens	–	Seo et al. (2008)
JC	miR-J1	T-Antigens	–	Seo et al. (2008)
MC	miR-S1	T-Antigens	–	Seo et al. (2009)
Adenovirus	mivaRI-138	–	TIA-1	Aparicio et al. (2010)

1.3.2 Human Cytomegalovirus

Human cytomegalovirus (HCMV) has the largest genome of the human herpesvirus at 230 kb and is the prototype of beta herpesviruses. To date, experimental evidence has uncovered 14 HCMV miRNAs from lytically infected primary cells (Table 1.1), 3 of which are transcribed from the antisense strand of known ORFs, 5 miRNAs are located in intergenic regions, and 1 is situated within an intron (Pfeffer et al. 2005).

HCMV miR-UL112-1 was reported to inhibit the transactivation of early gene expression during lytic infection by binding to the 3′ UTR of viral IE1 (Murphy et al. 2008). This suggests that miR-UL112-1 may contribute to the establishment and maintenance of latency, but expression of miR-UL112-1 during latent infection remains unknown (Murphy et al. 2008). miR-UL112-1 has recently been implicated in downregulating IL-32, which is critical for both innate and adaptive immune responses (Huang et al. 2013). miR-UL112-1 also has been shown to target the cellular major histocompatibility complex class-I-related chain B (MICB), a cell-surface protein recognized by natural killer (NK) cells, resulting in a decline in MICB protein levels (Stern-Ginossar et al. 2007). Additionally, miR-US25-2-3p was shown to downregulate tissue inhibitors of metalloprotease 3 (TIMP3) expression resulting in an increased shedding of soluble major histocompatibility complex class-I-related chain A (MICA) in patient serum (Esteso et al. 2014). The targeting of MICB by miR-UL112-1 and TIMP3 by miR-US25-2-3p most likely prevents NK cells from killing HCMV-infected cells. Another HCMV-encoded miRNA that targets NK cell activity is miR-UL148, which was determined to bind to the 3' UTR of RANTES, a cellular protein that induces proliferation and activation of NK cells (Kim et al. 2012).

HCMV miR-US4-1 was also shown to indirectly inhibit cytotoxic T lymphocyte (CTL) response by targeting ERAP1 transcripts (Kim et al. 2012). ERAP1 is essential in antigenic peptide production in the ER and mediates the stability of MHC class I-β_2-microglobulin-peptide heterotrimer. Thus, miR-US4-1 targetting of ERAP1 would have profound affects on antigen presentation and the resultant CTL response.

HCMV miRNA activity also affects other signalling cascades such as NF-κB signalling. miR-UL112-3p inhibits NF-κB signalling by targeting Toll-like receptor 2 (TLR2), a major pathogen recognition receptor (PRR) of NF-κB signalling (Landais et al. 2015). There is also miR-US25-1, which has been implicated in the downregulation of cell cycle control protein cyclin E2 (Grey et al. 2010). Thus, HCMV encodes for several miRNAs which have been shown to target both viral and cellular transcripts that affect a number of key signalling pathways.

1.3.3 Epstein–Barr Virus

Epstein-Barr virus (EBV) is a gamma herpesvirus that chronically persists in human B-lymphocytes after primary infection (Young and Rickinson 2004). In vitro

studies have shown that EBV is capable of transforming normal human B-cells into malignant cells, and infection with EBV is associated with malignant diseases such as Hodgkin's lymphoma, endemic Burkitt's lymphoma, and nasopharyngeal carcinoma (Young and Rickinson 2004). During latency, EBV expresses latency-associated genes that include latency-associated membrane proteins (LMP) and Epstein–Barr virus-associated nuclear antigens (EBNAs) (Young and Rickinson 2004). Two regions within the EBV genome, the Bam HI fragment H rightward open reading frame 1 (BHRF1) gene and in the Bam HI-A-region rightward transcript (BART) gene, encode for viral miRNAs (Table 1.1). miR-BHRF1-3 activity may be associated with host immune evasion by downregulating a T-cell attractant CXCL-11 (Xia et al. 2008). miR-BART2 targets *BALF5* transcripts resulting in a reduction of BALF5 viral DNA polymerase, which may serve to stabilize latency in EBV by suppressing lytic cycle viral replication (Barth et al. 2008). The late membrane protein (LMP1) modulates NF-κB signalling and contributes to EBV-mediated transformation. However, several BART miRNAs target the 3' UTR of EBV *LMP1* transcripts, resulting in a reduction of late membrane protein (LMP1) levels thereby affecting the protein's activity in viral-mediated transformation and NF-kB signalling. miR-BART5 has been shown to target apoptosis signalling by downregulating pro-apoptotic protein PUMA (Choy et al. 2008). In contrast, a relatively recent study reported EBV miR-BART15-3p promotes apoptosis by targeting anti-apoptotic protein BRUCE (Choi et al. 2013). The targeting of PUMA implies that miR-BART5 inhibits apoptosis to promote infected cell survival and persistent viral progeny production. However, it remains to be seen what role miR-BART15-3p plays in EBV lifecycle by promoting apoptosis.

1.3.4 Kaposi's Sarcoma-Associated Herpesvirus

The oncogenic Kaposi's sarcoma-associated herpesvirus (KSHV or HHV8) is also a gamma herpesvirus associated with the development of several human malignancies including Kaposi's sarcoma, primary effusion lymphoma (PEL), and Castleman's disease (Ensser and Fleckenstein 2005). KSHV expresses 12 viral miRNAs in latently infected cells from a 5-kb latency-associated region of the viral genome (Table 1.1). Cellular mRNA *BCLAF1* is believed to be involved in KSHV latency, since its downregulation by miR-K5, miR-K9, and mir-K10 resulted in significant reduction in progeny virus recovery after exit from latency (Ziegelbauer et al. 2009). Similarly, miR-K9 targets an important lytic switch protein RTA, which suggests this miRNA may have a role in regulating KSHV latency (Bellare and Ganem 2009). KSHV miR-K1, miR-K3-3p, and miR-K6-3p may be involved in host immune response evasion by targeting thrombospondin-1 (THBS1), which acts as a chemoattractant to recruit monocytes and T cells to sites of infection

(Narizhneva et al. 2005). miR-K1 inhibits cell arrest by downregulating the cyclin-dependent kinase inhibitor p21 in B cells latently infected with KSHV (Gottwein and Cullen 2010).

1.4 Polyomavirus, Adenovirus, and Papillomavirus

Polyomaviruses are small dsDNA viruses that can establish persistent infections as well as immortalize infected cells (White et al. 2013). Current human diseases caused by polyomaviruses are limited to 4 virus strains: BK polyomavirus, which causes kidney and urinary tract diseases; JC polyomavirus, which causes progressive multifocal leukoencephalopathy in immunocompromised individuals; Merkel cell polyomavirus, which causes Merkel cell carcinoma; and trichodysplasia spinulosa-associated polyomavirus, which causes a rare condition of its namesake (White et al. 2013). miRNAs encoded by polyomaviruses were first identified in SV40, a well-studied monkey polyomavirus (Alwine and Khoury 1980; Sullivan et al. 2005). Since then, many polyomaviruses are reported to encode for miRNAs including medically significant BK, JC and Merkel cell polyomaviruses (Seo et al. 2008, 2009). SV40 miRNA appears to be involved in immune evasion. SV40 miRNAs target the viral large T-antigen transcript, leading to cleavage of the transcription and a reduction in both transcript and protein levels (Sullivan et al. 2005). Since the large T-antigen is a target for CTL, miRNA-mediated reduction of large T-antigen levels may allow infected cells to escape immune detection.

Adenoviruses are also small dsDNA viruses that encode for 2 noncoding virus associated RNA (VAI and VAII). VAI confers resistance to cellular interferon-related defenses and contributes to viral replication. Both VAI and VAII have been reported to be processed by Dicer to yield miRNAs that are loaded onto miRISC (Aparicio et al. 2006, 2010; Xu et al. 2007; Sano et al. 2006). However, the functions of these miRNAs remain to be determined.

Initial pursuit to discover miRNAs in papillomaviruses (HPV-31), another DNA virus family, did not yield any positive identifications, suggesting this group of virus may not encode for any miRNAs (Cai et al. 2006). However, most recently, a study reported the identification of viral miRNAs encoded by HPV-16, HPV-38, and HPV-68 using advance sequencing technology SOLiD 4 (Qian et al. 2013). It remains to be seen how these miRNAs affect HPV infection and whether other groups may discover more miRNAs in other HPV genotypes.

1.5 RNA Viruses and Retroviruses

Contrary to DNA viruses, encryption of functional miRNAs in RNA virus genomes remains controversial. Functional miRNAs have not been detected from studies with RNA viruses such as hepatitis C virus (HCV), yellow fever virus, and human

immunodeficiency virus (HIV-1) (Swaminathan et al. 2013). The absence of miRNAs in RNA virus genomes may be due to the fact that most RNA viruses replicate in the cytoplasm, which seclude them from nuclear microprocessing machineries such as Drosha and TRBP. However, even with nuclear replicating RNA viruses such as influenza virus, no functional miRNAs have been identified so far (Umbach et al. 2010; Tycowski et al. 2015; Perez et al. 2010, 2012). It is quite possible that inclusion of non-coding miRNA regions within an RNA virus genome may lead to degradation of the entire viral genome by RISC-mediated mechanisms. Therefore, it has been generally accepted that RNA viruses do not encode vmiRNAs.

However, retroviruses, such as HIV-1, the causative agent of the acquired immunodeficiency syndrome (AIDS), are suspected to potentially encode for vmiRNAs due to the way they replicate within a host cell. HIV-1 viral RNA is reverse transcribed into double-stranded DNA that gets transported into the nucleus where it integrates into the host genome. Thus, HIV-1 should have access to RNAi machinery components present in both the nucleus and cytoplasm, similar to host miRNAs. As such, computational analysis predicts HIV-1 is capable of encoding several miRNA precursors (Bennasser et al. 2004, 2006). The vmiRNA Nef-U3-miR-N367, encoded within *nef*, was reported to target *nef* transcripts that resulted in its degradation and reduction in Nef protein levels, which led to a significant decline in viral replication through various mechanisms (Omoto et al. 2004). The inhibition of HIV-1 replication by Nef-U3-miR-N367 suggests this vmiRNA may play a crucial role in establishing persistent HIV-1 infection. Several other studies have reported the discovery of other vimiRNA encoded within the HIV-1 genome. Kaul et al. reported the identification of a pre-miRNA sequence in the 3′-end of the viral genome called hiv1-miR-H1. They observed this vmiRNA specifically targets a transcription factor that plays an important role negatively regulating cellular apoptosis (Kaul et al. 2009). The authors further noted hiv1-miR-H1 also suppresses other cellular proteins such as c-myc, Par-4, Bcl-2 and Dicer as well as downregulates cellular miR-149, which targets the HIV-1 Vpr protein (Kaul et al. 2009). The authors propose that this vmiRNA activates apoptosis in mononuclear cells by initiating an epigenomic pathway. Interestingly, the Env and Gag-Pol coding regions of different HIV-1 strains are reported to contain miRNA-like sequences with a surprisingly high similarity to human miR-196, miR-30d, miR-30e, miR-374a, and miR-424 (Holland et al. 2013). Additionally, the HIV-1 transactivation RNA (TAR) element has been found to undergo asymmetrical processing by Dicer to produce TAR-miR-5p and -3p (Ouellet et al. 2008). Recent studies using the extremely sensitive SOLiD™ 3 Plus System and a novel, sequence-targeted enrichment strategy identified several hundred small non-coding RNAs (sncRNAs) in infected T lymphocytes and macrophages ranging between 16 and 89 nt in length (Althaus et al. 2012; Schopman et al. 2012). However, controversy remains over whether HIV-1 truly encodes functional vmiRNAs given that some labs failed to detect the expression of any vmiRNA in HIV-1 infected cells

(Lin and Cullen 2007; Whisnant et al. 2013) in addition to the lack of experimental demonstrations on the physiological relevance of these reported vmiRNAs.

1.6 Activities of Cellular MiRNAs and PiRNAs in Viral Infections

1.6.1 Influence of Cellular MiRNAs and PiRNAs on Virus Replication in Mammalian Host

The possibility that cellular miRNAs may modulate virus replication has been the subject of intense investigation. Relatively recent experimental evidence provides support that mammalian sncRNAs regulate virus replication but whether this is physiologically relevant in vivo is still debatable (Yeung et al. 2007). Nevertheless, the impact of cellular miRNA on viral replication was first hinted at when the knockdown of major miRNA-processing enzymes Drosha and Dicer led to more robust replication of viruses such as influenza, vesicular stomatitis virus (VSV), and HIV-1 as a result of reduced mature miRNAs levels in mammalian cells (Matskevich and Moelling 2007; Otsuka et al. 2007; Triboulet et al. 2007). Moreover, the overexpression of virus-encoded RNAi suppressors resulted in a 5- to 10 fold increase in virus replication (de Vries et al. 2008). Antagomirs, chemically modified antisense-oligoribonucleotides, are the current stable method used to inactivate individual cellular miRNAs to assess their functional roles in targeting viruses. (Bennasser et al. 2005; Krutzfeldt et al. 2005). These observations suggest mammalian miRNA/RNAi functionally regulates viral replication.

Studies on miR-122 in HCV replication provide the only current evidence for direct regulation of virus replication by cellular miRNA (Jopling et al. 2005). HCV establishes persistent infections in the liver that may lead to liver cirrhosis and hepatocellular carcinoma. This cellular miRNA is highly expressed in the liver where it has been shown to positively regulate HCV replication in Huh7 human liver cells (Jopling et al. 2005) by directly interacting with two adjacent sites in the 5′ UTR of the viral RNA (Jopling et al. 2005). The mechanism that underlies miR-122`s affect on HCV replication remains unknown but recent studies begin to shed light on this process. miR-122 binding to sites located upstream of the HCV IRES, which directs translation of the positive-sense viral RNA genome, resulted in moderate upregulation of viral protein translation (Henke et al. 2008). However, another study has shown that miR-122 may act at another, yet undefined, stage of viral replication (Jangra et al. 2010).

Many cellular miRNAs have been found to indirectly affect HIV-1 replication by targeting factors known to be important to HIV infection called HIV dependency factors (HDF). Evidence for the anti-viral nature of cellular miRNA in HIV-1 infection comes from a study carried out by Triboulet et al. where they knocked down Drosha and Dicer, two important miRNA processing proteins, and found

viral replication was greatly enhanced in peripheral blood mononuclear cells (PBMCs) from HIV-1-infected patients and in latently infected cells (Triboulet et al. 2007). The same group further showed that viral mRNA co-localized with components of the miRISC complex. Interestingly, they reported virus reactivation in PBMCs isolated from patients undergoing cART when they knocked down cellular factors involved in miRNA-mediated silencing. These factors were also shown to inhibit viral gene expression by interfering with viral mRNA association with polysomes (Chable-Bessia et al. 2009). Also, a cluster of human miRNAs that include miR-28, miR-125b, miR-150, miR-223 and miR-382 was shown to target the 3′ end of HIV-1 mRNAs for repression (Huang et al. 2007). These observations highlight the important role miRNAs have in regulating HIV-1 infection.

Knockdown of Dicer greatly enhances influenza A virus (IAV) replication, suggesting cellular RNAi also has antiviral effects on influenza A virus replication (Matskevich and Moelling 2007). Song et al. used a 3′ UTR reporter construct to show that host miR-323, miR-491, and miR-654 inhibited influenza strain A/WSN/33 H1N1 replication through the binding of the PB1 transcript, resulting in degradation of the PB1 mRNA (Song et al. 2010). PB1 forms part of the IAV viral polymerase complex necessary for viral genome transcription and replication (Kobayashi et al. 1996). Several studies have reported modulation of host miRNA expression profiles upon viral infection, where infection of specific IAV subtypes resulted in different expression profiles of cellular miRNAs (Loveday et al. 2012; Tambyah et al. 2013; Li et al. 2010). This raises the question of whether the differing pathogenicities exhibited by different IAV subtypes may be associated with the differential expressions of specific cellular miRNAs. One recent study suggests that miR-24 may modulate the highly pathogenic H5N1 virus by down-regulating the furin secretory pathway. This pathway is exploited by the virus to proteolytically cleave its haemagglutinin precursor to active forms that allow viral entry in host cells via viral-host membrane fusion (Loveday et al. 2015). Surprisingly, miR-24 was reported to have little effect on low-pathogenic 2009 pandemic H1N1 infection, suggesting its activity may be strain-specific.

The cellular RNAi arsenal is not limited to miRNAs but include short interfering RNAs (siRNAs) and a separate class of sncRNAs called piRNAs. In the nematode model organism *Caenorhabditis elegans*, RNAi has been demonstrated to be the main antiviral response to Flock House virus and the natural pathogen Orsay virus. This antiviral response involves the PIWI family of genes such as *rde-1*, which encodes an Argonaute protein, and *rde-4*, which encodes a dsRNA-binding protein (Felix et al. 2011). While there are no homologues of these proteins in mammalian cells, there are evidence that piRNAs or repeat-associated small interfering RNAs (rasiRNAs) functionally suppress mammalian endogenous retroviruses in germ and somatic cells (Carmell et al. 2007; Watanabe et al. 2006). It has also been shown that mouse ES cells, and possibly a subset of somatic stem cells, retain the ability to generate small RNA species from exogenous dsRNA and viral RNA, which are also implicated in repressing retrotransposon activities (Calabrese et al. 2007). However, the importance of these small RNA species as antiviral agents against human viral infections remains to be fully addressed.

1.6.2 Interferon-Mediated Antiviral Activity Via MiRNAs

The interferon (IFN) signalling pathway is an important part of cellular antiviral immunity. This pathway is activated upon the production of long dsRNAs within a cell. Interestingly, there is evidence to suggest the IFN pathway partially overlaps with the RNAi pathway. It has been noted that proteins such as PACT and TRBP, which function in the IFN antiviral network, also has a role in miRNA processing and function (Chendrimada et al. 2005; Haase et al. 2005; Kok et al. 2007; Lee et al. 2006). Additionally, RNA-binding proteins are capable of suppressing both the IFN effector protein kinase R (PKR) (Cai et al. 2000; Li et al. 2004; McMillan et al. 1995) and RNAi (Bennasser et al. 2005; Haasnoot et al. 2007). Most notably was the report that interferon beta (IFN-β) can regulate the expression of several cellular miRNAs, 8 of which are predicted to have binding sites within the HCV genomic RNA. In support of this, the introduction of synthetic miRNA-mimics resembling these IFN-β-induced miRNAs recapitulated the antiviral effects of IFN-β on HCV infection. Additionally, IFN-β's antiviral effects against HCV were attenuated when antagomirs to these IFN-β-induced miRNAs were introduced into the infected cells (Pedersen et al. 2007). This evidence support the potential overlap of miRNAs and the IFN network in mammalian cells and possible IFN-dependent and IFN-independent defenses to viral infection. There also is evidence in a number of viruses that show vmiRNAs modulate the innate immune system by directly targeting the downregulation of immune factors, which affect processes such as the recruitment of effector cells of the immune system (Cullen 2013).

1.6.3 Viral Counter-Responses to Cellular RNAi Restriction

Viruses appear capable of countering the cell's RNAi restriction by actively suppressing RNAi activation through viral dsRNA-binding proteins that sequester and neutralize the antiviral RNAi activities. An example of this may be influenza A virus's NS1 protein which has been suggested to suppress host RNAi activity possibly through its dsRNA-binding capability (Haasnoot et al. 2007). HIV-1 counteracts this host defense by developing strategies to modulate cellular miRNA expression and interfere with the overall biogenesis of miRNAs. Viral protein Tat inhibits Dicer activity through its physical interaction with the helicase domain of Dicer and its binding to host mRNAs. However, the authors did not determine whether this was a direct protein-protein interaction (Chable-Bessia et al. 2009; Bennasser et al. 2005). Contradicting observations have been made in whether the SRS/RSS function of Tat alters miRNA expression profiles. Some have reported this function of Tat does result in altered miRNA expression in T cell lines and primary PBMCs, whereas others have shown that it does not affect miRNA expression in PBMCs (Mishra et al. 2012). AlteRNAtively, HCV evolved to adapt

to miRNA-restriction in a way that benefitted the virus' replication, for example the enhancement of HCV replication by liver-specific miR-122 (Jopling et al. 2005).

1.7 Concluding Remarks and Future Perspectives

Given the relatively recent discovery of sncRNAs, tremendous advances have been made in the discovery and characterization of these molecules and their effects on viral replication. However, the more we discover the clearer it becomes that viruses of diverse origin have evolved to exploit the host miRNA pathway in vast numbers of ways to effectively regulate infection and the host response. Nevertheless, several prominent issues remain to be addressed. The physiological importance of viral miRNAs in many groups of viruses aside from herpesviruses and polyomaviruses is still unclear. Likewise, it is very likely that the present list of targets identified for herpesvirus miRNAs represent only a partial number of potential targets; that many more targets remain to be discovered. Notably, the relationship between herpesvirus miRNA expression and viral latency and immune evasion remains uncertain. Although it is known that miRNAs from different herpesviruses target common pathways, it remains unclear whether these miRNAs also share a common function.

Another major area to expand future research is cellular miRNA regulation of viral infection. It appears that a broad spectrum of viruses employ complex mechanisms to regulate cellular miRNA expression to mold the host environment into something more permissible and advantageous for viral replication. The mechanistic details specific viruses use to modulate cellular miRNA expressions are largely unknown. Viral proteins may directly modulate host miRNA expression or indirectly regulate it due to a general host response to the infection. There have been several reports of direct cellular miRNA binding to viral RNA and negatively regulating the virus but most of these studies were performed in tissue culture or in organisms other than the natural host, which brings into question the physiological relevance of these interactions. It would be rational to imagine that evolutionary pressure would preclude viruses retaining cellular miRNA target sites that negatively regulate its replication, but if such sites were beneficial to the pathogen, then it seems possible the viruses might want to retain them in their genomes.

Although tremendous advances have been made to understand the relationship between miRNAs and viral replication, much remains to be discovered. This poses exciting potentials in terms of scientific discovery as well as prospects for novel therapeutic strategies to control viral infections.

Acknowledgments I would like to thank Dr. Brent Derry and Dr. Kevin Coombs for their reviews of the manuscript. ATT is supported by the Garron Family Cancer Centre and SickKids Research Training Centre Fellowship.

References

Agarwal V, Bell GW, Nam JW, Bartel DP (2015) Predicting effective microRNA target sites in mammalian mRNAs. eLife 4. doi:10.7554/eLife.05005

Althaus CF, Vongrad V, Niederost B, Joos B, Di Giallonardo F, Rieder P, Pavlovic J, Trkola A, Gunthard HF, Metzner KJ, Fischer M (2012) Tailored enrichment strategy detects low abundant small noncoding RNAs in HIV-1 infected cells. Retrovirology 9:27. doi:10.1186/1742-4690-9-27

Alwine JC, Khoury G (1980) Simian virus 40-associated small RNA: mapping on the simian virus 40 genome and characterization of its synthesis. J Virol 36(3):701–708

Anselmo A, Flori L, Jaffrezic F, Rutigliano T, Cecere M, Cortes-Perez N, Lefevre F, Rogel-Gaillard C, Giuffra E (2011) Co-expression of host and viral microRNAs in porcine dendritic cells infected by the pseudorabies virus. PLoS ONE 6(3):e17374. doi:10.1371/journal.pone.0017374

Aparicio O, Razquin N, Zaratiegui M, Narvaiza I, Fortes P (2006) Adenovirus virus-associated RNA is processed to functional interfering RNAs involved in virus production. J Virol 80 (3):1376–1384. doi:10.1128/jvi.80.3.1376-1384.2006

Aparicio O, Carnero E, Abad X, Razquin N, Guruceaga E, Segura V, Fortes P (2010) Adenovirus VA RNA-derived miRNAs target cellular genes involved in cell growth, gene expression and DNA repair. Nucleic Acids Res 38(3):750–763. doi:10.1093/nar/gkp1028

Aravin A, Gaidatzis D, Pfeffer S, Lagos-Quintana M, Landgraf P, Iovino N, Morris P, Brownstein MJ, Kuramochi-Miyagawa S, Nakano T, Chien M, Russo JJ, Ju J, Sheridan R, Sander C, Zavolan M, Tuschl T (2006) A novel class of small RNAs bind to MILI protein in mouse testes. Nature 442(7099):203–207. doi:10.1038/nature04916

Bagga S, Bracht J, Hunter S, Massirer K, Holtz J, Eachus R, Pasquinelli AE (2005) Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. Cell 122(4):553–563. doi:10.1016/j.cell.2005.07.031

Baril P, Ezzine S, Pichon C (2015) Monitoring the spatiotemporal activities of miRNAs in small animal models using molecular imaging modalities. Int J Mol Sci 16(3):4947–4972. doi:10.3390/ijms16034947

Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. Cell 136(2):215–233. doi:10.1016/j.cell.2009.01.002

Barth S, Pfuhl T, Mamiani A, Ehses C, Roemer K, Kremmer E, Jaker C, Hock J, Meister G, Grasser FA (2008) Epstein-Barr virus-encoded microRNA miR-BART2 down-regulates the viral DNA polymerase BALF5. Nucleic Acids Res 36(2):666–675. doi:10.1093/nar/gkm1080

Bellare P, Ganem D (2009) Regulation of KSHV lytic switch protein expression by a virus-encoded microRNA: an evolutionary adaptation that fine-tunes lytic reactivation. Cell Host Microbe 6(6):570–575. doi:10.1016/j.chom.2009.11.008

Bennasser Y, Le SY, Yeung ML, Jeang KT (2004) HIV-1 encoded candidate micro-RNAs and their cellular targets. Retrovirology 1:43. doi:10.1186/1742-4690-1-43

Bennasser Y, Le SY, Benkirane M, Jeang KT (2005) Evidence that HIV-1 encodes an siRNA and a suppressor of RNA silencing. Immunity 22(5):607–619. doi:10.1016/j.immuni.2005.03.010

Bennasser Y, Le SY, Yeung ML, Jeang KT (2006) MicroRNAs in human immunodeficiency virus-1 infection. Meth Mol Biol 342:241–253. doi:10.1385/1-59745-123-1:241

Betel D, Wilson M, Gabow A, Marks DS, Sander C (2008) The microRNA.org resource: targets and expression. Nucleic Acids Res 36 (Database issue):D149–153. doi:10.1093/nar/gkm995

Brennecke J, Stark A, Russell RB, Cohen SM (2005) Principles of microRNA-target recognition. PLoS Biol 3(3):e85. doi:10.1371/journal.pbio.0030085

Cai R, Carpick B, Chun RF, Jeang KT, Williams BR (2000) HIV-I TAT inhibits PKR activity by both RNA-dependent and RNA-independent mechanisms. Arch Biochem Biophys 373 (2):361–367. doi:10.1006/abbi.1999.1583

Cai X, Li G, Laimins LA, Cullen BR (2006) Human papillomavirus genotype 31 does not express detectable microRNA levels during latent or productive virus replication. J Virol 80 (21):10890–10893. doi:10.1128/jvi.01175-06

Calabrese JM, Seila AC, Yeo GW, Sharp PA (2007) RNA sequence analysis defines Dicer's role in mouse embryonic stem cells. Proc Natl Acad Sci USA 104(46):18097–18102. doi:10.1073/pnas.0709193104

Carmell MA, Xuan Z, Zhang MQ, Hannon GJ (2002) The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. Genes Dev 16 (21):2733–2742. doi:10.1101/gad.1026102

Carmell MA, Girard A, van de Kant HJ, Bourc'his D, Bestor TH, de Rooij DG, Hannon GJ (2007) MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. Dev Cell 12(4):503–514. doi:10.1016/j.devcel.2007.03.001

Chable-Bessia C, Meziane O, Latreille D, Triboulet R, Zamborlini A, Wagschal A, Jacquet JM, Reynes J, Levy Y, Saib A, Bennasser Y, Benkirane M (2009) Suppression of HIV-1 replication by microRNA effectors. Retrovirology 6:26. doi:10.1186/1742-4690-6-26

Chen CY, Shyu AB (2013) Deadenylation and P-bodies. Adv Exp Med Biol 768:183–195. doi:10.1007/978-1-4614-5107-5_11

Chendrimada TP, Gregory RI, Kumaraswamy E, Norman J, Cooch N, Nishikura K, Shiekhattar R (2005) TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature 436(7051):740–744. doi:10.1038/nature03868

Choi H, Lee H, Kim SR, Gho YS, Lee SK (2013) Epstein-Barr virus-encoded microRNA BART15-3p promotes cell apoptosis partially by targeting BRUCE. J Virol 87(14):8135–8144. doi:10.1128/jvi.03159-12

Choy EY, Siu KL, Kok KH, Lung RW, Tsang CM, To KF, Kwong DL, Tsao SW, Jin DY (2008) An Epstein-Barr virus-encoded microRNA targets PUMA to promote host cell survival. J Exp Med 205(11):2551–2560. doi:10.1084/jem.20072581

Cullen BR (2013) MicroRNAs as mediators of viral evasion of the immune system. Nature Immunol 14(3):205–210. doi:10.1038/ni.2537

Das PP, Bagijn MP, Goldstein LD, Woolford JR, Lehrbach NJ, Sapetschnig A, Buhecha HR, Gilchrist MJ, Howe KL, Stark R, Matthews N, Berezikov E, Ketting RF, Tavare S, Miska EA (2008) Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the Caenorhabditis elegans germline. Mol Cell 31(1):79–90. doi:10.1016/j.molcel.2008.06.003

de Vries W, Haasnoot J, van der Velden J, van Montfort T, Zorgdrager F, Paxton W, Cornelissen M, van Kuppeveld F, de Haan P, Berkhout B (2008) Increased virus replication in mammalian cells by blocking intracellular innate defense responses. Gene Therap 15(7):545–552. doi:10.1038/gt.2008.12

Djuranovic S, Nahvi A, Green R (2012) miRNA-mediated gene silencing by translational repression followed by mRNA deadenylation and decay. Science 336(6078):237–240. doi:10.1126/science.1215691

Easow G, Teleman AA, Cohen SM (2007) Isolation of microRNA targets by miRNP immunopurification. RNA 13(8):1198–1204. doi:10.1261/rna.563707

Eichhorn SW, Guo H, McGeary SE, Rodriguez-Mias RA, Shin C, Baek D, Hsu SH, Ghoshal K, Villen J, Bartel DP (2014) mRNA destabilization is the dominant effect of mammalian microRNAs by the time substantial repression ensues. Mol Cell 56(1):104–115. doi:10.1016/j.molcel.2014.08.028

Ensser A, Fleckenstein B (2005) T-cell transformation and oncogenesis by gamma2-herpesviruses. Adv Cancer Res 93:91–128. doi:10.1016/s0065-230x(05)93003-0

Erson-Bensan AE (2014) Introduction to microRNAs in biological systems. Meth Mol Biol 1107:1–14. doi:10.1007/978-1-62703-748-8_1

Esteso G, Luzon E, Sarmiento E, Gomez-Caro R, Steinle A, Murphy G, Carbone J, Vales-Gomez M, Reyburn HT (2014) Altered microRNA expression after infection with human cytomegalovirus leads to TIMP3 downregulation and increased shedding of metalloprotease substrates, including MICA. J Immunol 193(3):1344–1352. doi:10.4049/jimmunol.1303441

Feldman ER, Tibbetts SA (2015) Emerging roles of herpesvirus microRNAs during in vivo infection and pathogenesis. Curr Pathobiol Rep 3(3):209–217. doi:10.1007/s40139-015-0085-z

Felix MA, Ashe A, Piffaretti J, Wu G, Nuez I, Belicard T, Jiang Y, Zhao G, Franz CJ, Goldstein LD, Sanroman M, Miska EA, Wang D (2011) Natural and experimental infection of *Caenorhabditis* nematodes by novel viruses related to nodaviruses. PLoS Biol 9(1):e1000586. doi:10.1371/journal.pbio.1000586

Frappier L (2015) Regulation of herpesvirus reactivation by host microRNAs. J Virol 89(5):2456–2458. doi:10.1128/jvi.03413-14

Girard A, Sachidanandam R, Hannon GJ, Carmell MA (2006) A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature 442(7099):199–202. doi:10.1038/nature04917

Glazov EA, Horwood PF, Assavalapsakul W, Kongsuwan K, Mitchell RW, Mitter N, Mahony TJ (2010) Characterization of microRNAs encoded by the bovine herpesvirus 1 genome. J Gen Virol 91(Pt 1):32–41. doi:10.1099/vir.0.014290-0

Gottwein E, Cullen BR (2010) A human herpesvirus microRNA inhibits p21 expression and attenuates p21-mediated cell cycle arrest. J Virol 84(10):5229–5237. doi:10.1128/jvi.00202-10

Grey F, Tirabassi R, Meyers H, Wu G, McWeeney S, Hook L, Nelson JA (2010) A viral microRNA down-regulates multiple cell cycle genes through mRNA 5'UTRs. PLoS Pathog 6 (6):e1000967. doi:10.1371/journal.ppat.1000967

Haase AD, Jaskiewicz L, Zhang H, Laine S, Sack R, Gatignol A, Filipowicz W (2005) TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. EMBO Rep 6(10):961–967. doi:10.1038/sj.embor.7400509

Haasnoot J, de Vries W, Geutjes EJ, Prins M, de Haan P, Berkhout B (2007) The Ebola virus VP35 protein is a suppressor of RNA silencing. PLoS Pathog 3(6):e86. doi:10.1371/journal.ppat.0030086

Henke JI, Goergen D, Zheng J, Song Y, Schuttler CG, Fehr C, Junemann C, Niepmann M (2008) microRNA-122 stimulates translation of hepatitis C virus RNA. EMBO J 27(24):3300–3310. doi:10.1038/emboj.2008.244

Hirakata S, Siomi MC (2015) piRNA biogenesis in the germline: from transcription of piRNA genomic sources to piRNA maturation. Biochim Biophys Acta. doi:10.1016/j.bbagrm.2015.09.002

Holland B, Wong J, Li M, Rasheed S (2013) Identification of human microRNA-like sequences embedded within the protein-encoding genes of the human immunodeficiency virus. PLoS ONE 8(3):e58586. doi:10.1371/journal.pone.0058586

Houwing S, Kamminga LM, Berezikov E, Cronembold D, Girard A, van den Elst H, Filippov DV, Blaser H, Raz E, Moens CB, Plasterk RH, Hannon GJ, Draper BW, Ketting RF (2007) A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. Cell 129 (1):69–82. doi:10.1016/j.cell.2007.03.026

Huang J, Wang F, Argyris E, Chen K, Liang Z, Tian H, Huang W, Squires K, Verlinghieri G, Zhang H (2007) Cellular microRNAs contribute to HIV-1 latency in resting primary CD4+T lymphocytes. Nature Med 13(10):1241–1247. doi:10.1038/nm1639

Huang Y, Qi Y, Ma Y, He R, Ji Y, Sun Z, Ruan Q (2013) The expression of interleukin-32 is activated by human cytomegalovirus infection and down regulated by hcmv-miR-UL112-1. Virol J 10:51. doi:10.1186/1743-422x-10-51

Jakymiw A, Pauley KM, Li S, Ikeda K, Lian S, Eystathioy T, Satoh M, Fritzler MJ, Chan EK (2007) The role of GW/P-bodies in RNA processing and silencing. J Cell Sci 120(Pt 8):1317–1323. doi:10.1242/jcs.03429

Jangra RK, Yi M, Lemon SM (2010) Regulation of hepatitis C virus translation and infectious virus production by the microRNA miR-122. J Virol 84(13):6615–6625. doi:10.1128/jvi.00417-10

John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS (2004) Human MicroRNA targets. PLoS Biol 2(11):e363. doi:10.1371/journal.pbio.0020363

Jopling CL, Yi M, Lancaster AM, Lemon SM, Sarnow P (2005) Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. Science 309(5740):1577–1581. doi:10.1126/science.1113329

Jurak I, Silverstein LB, Sharma M, Coen DM (2012) Herpes simplex virus is equipped with RNA- and protein-based mechanisms to repress expression of ATRX, an effector of intrinsic immunity. J Virol 86(18):10093–10102. doi:10.1128/jvi.00930-12

Kaul D, Ahlawat A, Gupta SD (2009) HIV-1 genome-encoded hiv1-mir-H1 impairs cellular responses to infection. Mol Cell Biochem 323(1–2):143–148. doi:10.1007/s11010-008-9973-4

Kim VN, Nam JW (2006) Genomics of microRNA. Trends Genet: TIG 22(3):165–173. doi:10.1016/j.tig.2006.01.003

Kim Y, Lee S, Kim S, Kim D, Ahn JH, Ahn K (2012) Human cytomegalovirus clinical strain-specific microRNA miR-UL148D targets the human chemokine RANTES during infection. PLoS Pathog 8(3):e1002577. doi:10.1371/journal.ppat.1002577

Kincaid RP, Sullivan CS (2012) Virus-encoded microRNAs: an overview and a look to the future. PLoS Pathog 8(12):e1003018. doi:10.1371/journal.ppat.1003018

Kirino Y, Mourelatos Z (2007) The mouse homolog of HEN1 is a potential methylase for Piwi-interacting RNAs. RNA 13(9):1397–1401. doi:10.1261/rna.659307

Kobayashi M, Toyoda T, Ishihama A (1996) Influenza virus PB1 protein is the minimal and essential subunit of RNA polymerase. Arch Virol 141(3–4):525–539

Kok KH, Ng MH, Ching YP, Jin DY (2007) Human TRBP and PACT directly interact with each other and associate with dicer to facilitate the production of small interfering RNA. J Biol Chem 282(24):17649–17657. doi:10.1074/jbc.M611768200

Krutzfeldt J, Rajewsky N, Braich R, Rajeev KG, Tuschl T, Manoharan M, Stoffel M (2005) Silencing of microRNAs in vivo with 'antagomirs'. Nature 438(7068):685–689. doi:10.1038/nature04303

Landais I, Pelton C, Streblow D, DeFilippis V, McWeeney S, Nelson JA (2015) Human cytomegalovirus miR-UL112-3p targets TLR2 and modulates the TLR2/IRAK1/NFkappaB signaling pathway. PLoS Pathog 11(5):e1004881. doi:10.1371/journal.ppat.1004881

Lau NC, Seto AG, Kim J, Kuramochi-Miyagawa S, Nakano T, Bartel DP, Kingston RE (2006) Characterization of the piRNA complex from rat testes. Science 313(5785):363–367. doi:10.1126/science.1130164

Le Thomas A, Rogers AK, Webster A, Marinov GK, Liao SE, Perkins EM, Hur JK, Aravin AA, Toth KF (2013) Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. Genes Dev 27(4):390–399. doi:10.1101/gad.209841.112

Lee Y, Hur I, Park SY, Kim YK, Suh MR, Kim VN (2006) The role of PACT in the RNA silencing pathway. EMBO J 25(3):522–532. doi:10.1038/sj.emboj.7600942

Leung AK, Sharp PA (2013) Quantifying Argonaute proteins in and out of GW/P-bodies: implications in microRNA activities. Adv Exp Med Biol 768:165–182. doi:10.1007/978-1-4614-5107-5_10

Li WX, Li H, Lu R, Li F, Dus M, Atkinson P, Brydon EW, Johnson KL, Garcia-Sastre A, Ball LA, Palese P, Ding SW (2004) Interferon antagonist proteins of influenza and vaccinia viruses are suppressors of RNA silencing. Proc Natl Acad Sci USA 101(5):1350–1355. doi:10.1073/pnas.0308308100

Li Y, Chan EY, Li J, Ni C, Peng X, Rosenzweig E, Tumpey TM, Katze MG (2010) MicroRNA expression and virulence in pandemic influenza virus-infected mice. J Virol 84(6):3023–3032. doi:10.1128/jvi.02203-09

Lim RS, Kai T (2015) A piece of the pi(e): The diverse roles of animal piRNAs and their PIWI partners. Sem Cell Dev Biol. doi:10.1016/j.semcdb.2015.10.025

Lin J, Cullen BR (2007) Analysis of the interaction of primate retroviruses with the human RNA interference machinery. J Virol 81(22):12218–12226. doi:10.1128/jvi.01390-07

Lo AK, To KF, Lo KW, Lung RW, Hui JW, Liao G, Hayward SD (2007) Modulation of LMP1 protein expression by EBV-encoded microRNAs. Proc Natl Acad Sci USA 104 (41):16164–16169. doi:10.1073/pnas.0702896104

Loveday EK, Svinti V, Diederich S, Pasick J, Jean F (2012) Temporal- and strain-specific host microRNA molecular signatures associated with swine-origin H1N1 and avian-origin H7N7 influenza A virus infection. J Virol 86(11):6109–6122. doi:10.1128/jvi.06892-11

Loveday EK, Diederich S, Pasick J, Jean F (2015) Human microRNA-24 modulates highly pathogenic avian-origin H5N1 influenza A virus infection in A549 cells by targeting secretory pathway furin. J Gen Virol 96(Pt 1):30–39. doi:10.1099/vir.0.068585-0

Malone CD, Hannon GJ (2009) Small RNAs as guardians of the genome. Cell 136(4):656–668. doi:10.1016/j.cell.2009.01.045

Matskevich AA, Moelling K (2007) Dicer is involved in protection against influenza A virus infection. J Gen Virol 88(Pt 10):2627–2635. doi:10.1099/vir.0.83103-0

McMillan NA, Chun RF, Siderovski DP, Galabru J, Toone WM, Samuel CE, Mak TW, Hovanessian AG, Jeang KT, Williams BR (1995) HIV-1 Tat directly interacts with the interferon-induced, double-stranded RNA-dependent kinase, PKR. Virology 213(2):413–424. doi:10.1006/viro.1995.0014

Meister G (2013) Argonaute proteins: functional insights and emerging roles. Nature Rev Genet 14 (7):447–459. doi:10.1038/nrg3462

Mishra R, Chhatbar C, Singh SK (2012) HIV-1 Tat C-mediated regulation of tumor necrosis factor receptor-associated factor-3 by microRNA 32 in human microglia. J Neuroinflam 9:131. doi:10.1186/1742-2094-9-131

Moazed D (2009) Small RNAs in transcriptional gene silencing and genome defence. Nature 457 (7228):413–420. doi:10.1038/nature07756

Murphy E, Vanicek J, Robins H, Shenk T, Levine AJ (2008) Suppression of immediate-early viral gene expression by herpesvirus-coded microRNAs: implications for latency. Proc Natl Acad Sci USA 105(14):5453–5458. doi:10.1073/pnas.0711910105

Narizhneva NV, Razorenova OV, Podrez EA, Chen J, Chandrasekharan UM, DiCorleto PE, Plow EF, Topol EJ, Byzova TV (2005) Thrombospondin-1 up-regulates expression of cell adhesion molecules and promotes monocyte binding to endothelium. FASEB J 19(9):1158–1160. doi:10.1096/fj.04-3310fje

Nilsen TW (2007) Mechanisms of microRNA-mediated gene regulation in animal cells. Trends Genet 23(5):243–249. doi:10.1016/j.tig.2007.02.011

Nishi K, Takahashi T, Suzawa M, Miyakawa T, Nagasawa T, Ming Y, Tanokura M, Ui-Tei K (2015) Control of the localization and function of a miRNA silencing component TNRC6A by Argonaute protein. Nucleic Acids Res 43(20):9856–9873. doi:10.1093/nar/gkv1026

Ohtani H, Iwasaki YW, Shibuya A, Siomi H, Siomi MC, Saito K (2013) DmGTSF1 is necessary for Piwi-piRISC-mediated transcriptional transposon silencing in the *Drosophila* ovary. Genes Dev 27(15):1656–1661. doi:10.1101/gad.221515.113

Omoto S, Ito M, Tsutsumi Y, Ichikawa Y, Okuyama H, Brisibe EA, Saksena NK, Fujii YR (2004) HIV-1 nef suppression by virally encoded microRNA. Retrovirology 1:44. doi:10.1186/1742-4690-1-44

Otsuka M, Jing Q, Georgel P, New L, Chen J, Mols J, Kang YJ, Jiang Z, Du X, Cook R, Das SC, Pattnaik AK, Beutler B, Han J (2007) Hypersusceptibility to vesicular stomatitis virus infection in Dicer1-deficient mice is due to impaired miR24 and miR93 expression. Immunity 27 (1):123–134. doi:10.1016/j.immuni.2007.05.014

Ouellet DL, Plante I, Landry P, Barat C, Janelle ME, Flamand L, Tremblay MJ, Provost P (2008) Identification of functional microRNAs released through asymmetrical processing of HIV-1 TAR element. Nucleic Acids Res 36(7):2353–2365. doi:10.1093/nar/gkn076

Pedersen IM, Cheng G, Wieland S, Volinia S, Croce CM, Chisari FV, David M (2007) Interferon modulation of cellular microRNAs as an antiviral mechanism. Nature 449(7164):919–922. doi:10.1038/nature06205

Perez JT, Varble A, Sachidanandam R, Zlatev I, Manoharan M, Garcia-Sastre A, tenOever BR (2010) Influenza A virus-generated small RNAs regulate the switch from transcription to replication. Proc Natl Acad Sci USA 107(25):11525–11530. doi:10.1073/pnas.1001984107

Perez JT, Zlatev I, Aggarwal S, Subramanian S, Sachidanandam R, Kim B, Manoharan M, tenOever BR (2012) A small-RNA enhancer of viral polymerase activity. J Virol 86 (24):13475–13485. doi:10.1128/jvi.02295-12

Pfeffer S, Zavolan M, Grasser FA, Chien M, Russo JJ, Ju J, John B, Enright AJ, Marks D, Sander C, Tuschl T (2004) Identification of virus-encoded microRNAs. Science 304 (5671):734–736. doi:10.1126/science.1096781

Pfeffer S, Sewer A, Lagos-Quintana M, Sheridan R, Sander C, Grasser FA, van Dyk LF, Ho CK, Shuman S, Chien M, Russo JJ, Ju J, Randall G, Lindenbach BD, Rice CM, Simon V, Ho DD, Zavolan M, Tuschl T (2005) Identification of microRNAs of the herpesvirus family. Nat Methods 2(4):269–276. doi:10.1038/nmeth746

Qian K, Pietila T, Ronty M, Michon F, Frilander MJ, Ritari J, Tarkkanen J, Paulin L, Auvinen P, Auvinen E (2013) Identification and validation of human papillomavirus encoded microRNAs. PLoS ONE 8(7):e70202. doi:10.1371/journal.pone.0070202

Rajewsky N (2006) microRNA target predictions in animals. Nature Genet 38(Suppl):S8–13. doi:10.1038/ng1798

Rehmsmeier M, Steffen P, Hochsmann M, Giegerich R (2004) Fast and effective prediction of microRNA/target duplexes. RNA 10(10):1507–1517. doi:10.1261/rna.5248604

Roizman B, Whitley RJ (2013) An inquiry into the molecular basis of HSV latency and reactivation. Ann Rev Microbiol 67:355–374. doi:10.1146/annurev-micro-092412-155654

Sano M, Kato Y, Taira K (2006) Sequence-specific interference by small RNAs derived from adenovirus VAI RNA. FEBS Lett 580(6):1553–1564. doi:10.1016/j.febslet.2006.01.085

Schopman NC, Willemsen M, Liu YP, Bradley T, van Kampen A, Baas F, Berkhout B, Haasnoot J (2012) Deep sequencing of virus-infected cells reveals HIV-encoded small RNAs. Nucleic Acids Res 40(1):414–427. doi:10.1093/nar/gkr719

Seo GJ, Fink LH, O'Hara B, Atwood WJ, Sullivan CS (2008) Evolutionarily conserved function of a viral microRNA. J Virol 82(20):9823–9828. doi:10.1128/jvi.01144-08

Seo GJ, Chen CJ, Sullivan CS (2009) Merkel cell polyomavirus encodes a microRNA with the ability to autoregulate viral gene expression. Virology 383(2):183–187. doi:10.1016/j.virol.2008.11.001

Shukla GC, Singh J, Barik S (2011) MicroRNAs: processing, maturation, target recognition and regulatory functions. Mol Cell Pharmacol 3(3):83–92

Skalsky RL, Cullen BR (2010) Viruses, microRNAs, and host interactions. Ann Rev Microbiol 64:123–141. doi:10.1146/annurev.micro.112408.134243

Song L, Liu H, Gao S, Jiang W, Huang W (2010) Cellular microRNAs inhibit replication of the H1N1 influenza A virus in infected cells. J Virol 84(17):8849–8860. doi:10.1128/jvi.00456-10

Stern-Ginossar N, Elefant N, Zimmermann A, Wolf DG, Saleh N, Biton M, Horwitz E, Prokocimer Z, Prichard M, Hahn G, Goldman-Wohl D, Greenfield C, Yagel S, Hengel H, Altuvia Y, Margalit H, Mandelboim O (2007) Host immune system gene targeting by a viral miRNA. Science 317(5836):376–381. doi:10.1126/science.1140956

Sullivan CS, Grundhoff AT, Tevethia S, Pipas JM, Ganem D (2005) SV40-encoded microRNAs regulate viral gene expression and reduce susceptibility to cytotoxic T cells. Nature 435 (7042):682–686. doi:10.1038/nature03576

Swaminathan G, Martin-Garcia J, Navas-Martin S (2013) RNA viruses and microRNAs: challenging discoveries for the 21st century. Physiol Genom 45(22):1035–1048. doi:10.1152/physiolgenomics.00112.2013

Tang S, Bertke AS, Patel A, Wang K, Cohen JI, Krause PR (2008) An acutely and latently expressed herpes simplex virus 2 viral microRNA inhibits expression of ICP34.5, a viral neurovirulence factor. Proc Natl Acad Sci USA 105(31):10931–10936. doi:10.1073/pnas.0801845105

Tang S, Patel A, Krause PR (2009) Novel less-abundant viral microRNAs encoded by herpes simplex virus 2 latency-associated transcript and their roles in regulating ICP34.5 and ICP0 mRNAs. J Virol 83(3):1433–1442. doi:10.1128/jvi.01723-08

Tambyah PA, Sepramaniam S, Mohamed Ali J, Chai SC, Swaminathan P, Armugam A, Jeyaseelan K (2013) MicroRNAs in circulation are altered in response to influenza A virus infection in humans. PLoS ONE 8(10):e76811. doi:10.1371/journal.pone.0076811

Triboulet R, Mari B, Lin YL, Chable-Bessia C, Bennasser Y, Lebrigand K, Cardinaud B, Maurin T, Barbry P, Baillat V, Reynes J, Corbeau P, Jeang KT, Benkirane M (2007)

Suppression of microRNA-silencing pathway by HIV-1 during virus replication. Science 315 (5818):1579–1582. doi:10.1126/science.1136319

Tycowski KT, Guo YE, Lee N, Moss WN, Vallery TK, Xie M, Steitz JA (2015) Viral noncoding RNAs: more surprises. Genes Dev 29(6):567–584. doi:10.1101/gad.259077.115

Umbach JL, Kramer MF, Jurak I, Karnowski HW, Coen DM, Cullen BR (2008) MicroRNAs expressed by herpes simplex virus 1 during latent infection regulate viral mRNAs. Nature 454 (7205):780–783. doi:10.1038/nature07103

Umbach JL, Nagel MA, Cohrs RJ, Gilden DH, Cullen BR (2009) Analysis of human alphaherpesvirus microRNA expression in latently infected human trigeminal ganglia. J Virol 83(20):10677–10683. doi:10.1128/jvi.01185-09

Umbach JL, Yen HL, Poon LL, Cullen BR (2010) Influenza A virus expresses high levels of an unusual class of small viral leader RNAs in infected cells. mBio 1 (4). doi:10.1128/mBio.00204-10

Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD (2006) A distinct small RNA pathway silences selfish genetic elements in the germline. Science 313(5785):320–324. doi:10.1126/science.1129333

Valencia-Sanchez MA, Liu J, Hannon GJ, Parker R (2006) Control of translation and mRNA degradation by miRNAs and siRNAs. Genes Dev 20(5):515–524. doi:10.1101/gad.1399806

Van Wynsberghe PM, Chan SP, Slack FJ, Pasquinelli AE (2011) Analysis of microRNA expression and function. Meth Cell Biol 106:219–252. doi:10.1016/b978-0-12-544172-8.00008-6

Watanabe T, Takeda A, Tsukiyama T, Mise K, Okuno T, Sasaki H, Minami N, Imai H (2006) Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. Genes Dev 20 (13):1732–1743. doi:10.1101/gad.1425706

Whisnant AW, Bogerd HP, Flores O, Ho P, Powers JG, Sharova N, Stevenson M, Chen CH, Cullen BR (2013) In-depth analysis of the interaction of HIV-1 with cellular microRNA biogenesis and effector mechanisms. mBio 4(2):e000193. doi:10.1128/mBio.00193-13

White MK, Gordon J, Khalili K (2013) The rapidly expanding family of human polyomaviruses: recent developments in understanding their life cycle and role in human pathology. PLoS Pathog 9(3):e1003206. doi:10.1371/journal.ppat.1003206

Xia T, O'Hara A, Araujo I, Barreto J, Carvalho E, Sapucaia JB, Ramos JC, Luz E, Pedroso C, Manrique M, Toomey NL, Brites C, Dittmer DP, Harrington WJ Jr (2008) EBV microRNAs in primary lymphomas and targeting of CXCL-11 by ebv-mir-BHRF1-3. Cancer Res 68(5):1436–1442. doi:10.1158/0008-5472.can-07-5126

Xu N, Segerman B, Zhou X, Akusjarvi G (2007) Adenovirus virus-associated RNAII-derived small RNAs are efficiently incorporated into the rna-induced silencing complex and associate with polyribosomes. J Virol 81(19):10540–10549. doi:10.1128/jvi.00885-07

Yeung ML, Benkirane M, Jeang KT (2007) Small non-coding RNAs, mammalian cells, and viruses: regulatory interactions? Retrovirology 4:74. doi:10.1186/1742-4690-4-74

Young LS, Rickinson AB (2004) Epstein-Barr virus: 40 years on. Nature Rev Cancer 4(10):757–768. doi:10.1038/nrc1452

Ziegelbauer JM, Sullivan CS, Ganem D (2009) Tandem array-based expression screens identify host mRNA targets of virus-encoded microRNAs. Nature Genet 41(1):130–134. doi:10.1038/ng.266

Chapter 2
Microbial Manipulation Host Dark Matter

Samantha Barichievy, Loretta Magagula, Youtaro Shibayama and Musa M. Mhlanga

Abstract In 2010, Francis Collins, director of the NIH, referenced the transcribed yet untranslated component of the human genome as 'dark matter', a term often used by astrophysicists to describe the vast quantities of invisible hypothetical matter known to make up the majority of our universe. Since then, geneticists have set out to shed light on this matter with remarkable success, in an array of biological contexts ranging from cancer to developmental biology. In recent years, rapid advances have been made towards uncovering the functional biological roles of long noncoding RNAs (lncRNAs), which have been estimated to represent 70–90 % of mammalian genomic dark matter. It has become increasingly evident that the primary function of our noncoding genome is to regulate the coding genome. This makes genomic dark matter an attractive evolutionary target for pathogens, who need to alter the cellular host environment in order to promote their survival and propagation. In this review, we focus on the constituents of the mammalian genomic dark matter that are manipulated by viral and microbial pathogens. We also dive deeper into the involvement of ncRNAs, including enhancer RNAs (eRNAs), in the innate immune response against intracellular pathogens. This commentary further highlights how dark and abstruse our noncoding genome still is, particularly in the context of infection biology.

S. Barichievy (✉) · L. Magagula · Y. Shibayama · M.M. Mhlanga (✉)
Gene Expression and Biophysics Group, CSIR Synthetic Biology ERA, Brummeria, Pretoria, South Africa
e-mail: samantha.barichievy@astrazeneca.com

M.M. Mhlanga
e-mail: musa@mhlangalab.org

S. Barichievy
Discovery Sciences, AstraZeneca, Gothenburg, Sweden

M.M. Mhlanga
Gene Expression and Biophysics Unit, Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Lisbon, Portugal

L. Magagula · Y. Shibayama · M.M. Mhlanga
Division of Chemical, Systems and Synthetic Biology, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa

A.L. Leitão and F.J. Enguita (eds.), *Non-coding RNAs and Inter-kingdom Communication*, DOI 10.1007/978-3-319-39496-1_2

2.1 Introduction

The English term ‘orchestrate’ is a verb that means ‘to arrange or direct the elements of a situation to produce a desired effect, especially surreptitiously’. The last part of the definition is intriguing, as it suggests the act of orchestration is clandestine or covert, perhaps even ‘hidden’. It follows then that while the conductor of an orchestra follows a score to guide all of the elements comprising a musical performance to a unified act, the piece is covertly directed according to the personal wishes of the conductor. In truth, conductors make adjustments all of which may not seem obvious to the orchestra members themselves. The ability to furtively direct the elements of a situation to produce a desired effect is also a hallmark of successful pathogenic infection. As obligate parasites, infectious viruses and bacteria need to orchestrate the host cellular response to their very presence, in order to ensure the completion of their life cycles. Numerous strategies exist whereby pathogens engineer the host response to infection, many of which are well characterized and indeed, have led to Nobel Prize winning observations. Notably, the majority of these described host-pathogen interactions involve cellular proteins and microbial manipulation of the coding portions of host cellular genomes. However, with the recent explosion in our understanding of the role noncoding RNAs have in gene regulation, it is unsurprising that pathogens can, and do, manipulate these elementary cellular tools to their own advantage. It seems then that much like the conductor of an orchestra, microbial pathogens use noncoding RNAs akin to a conductor’s baton, to chisel the intracellular environment to their liking.

2.2 Defining Cellular Dark Matter

The central dogma of gene expression has long been taught as the flow of information from DNA to an RNA intermediate and then to protein. The discovery of messenger RNA in 1961 confirmed the intermediate position of RNA within the central dogma (Jacob and Monod 1961). There followed several decades of research confirming the flow of genetic information to proteins via mRNAs, thereby cementing their absolute requirement for protein production, and reinforcing the notion that genetic output is almost completely executed by proteins. The ‘one gene, one enzyme’ paradigm thus helped to obscure our understanding of the human genome, despite being cautioned against by the Nobel laureate Barbara McClintock in a letter she penned in 1950 to Marcus Rhoades. In it she asked ‘Are we letting a philosophy of the protein-coding gene control (our) reasoning? What then is the philosophy of the gene?’ (Mhlanga 2012). Her reservations gained support as observations accumulated showing that cells, in particular eukaryotes, contain large amounts of transcribed RNA that does not code for proteins. While some of these transcripts were shown to comprise the RNA translation machinery (such as rRNAs and tRNAs), and others were shown to be spliced variants of

protein-coding mRNAs, the vast majority remained unexplained (Nickerson et al. 1989; Paul and Duerksen 1975; Salditt-Georgieff and Darnell 1982; Weinberg and Penman 1968). The predominating protein-centric view of the day with its mechanical orientation even led to the term 'junk DNA' being used to describe this unexplained portion of the genome (Mhlanga 2012). However, when the complete sequence of the human genome was published (Lander et al. 2001; Venter et al. 2001), and it became clear just how much non protein-coding RNA is encoded within our own DNA, early proponents of the importance of noncoding RNA gained traction and since then a more nuanced view is emerging regarding our understanding of 'genes'.

Conservatively, only 2 % of the human genome is comprised of protein-coding genes, leaving 98 % of our genome in the dark (Lander et al. 2001; Venter et al. 2001). It is this disproportionate abundance of DNA that we do not comprehend which prompted use of the phrase 'junk DNA' and more recently cellular 'dark matter' when describing the noncoding portion of eukaryotic genomes (Yamada et al. 2003). In astronomical terms, 'dark matter' refers to the matter known to make up perhaps 90 % of the mass of the universe, but which is not directly observable. While studies in prokaryotes neatly supported the idea that genes are generally convertible to proteins via mRNA, comparative genomic studies have revealed that this is not true of eukaryotes. The ratio of coding to noncoding DNA increases as a function of developmental complexity, with fungi and plant genomes comprising 50 % noncoding DNA, and the noncoding proportion in mammalian genomes approaching that of humans at 98 % (Mattick 2004). Without delving deeply into the arena of systems biology, complex organisms obviously require combinations of interrelated gene expression programs in order to layout and maintain the precisely patterned and positionally distinct cell types, body plans and structures of which they are comprised. The fact that typical eukaryotic life cycles span a period of time adds a layer of intricacy, and means that genetic regulation of complex organisms must occur in a four-dimensional space (Buchler et al. 2003; Levine and Tjian 2003). As all of that information must be encoded within the DNA, it is highly suggestive that the abundant cellular dark matter is required to regulate such complex genomes.

The difficulties in understanding noncoding DNA means that noncoding RNAs remain tricky to accurately define and thus quantify. However, a few studies have revealed that noncoding RNAs may be expressed in excess of 20-fold compared to protein-coding RNAs (Nagano and Fraser 2011; Ponting and Belgard 2010). Noncoding RNAs have been segregated into small and long noncoding fractions, based on a somewhat arbitrarily delineated size cutoff of 200 bp. Small noncoding RNAs (<200 bp) are the most abundant transcripts in a cell but are encoded by the lowest number of genes. They include microRNAs, tRNAs, rRNAs and snoRNAs, among others (Nickerson et al. 1989; Paul and Duerksen 1975; Salditt-Georgieff and Darnell 1982; Weinberg and Penman 1968). Long noncoding RNAs (>200 bp) are the least abundant transcripts but are encoded by the highest number of genes comprising nearly 10,000 of the total ~18,000 human genes, and only a handful have been functionally characterized (Cabili et al. 2011; Derrien et al. 2012).

Broadly put, small noncoding RNAs regulate 'how and how much of each gene is made', while long noncoding RNAs regulate 'what genes are expressed when'. These descriptions have been shaped by the burst of long noncoding RNA research within the last decade, which has shown them to be involved in diverse cellular functions regulating nearly all aspects of gene expression from transcription and mRNA degradation, to splicing, translational efficiency, and even regulation of the 3D nuclear architecture via chromatin modification. In addition, long noncoding RNAs have been shown to regulate critical aspects of development, regeneration and disease.

2.3 Cellular Dark Matter Function

Recently the definition of cellular 'dark matter' has been restricted to long noncoding RNAs (lncRNAs) (Derrien et al. 2012). By combining genomic characterisation with chromatin signatures and RNA sequencing across several mammals, consortium-led efforts by ENCODE and FANTOM have identified 15,931 lncRNAs in the human genome (GENCODE version 23). The detailed intricacies of lncRNA definitions are reviewed elsewhere (Guttman and Rinn 2012; Hu et al. 2012; Rinn and Chang 2012; Wang and Chang 2011) but briefly, lncRNAs are typically expressed in a highly tissue-specific manner, their DNA region displays characteristic chromatin signatures (histone 3 lysine 4 tri-methylation, H3K4me3, and histone 3 lysine 36 tri-methylation, H3K36me3), and they are often co-expressed with neighbouring genes (Cabili et al. 2011; Derrien et al. 2012; Guttman and Rinn 2012). In addition, lncRNAs are usually in close proximity to regions of the genome that are rich in protein-coding genes, and they seem to be particularly abundant near transcription factors (Guttman and Rinn 2012; Ponjavic et al. 2009). Early studies showed that certain lncRNAs regulate specific target genes via epigenetic modifications (Pandey et al. 2008; Penny et al. 1996; Sleutels et al. 2002; Wang and Chang 2011). Indeed, it is now appreciated that lncRNAs can alter histone proteins to induce gene activation or repression (Flynn and Chang 2012). Altering the chromatin may also extend to chromosomal looping as certain lncRNAs are transcribed from enhancer regions thereby allowing them to activate genes independently of distance or local genetic context (Orom et al. 2010; Wang et al. 2011). Intriguingly, the functions of lncRNAs are closely tied to their interactions with one or more protein-binding partners, and the associated mechanisms employed during lncRNA-mediated regulation are diverse (Cech and Steitz 2014; Guttman and Rinn 2012; Hu et al. 2012; Kornienko et al. 2013; Rinn and Chang 2012; Wang and Chang 2011).

As more lncRNAs are catalogued and functionally characterised, the assortment of mechanisms by which they regulate their target genes is being refined. At present, they are classed as 'decoys', 'scaffolds', 'guides' and 'enhancers' (Guttman and Rinn 2012; Hu et al. 2012; Rinn and Chang 2012; Wang and Chang 2011). LncRNA decoys bind to and thus titrate away DNA-binding proteins such as

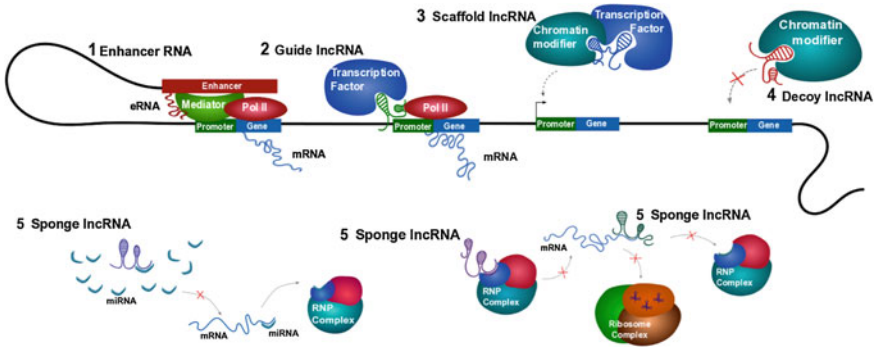


Fig. 2.1 Molecular mechanisms of lncRNAs. 1 Enhancer RNAs (eRNAs) are enhancer-derived transcripts that signal Mediator-directed chromatin looping to form long-range contacts between the parental enhancer sequence and targeted gene promoter thereby activating gene transcription. **2** Guide lncRNAs recruit proteins such as chromatin modifiers and transcription factors to target genes either *in cis* or in *trans*. **3** Scaffold lncRNAs bring multiple proteins within a single complex or spatial proximity. **4** Decoy lncRNAs titrate proteins such as transcription factors and chromatin modifiers away from gene loci. **5** Sponge lncRNAs compete for binding with miRNAs, RNP complexes and/or mRNAs thereby regulating translation

transcriptions factors, or lncRNA decoys can bind miRNAs thereby acting as molecular 'sponges'. Scaffold lncRNAs bring two or more proteins within a single complex or spatial proximity. As a scaffold, lncRNAs can also guide their bound protein(s), such as chromatin modifiers, to a specific DNA or RNA target sequence. Lastly, lncRNA enhancers guide chromosomal looping to exert a *cis* gene regulatory effect (Fig. 2.1). Each of these mechanisms underlies gene expression as they are at the heart of transcriptional control. This is also a critical point of manipulation for pathogens, which are particularly well adapted to altering the host cell in their favour. Historically our understanding of host-pathogen interactions has been informed by how viruses and bacteria manipulate cellular transcription and translation. This allows them to either trigger or suppress the host immune response, as well as to hijack the cellular machinery to ensure production of their own pathogenic components including virulence factors such as toxins. However, when considering that lncRNAs most likely regulate the very proteins that are targeted by pathogens, it is highly likely that host-pathogen interactions extend to the cellular dark matter as well.

2.4 Viruses and Cellular Dark Matter

As obligate intracellular parasites, viruses must exploit host components to complete portions of their life cycle. Typically, viral infection is detected by host cells early on and they mount an innate immune response that often includes inflammation. Unsurprisingly, several lncRNAs have been shown to regulate innate

immunity (reviewed in Atianand and Fitzgerald 2014; Aune and Spurlock 2016; Carpenter 2016; Zhang and Cao 2015). Notably, the identification of lncRNAs involved in innate immunity has typically been based on whole genome lncRNA profiling by RNA-seq or microarrays, and in response to specific stimuli such as toll-like receptor (TLR) agonists. A few of these lncRNAs have also been characterised in response to specific viral infections including the Human Immunodeficiency Virus (HIV), Influenza A Virus (IAV), and the Herpes simplex virus (HSV).

HIV is a retrovirus that begins and ends its infection cycle with two ssRNA copies of the genome, via a critical dsDNA intermediate that must be integrated within the host genome, at a chromatin region that is conducive to integration (Lusic and Giacca 2014; Marini et al. 2015). This action induces DNA damage in the cellular chromatin, alters its 3D structure, and triggers innate immunity, ultimately leading to latency and chronic infection by the virus (Ackerman et al. 2012; Jackson and Bartek 2009; Lilley et al. 2007; Mogensen et al. 2010). The viral integrase protein is responsible for cleaving the host DNA and together with host proteins, enables integration of the proviral genome (Demeulemeester et al. 2015). The generation of a double strand break (DSB) within the cellular chromatin is inherent in this, and because there is no intact sister chromatid to serve as a template, DSBs are the most severe DNA lesions for mammalian cells to endure (Jackson and Bartek 2009). DSBs are so poorly tolerated that a single such lesion within an essential gene can kill a cell (Khanna and Jackson 2001; Rich et al. 2000). As HIV must induce a DSB to complete its life cycle, apoptosis is a likely outcome of infection and indeed continuous decline of CD4 + T cells is used to monitor disease progression. Both viral-mediated integration (Cooper et al. 2013) and abortive infection (Doitsh et al. 2010) drive CD4 + T cell death, but there is also widespread dissemination of the virus throughout the host that is facilitated by infected macrophages. The virus is able to prevent TRAIL-induced apoptosis in macrophages via unknown mechanisms (Swingler et al. 2007) but more intriguingly, HIV is able to selectively impair apoptosis in macrophages by controlling a key apoptosis regulator, namely lncRNA-p21 (Barichievy, personal communication; Barichievy et al. 2015).

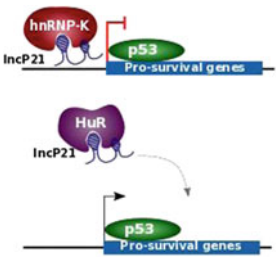
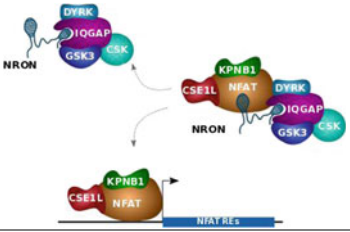
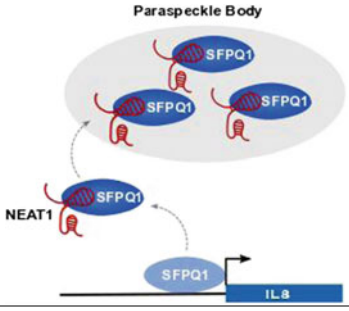
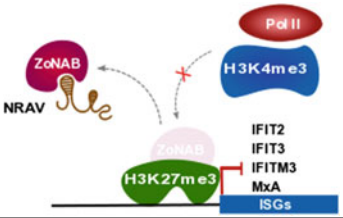
All metazoans have evolved sensitive mechanisms to detect all forms of DNA damage (Hartlerode and Scully 2009) and the tumour suppressor protein p53 is a core transcription factor within the subsequent cellular response to DNA damage (Meek 2004). The activation of p53 initiates various signaling cascades that culminate in either apoptosis, senescence or cell-cycle arrest (Zhou and Elledge 2000). The latter pathway provides time for cells to repair the DNA damage, while senescence and apoptosis are terminal pathways (Riley et al. 2008) and are thus tightly controlled and co-ordinated. Fairly recent data uncovered that the p53-mediated apoptosis response is regulated by an intergenic lncRNA that acts with the nuclear-localised protein hnRNP-K (Huarte et al. 2010). LncRNA-p21 guides hnRNP-K to specific prosurvival p53 target genes *in cis* (Huarte et al. 2010) and *in trans* (Dimitrova et al. 2014). One of these target genes is MAP2K1, which is the primary kinase responsible for phosphorylating ERK2 in healthy cells as part

of the regular survival cycle (Chang and Karin 2001). In healthy cells, both p53 and hnRNP-K are negatively regulated by HDM2 (Enge et al. 2009; Moumen et al. 2005) and therefore p53-transcribed genes, including lncRNA-p21, are not expressed. Concurrently, MAP2K1-activated ERK2 phosphorylates hnRNP-K causing it to accumulate in the cytoplasm (Habelhah et al. 2001) and thus prevent it from complexing with lncRNA-p21 in the nucleus. In addition, healthy cells further negatively regulate lncRNA-p21 by HuR-initiated Ago2/let7-mediated degradation (Yoon et al. 2012). The combined effects of these intersecting pathways ensures cellular survival. In contrast, DNA damage such as DSBs lead to alteRNAtive modifications of p53 thus nullifying HDM2 regulation (Enge et al. 2009). Different upstream signaling cascades prevent ERK2 activation and thus hnRNP-K is able to enter the nucleus (Moumen et al. 2005), interact with lncRNA-p21 and trigger apoptosis (Huarte et al. 2010). The ERK2/lncRNA-p21 intersection is the pivot point manipulated by HIV in order to evade apoptosis in macrophages (Barichievy et al. 2015).

In addition to cellular survival, activated ERK2 is required by HIV during the integration process, and forms part of the pre-integration complex (Bukong et al. 2010). As a consequence of the virus gaining control of this host protein prior to the actual integration event, any subsequent DSBs can be masked. Indeed, our recent unpublished data show that HIV integration does not cause ATM autophosphorylation or downstream activation of apoptosis-specific marks on p53, thus lncRNA-p21 is not transcribed (Barichievy et al. 2015). Central to this is HIV's control of ERK2 and its immediate upstream activator MAP2K1, as inhibition of these host factors leads to apoptosis only in the presence of virus. In addition, by controlling MAP2K1/ERK2, HIV ensures that hnRNP-K remains cytoplasmic and thus unavailable for binding with lncRNA-p21 and subsequent initiation of apoptosis. Intriguingly, integration of the provirus in CD4 + T cells is not facilitated by ERK2 but rather through the actions of JNK and Pin1 (Manganaro et al. 2010). Most likely this is because ERK2 expression is shut down when dual positive CD4 +/CD8 + cells differentiate into CD4 + cells (Chang et al. 2012; Fischer et al. 2005). The intimate tipping point between cellular survival and apoptosis at the intersection of lncRNA-p21 and ERK2 thus possibly only occurs in macrophages, and it seems that HIV has evolved a pivotal mechanism to conduct the cell in favour of viral survival (Table 2.1).

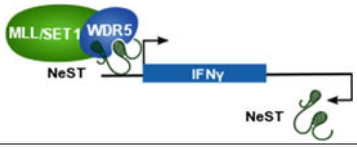
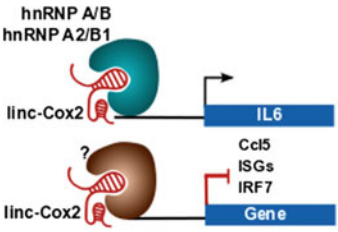
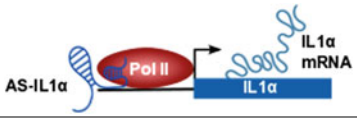
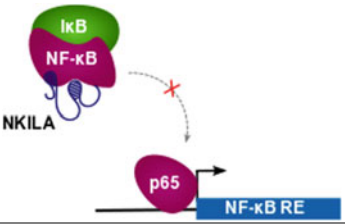
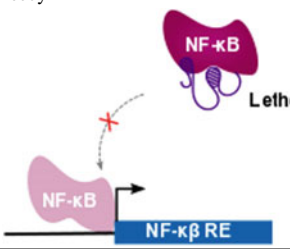
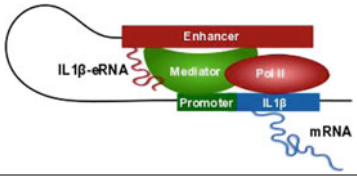
In addition to its manipulation of lncRNA-p21 in macrophages in order to evade apoptosis (Barichievy et al. 2015), HIV modulates the expression of two other lncRNAs in CD4 + T cells in order to enhance viral expression (Imam et al. 2015; Zhang et al. 2013). It has long been known that the nuclear factor of activated T cells (NFAT) transcription factor, which is specifically expressed in primary CD4 + T cells, enhances HIV gene expression by binding to the viral LTR (Cron et al. 2000). Intriguingly, the nuclear import of NFAT is repressed via the interaction of lncRNA NRON with importin-beta family proteins (Sharma et al. 2011; Willingham et al. 2005), thus NRON is a negative regulator of NFAT activity. HIV is able to leverage the gene enhancement function of NFAT by blocking NRON expression early during the infection cycle although the mechanism remains unclear

Table 2.1 LncRNA function during pathogen infection

LncRNA	Species	Stimulus	Type and description	Reference
lncP21	Mouse (human)	HIV	Guide 	Barichievy et al. (2015). Barichievy, submitted
NRON	Mouse (human)	HIV	Decoy, Scaffold 	Willingham et al. (2005), Sharma et al. (2011)
NEAT1	Human (mouse)	IAV HSV HIV	Guide, Scaffold 	Zhang et al. (2013), Imamura et al. (2014)
NRAV	Human	IAV	Decoy 	Ouyang et al. (2014)
VIN	Human	IAV	Unknown	Winterling et al. (2014)

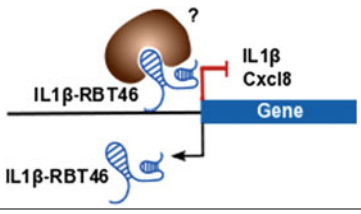
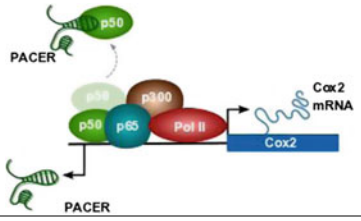
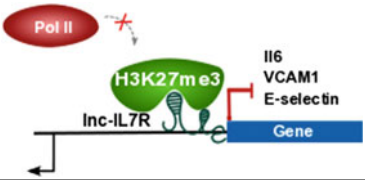
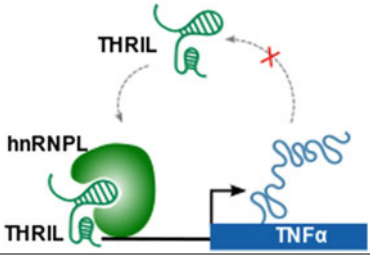
(continued)

Table 2.1 (continued)

LncRNA	Species	Stimulus	Type and description	Reference
NeST	Mouse	*Salmonella*	Guide, Scaffold 	Gomez et al. (2013)
lincRNA-COX2	Mouse	LPS *Listeria* Pam_3CSK_4	Guide 	Guttman et al. (2009), Carpenter et al. (2013)
AS-IL1α	Mouse	LPS *Listeria* Pam_3CSK_4	Guide 	Chan et al. (2015)
NKILA	Human	LPS TNFα	Scaffold 	Liu et al. (2015)
Lethe	Mouse	TNFα	Decoy 	Rapicavoli et al. (2013)
IL1b-eRNA	Human	LPS	Scaffold, Enhancer 	Ilott et al. (2014)

(continued)

Table 2.1 (continued)

LncRNA	Species	Stimulus	Type and description	Reference
IL1b-RBT46	Human	LPS	Guide 	Ilott et al. (2014)
PACER	Human	LPS	Decoy 	Krawczyk and Emerson (2014)
Lnc-IL7R	Human	LPS Pam_3CSK_4	Guide 	Cui et al. (2014)
THRIL	Human	Pam_3CSK_4	Guide 	Li et al. (2014)

(Imam et al. 2015) (Table 2.1). In addition, the viral accessory proteins Nef and Vpu seem to have contrasting effects on NRON depending on the stage of infection, although they may be acting as 'molecular rheostats' to finely tune T cell activation in favour of viral replication, not cell death.

Within all mammalian cells, the nucleus contains many distinct domains including nuclear bodies known as paraspeckles. Within paraspeckles, the nuclear paraspeckle assembly transcript 1 lncRNA, NEAT1, modulates the cellular response to stress via triage of specific proteins within paraspeckles (Imamura et al. 2014). In HIV-infected CD4 + T cells, NEAT1 increases overall viral expression by enhancing nuclear export of HIV transcripts although the underlying mechanism

remains obscure (Zhang et al. 2013). It may be that as NEAT1 also represses the RNA-specific adenosine deaminase B2 (ADARB2) gene, nucleocytoplasmic transport of all ADARB2-sensitive transcripts, including HIV mRNAs, may be affected (Zhang et al. 2013). Furthermore, as modulation by NEAT1 of the host protein splicing factor proline/glutamine rich (SFPQ) leads to released repression of the interleukin-8 (IL-8) cytokine (Imamura et al. 2014), it may be that NEAT1 is more broadly involved in innate immunity. Indeed, the NEAT1/SFPQ interaction has been noted for IAV and HSV as well (Imamura et al. 2014) (Table 2.1).

In addition to NEAT1, IAV infection has also been linked to two additional functionally characterised lncRNAs, namely NRAV and VIN (Ouyang et al. 2014; Winterling et al. 2014). Mammalian cells respond to infecting viral components via pathogen recognition receptors such as RIG-I, MDA5 and TLR3 (Collins and Mossman 2014). Their activation initiates a signaling cascade that culminates in the expression of hundreds of antiviral proteins encoded by interferon-stimulated genes (ISGs) including MxA (Yan and Chen 2012). Following the activation of this innate immune response, a regulatory cascade is switched on as part of a complex and interconnected network that also includes epigenetic factors, thus ensuring that a rapid antiviral defence is mounted with minimal inflammatory damage (Smale 2012). The lncRNA NRAV (negative regulator of antiviral response) promotes IAV replication by suppressing several ISGs including MxA, probably via modulation of the histone marks on these genes (Ouyang et al. 2014). Thus increased NRAV expression is favourable for IAV infection.The virus inducible lncRNA VIN is also induced during IAV infection but not in response to IBV strains (Winterling et al. 2014). While the nuclear location of VIN suggests a role in gene regulation of IAV specifically, the mechanism has not been at all explored. What is clear is that viruses interact with several mammalian lncRNAs (Table 2.1), and this will only increase as research into host-pathogen interactions expands into the cellular dark matter space.

2.5 Bacteria and Cellular Dark Matter

Analogous to their viral counterparts, the proclivity of bacterial pathogens to hijack their cellular host machinery is driven by their inherent parasitic nature. By commandeering aspects of the cellular regulatory networks, bacteria can subvert host cell responses thus creating a microbially-permissive environment. Similarly to viruses, bacteria have evolved several different mechanisms to control host gene expression and nuclear architecture, all of which promote intracellular bacterial survival and progressive infection (Hamon and Cossart 2008). In recent years, these mechanisms have expanded to include bacterial control of lncRNAs, which allow the pathogens to affect transcriptional control of a cohort of co-regulated genes (Carpenter et al. 2013; Cui et al. 2014; Guttman and Rinn 2012; IIott et al. 2014; Liu et al. 2015; Rapicavoli et al. 2013). Indeed, earlier studies hinted at this via the observation that expression of several innate immunity genes was increased in

response to the microbial cell wall component lipopolysaccharide (LPS) (Arbibe et al. 2007; Levine and Tjian 2003; Saccani et al. 2002; Weinmann et al. 2001).

The first example of bacterial-mediated lncRNA manipulation built upon observations in mice that identified a gene which controlled gamma interferon (IFN-γ) and subsequent susceptibility to persistent Theiler's virus infection (Vigneau et al. 2003). This gene was shown to encode the lncRNA NeST (*Nettoie Salmonella pas Theiler's*) and its increased expression caused extended persistence of Theiler's virus infection but also provided resistance to *Salmonella enterica* pathogenesis in infected mice (Gomez et al. 2013). Functionally, NeST was shown to act in *trans*, as an enhancer RNA by binding to the MLL/SET1 H3K4 methylase WDR5, leading to subsequent induction of IFN-γ (Gomez et al. 2013). Specifically, NeST/WDR5 binding resulted in increased deposition of an H3K4me3 chromatin activation mark at the IFN-γ locus in murine splenic and CD8 + T cells (Gomez et al. 2013). Notably, the observation that a single lncRNA can elicit both pro and anti-pathogenic effects reveals an intricately complex phenomenon for lncRNA functioning that is yet to be understood. Yet NeST is not an isolated case in this regard. LncRNA-Cox2 acts as both an activator and suppressor of innate immune response genes (Carpenter et al. 2013). Initially discovered in LPS-stimulated CD11C + bone marrow derived dendritic cells (BMDCs), the functional involvement of lncRNA-Cox2 in the innate immune response was largely uncharacterised (Carpenter et al. 2013). However, in response to bacterial-derived LPS or a synthetic bacterial lipoprotein Pam_3CSK_4, as well as in response to Listeria infection, lncRNA-Cox2 expression was significantly induced in murine BMDCs and macrophages (Carpenter et al. 2013). Furthermore, lncRNA-Cox2 expression and that of its proximal gene *Cox2* were specifically dependent on the TLR signaling adaptor protein MyD88. Collectively, these findings established the induction of lncRNA-Cox2 as part of the innate immune response.

The underlying complexity of lncRNA-Cox2 function as an activator or repressor of immunity was only revealed by careful dissection of specific immune response genes under different stimulation conditions. RNAi-mediated lncRNA-Cox2 knockdown in unstimulated cells upregulated expression of IRF7, CCL5 and other selected ISGs (Carpenter et al. 2013). In contrast, similar lncRNA-Cox2 knockdown in Pam_3CSK_4 stimulated cells decreased expression of TLR1 and IL6. This suggested that lncRNA-Cox2 represses IRF7, CCL5 and other selected ISGs, while activating TLR1 and IL6, which was supported by transcriptomic analysis of macrophages ectopically expressing lncRNA-Cox2. In addition, by overlaying differential gene expression and RNA polymerase II (RNPII) occupancy profiles for various cell conditions, it was established that lncRNA-Cox2 functions as a repressor in the complex with hnRNP A/B and hnRNP A2/B1 (Carpenter et al. 2013). While a functional mechanism for lncRNA-Cox2's activating state has not yet been characterised, it is likely that a different protein partner will be involved. Furthermore, it would be interesting to decode where the divergence lies in the lncRNA-Cox2 induction cascade to identify how this lncRNA determines its immune-related role.

Several other studies have been published revealing the possibility of additional lncRNA involvement in bacterial infection. Notably, only one of these involved infectious bacteria while the remainder utilised LPS or related innate immunity agonists. However, the lessons remain valuable given the paucity of data that covers host-pathogen interactions in the cellular dark matter space. In the only other study to include bacterial infection, the lncRNA AS-IL1α (which is also a natural antisense transcript) was shown to act as a guide for RNPII II to bind the ILα promoter in response to *Listeria monocytogenes* infection or LPS or Pam_3CSK_4 treatment of murine macrophages (Chan et al. 2015). While AS-IL1α was shown to enhance IL1α expression, the possible manipulation of this interaction by *Listeria* was not explored. In a separate study in the human MCF breast cancer cell line, LPS-mediated activation of innate immunity was related to increased expression of a lncRNA termed NKILA (NF-KappaB Interacting LncRNA) which acts as a scaffold to maintain the inhibitory complex comprising NF-κβ and Iκβ (Liu et al. 2015). In this complexed form, NF-κβ localisation is restricted to the nucleus, which prevents NF-κβ-mediated transcriptional activation of genes typically involved in the innate immune response. Interestingly, LPS stimulation resulted in a 12-fold upregulation of NKILA compared to unstimulated MCFs (Liu et al. 2015). Given that LPS is a well-described and potent NF-κβ activator, and Gram-negative bacterial cell wall component, this suggests that such bacteria may upregulate the expression of NKILA in order to block the host from secreting essential innate immunity cytokines and chemokines.

Acute systemic inflammatory diseases such as Kawasaki disease are characterised by elevated circulating TNFα, which was recently shown to be related to the expression of a lnRNA termed THRIL (TNFα and hnRNPL related immunoregulatory LincRNA) (Li et al. 2014). While neither the study, nor the disease touched on pathogenesis, the involvement of macrophages and TNFα make this data interesting. Indeed, microarray analysis revealed that THRIL expression was significantly downregulated in Pam_3CSK_4 stimulated macrophages, and RNAi-mediated knockdown of THRIL strongly reduced TNFα mRNA and protein expression (Li et al. 2014). Interestingly, the downregulation of THRIL expression resulted in reduced expression of a proximal coding gene, Bri3 bp, which also contributed to a reduction of TNFα transcription. At a molecular level, THRIL was found to directly bind hnRNPL, forming a complex whose occupancy at the TNFα promoter was required to maintain basal levels of TNFα expression (Li et al. 2014). These findings led to the hypothesis that binding of the THRIL-hnRNPL complex at the TNFα promoter is required for basal TNFα transcription, and that under stimulated conditions the high expression of TNFα initiates a negative feedback loop in which both THRIL and TNFα are downregulated. Overall, the involvement of THRIL in an inflammatory disease as well as the ability for bacterial lipoprotein mimic, Pam_3CSK_4, to alter its expression suggests that bacteria may reduce THRIL expression in an attempt to deregulate the innate immune response by targeting TNFα.

The pseudogene lncRNA, Lethe, was also shown to be significantly upregulated upon TNFα stimulation, as well as in response to IL-1β activation or LPS treatment

in mouse embryonic fibroblasts (MEFs) (Rapicavoli et al. 2013). Under these conditions, Lethe functions as a decoy and negative inhibitor of NF-κβ signalling by titrating RelA (p65) away from NF-κβ responsive elements, including the Cox2 promoter (Rapicavoli et al. 2013). By displacing the activating NF-κβ subunit away from its responsive elements, no transcription occurs at these loci. Given the critical role of NF-κβ in the inflammatory response, it is unsurprising that three other lncRNAs, namely IL1β-eRNA, IL1β-RBT46 and PACER (p50-associated COX-2 extragenic RNA) regulate transcription of specific NF-κβ target genes (IIott et al. 2014; Krawczyk and Emerson 2014). Both IL1β-eRNA and IL1β-RBT46 were described in LPS-stimulated monocyte-like ThP1 cells, and their knockdown attenuated transcription of CXCL8 and IL6 although their protein binding partners and molecular functions remain obscure (IIott et al. 2014). PACER (p50-associated Cox2 extragenic RNA) was described in LPS-treated human epithelial cells and shown to act as a decoy for the NF-κβ repressive subunit p50 thus occluding it from the Cox2 promoter (Krawczyk and Emerson 2014). This enabled p300 histone acetyltransferase recruitment and assembly of initiating RNPII complexes thus promoting Cox2 expression.

While none of these studies in bacteria involved complete pathogenic stimulation of NF-κβ, it is tempting to speculate that bacteria may exploit the lncRNA-mediated regulation of NF-κβ to control cellular inflammation. From the host cell perspective, these lncRNAs may also be negatively regulating the sustained NF-κβ stimulation that sometimes follows bacterial infection and which can lead to sepsis. In support of this, the cellular lnc-IL7R was shown to attenuate LPS-induced inflammation in ThP1 cells (Cui et al. 2014). Mechanistically, lnc-IL7R does so by increasing deposition of the epigenetic transcription silencing mark H3K27me3 at promoters of inflammatory mediators such as IL6, VCAM1 and E-selectin (Cui et al. 2014). Collectively, these observations serve to establish a clear role for lncRNAs in innate immunity, and particularly in the inflammatory response that is central to microbial infection. While more examples related to bacterial-lncRNA interactions exist as compared to those for viral infections (Table 2.1), the lack of data generated in the presence of whole microbes does underscore that much remains to be discovered in this dark space.

2.6 Enhancer-Derived Short lncRNAs and Their Involvement in the LPS Response

Nearly all lncRNAs that have been explored above influence transcriptional outcome regardless of mechanism. Another group of lncRNAs that has recently emerged as potential principal regulators of transcription are the enhancer RNAs (eRNAs). Although enhancers are known to be indispensable transcriptional regulatory elements in the genome, the current understanding on the presence, dynamics and function of eRNAs is obscure at best, and their involvement in

inter-kingdom interactions is unexplored. However, studies using LPS as summarised below argue in favour of their participation in the innate immune response specifically in response to bacterial pathogens. In addition, where necessary, studies from other contexts such as cancer also facilitate interpretation of the limited data that is available in current literature. Yet despite the paucity of research in this area, there are strong suggestions that they play a role in host-pathogen interactions, although they do so via different functional mechanisms.

Enhancer elements were first observed when transcription of the beta-globin gene was activated by a piece of SV40 DNA acting in *cis* as far as thousands of bases away from the gene (Banerji et al. 1981). From this remarkable observation, the authors of the study correctly predicted that, given their regulatory potential, similar elements may be widespread throughout the genome. Following decades of research, it has now been established that enhancers outnumber protein-coding genes, and regulate the temporal and spatial expression of genes during development, differentiation and homeostasis (Bulger and Groudine 2010). Transcription at an enhancer was first reported at the beta-globin locus (Collis et al. 1990; Tuan et al. 1992), but more surprisingly enhancers were recently found to be pervasively transcribed bi-directionally by RNAPII to produce a class of lncRNAs called eRNAs (Kim et al. 2010). These are mostly unspliced, non-polyadenylated and have a median length of 346 nt (Andersson et al. 2014). As the levels of eRNA expression correlates with those of mRNA at nearby genes (Kim et al. 2010), the eRNAs have been suggested to be used as a predictor for active enhancers, in addition to the canonical chromatin marks such as H3K4me1, H3K4me2 and H3K27ac as well as transcription factor binding. By using enhancer transcription to delineate active enhancers, a recent enhancer atlas that includes over 43,000 such elements across the majority of human cell types and tissues has been compiled (Andersson et al. 2014).

A number of studies have reported important observations on the dynamics in enhancer landscape during signal-dependent gene activation, including TLR4 signaling. In mouse macrophages stimulated with LPS in the presence of IFNγ, 70 % of extragenic RNAPII peaks along the entire genome were shown to be associated with canonical enhancer marks (De Santa et al. 2010). Enhancer transcription was stimulus-regulated and located nearby the induced protein-coding genes, suggesting eRNAs as an important class of lncRNAs in regulating the LPS response. In another study, KLA-stimulation of TLR4 in mouse macrophages led to the appearance of $\sim$3000 new enhancers as identified by the gain of H3K4me2 chromatin marks and the loss of $\sim$1000 enhancers (Kaikkonen et al. 2013). Inhibition of enhancer transcription by BET inhibitors and flavopiridol reduced H3K4me1 and H3K4me2, suggesting that histone methylation is preceded by enhancer transcription (Kaikkonen et al. 2013). This sequence of events, however, may be stimulus and/or cell type specific, as enhancer transcriptional inhibition in estrogen-stimulated breast cancer cells does not lead to changes in histone modifications and other molecular features of enhancers (Hah et al. 2013). In either case, the presence of eRNAs seems to correlate well with other molecular features of active enhancers, arguing in favor of detection of eRNAs as a good measure of enhancer activity. In

summary, the emergence of new transcribed enhancers during TLR4 signaling and their association with nearby coding genes involved in inflammation have hinted at the importance of eRNAs in regulating the innate immune response.

The study describing IL1β-eRNA and IL1β-RBT46 in LPS-stimulated monocyte-like ThP1 cells (IIott et al. 2014) has also correlated the expression of eRNAs to those of nearby coding genes in a genome-wide fashion, using human macrophages stimulated with LPS. This correlation was stronger than that between canonical lncRNAs and coding genes, particularly for genes involved in monocyte inflammatory responses, confirming the potential importance of eRNAs in regulating the temporal nature of innate immune responses. The regulatory nature of transcribing enhancers was supported by an observation where levels of *IL1β* mRNA were significantly attenuated by the TPCA-1-mediated inhibition of NF-κβ, which binds at the *IL1β* enhancer but not its promoter. This observation suggests a sequence of events that starts with NF-κβ binding to the enhancer in order to promote eRNA transcription, which is then followed by transcription of the target coding gene. Although knocking down rapidly induced eRNAs is not technically trivial, the authors validated a functional role for eRNAs by successfully knocking down those associated with *IL1β* and *CXCL8*, leading to a reduction in the levels of those specific mRNAs (Table 2.1) (IIott et al. 2014).

Super-enhancers, or stretch-enhancers, consist of clusters of enhancers that are densely occupied by key transcription factors. Super-enhancers share most features of regular enhancers, but at a much larger scale. Both their coverage of DNA regions and levels of chromatin marks such as H3K27ac and H3K4me1 are on average an order of magnitude greater than those of regular enhancers (Whyte et al. 2013). Consequently, their ability to activate transcription of coding genes and sensitivity to perturbation are also greater. Like regular enhancers, super-enhancers seem cell type specific and therefore are most likely involved in regulating cellular identity. A few studies thus far have shown involvement of super-enhancers in driving the expression of innate immunity genes. In primary human umbilical vein endothelial cells, stimulation with TNFα causes a recruitment of both p65 and BRD4 to regions nearby pro-inflammatory coding genes, forming de novo super-enhancers (Brown et al. 2014). This recruitment came at the expense of pre-existing basal super-enhancers which were 'decommissioned' upon stimulation. Down-regulated genes nearby these 'lost' super-enhancers were involved in angiogenesis and endothelial barrier function. The gain and loss of these super-enhancers resulted in the largest changes to RNAPII occupancy and expression changes of nearby genes (Brown et al. 2014), implicating these elements as having a crucial role in inflammation.

A similar observation has been made in LPS stimulated mouse macrophages, using eRNAs arising from super-enhancers to identify their active status (Hah et al. 2013). Multiple eRNAs are generally transcribed from super-enhancers, and following stimulation were dynamically induced near most innate immunity genes, but reduced near genes involved in cell metabolism and nuclear organization. Although comprising only 3 % of total enhancers, super-enhancers were strongly enriched near genes that were either induced or repressed in response to TLR4 signaling,

raising the possibility that super-enhancers are potential contributors to not only cellular identity but also functional identity. Using global run-on sequencing (GRO-seq), up to nearly 30 % of total nascent RNAs were identified to be eRNAs, the majority of which were produced at super-enhancers. All of the multiple eRNAs arising from individual super-enhancers were observed to be induced or repressed from a population of cells (Hah et al. 2013), but only single-cell or single-allele analyses will elucidate whether all eRNAs within a super-enhancer coordinately respond to stimuli. Nonetheless, all together, these studies demonstrate that super-enhancers and their transcripts are potentially important regulators of innate immunity and are thus implicated in pathogenesis.

A model of enhancer activity whereby they exert their effect on distal promoters by being in close proximity in three dimensional space, is now widely accepted (Lam et al. 2014). This has raised the possibility that eRNAs are mere transcriptional noise that happen to correlate with the induction of nearby genes. A number of studies, however, suggest their functional contribution to activation of gene transcription. For example, as previously mentioned, enhancer knockdowns have caused a reduction in transcription of specific nearby genes (IIott et al. 2014; Lam et al. 2013; Li et al. 2013; Melo et al. 2013; Mousavi et al. 2013). Additionally, eRNA tethering to reporter genes has shown that the eRNA itself, rather than the act of enhancer transcription, is required for reporter activation (Li et al. 2013; Melo et al. 2013). Moreover, an inversion of enhancer sequence, leading to an eRNA with a completely different sequence, abolished enhancer activity, suggesting that a specific eRNA sequence is necessary for its function (Lam et al. 2013). Furthermore, how eRNAs mechanistically activate nearby genes is currently not very clear. A few studies report the ablation of enhancer-promoter contacts upon knockdown of eRNAs (Lai et al. 2013; Li et al. 2013), whereas others report such contacts being unaffected by eRNA levels (Hah et al. 2013; Schaukowitch et al. 2014). Knocking down subunits of the Integrator complex, which is necessary for 3' cleavage of eRNAs to produce their mature form, leads to accumulation of unprocessed, longer forms of the transcripts, resulting in abrogation of EGF-induced enhancer-promoter chromatin looping in HeLa cells (Lai et al. 2015).

It is possible that eRNA functions depend on context, as the above studies were performed on different loci in various cell types using different stimuli. One mechanism of eRNA function may therefore be to initiate and/or maintain looping between enhancers and promoters. Currently a common observation among all studies is that eRNA knockdown causes a reduction in the transcription of specific nearby target genes. A study using neurons has suggested that, upon membrane depolarization, eRNAs act as a decoy for the NELF complex, which otherwise binds nascent RNAs to cause promoter-proximal pausing of RNAPII, thus facilitating the transition from paused RNAPII to productive elongation (Schaukowitch et al. 2014). It has yet to be established whether this is a widespread mechanism of eRNAs or is specific to the conditions tested in this study. Another potential mechanism of eRNAs that has been proposed is trapping of transcription factors thus leading to a positive feedback loop that contributes to stability of gene expression programs (Sigova et al. 2015). More studies addressing mechanistic

details of eRNA functions are expected in the near future. Furthermore, to our knowledge, there is no study to date that has addressed the status of eRNAs specifically upon infection by a pathogen. It will be interesting to see whether eRNA levels are altered as a means to manipulate downstream gene expression in the host, as the use of LPS to stimulate TLR4 signaling cannot provide this information. Given their transcriptional regulatory capacity however, it will not be surprising to find eRNAs as a target by which pathogens control gene expression in the host.

2.7 Discussion and Outlook

In surveying the landscape of transcriptional regulation in eukaryotes, it is evident that diverse cellular modules are implicated. Transcription in eukaryotic organisms and more specifically in mammals, involves several core molecular players in the nucleus. Though still inchoate, our understanding of lncRNAs has shed light on how lncRNAs interact with a fraction of these core molecular players. However, since lncRNA function is deeply entwined within transcriptional regulation, this introduces the potential for the entire landscape of nuclear and cytoplasmic molecular players involved in transcription to be implicated. This may extend well beyond nuclear architecture associated with transcriptional regulation. Broadly speaking classification of lncRNAs beyond the eRNA and non-eRNA variety is one of the key distinctions. Outside this, the diverse mechanisms of action that they possess as evidenced by this review are indications of how poor our current ability to classify lncRNAs is. In terms of transcriptional regulation by enhancers, eRNAs represent a major class of lncRNAs that could potentially be a target of pathogen manipulation of host transcription. This could potentially link pathogen manipulation of lncRNA to 3D chromatin structure and the disruption of long range chromatin contact. However, to date no evidence of such linkages exist.

In this review we have highlighted several bona fide mechanisms implicating transcriptional regulators in the nucleus that have been experimentally validated in viral and bacterial pathogenesis. Many involve clever adaptations that pathogens have made to prevent lncRNA function; those from the Herpes Simplex Virus and its regulation of innate immune genes come to mind. In surveying the lncRNAs so far identified that are implicated in HIV infection, what is striking is how although lncRNAs are significant targets of pathogenesis, this may occur via rather circumspect interventions, although with devastating effects. Viral manipulation of lncRNAs regulating apoptosis has severe consequences in CD4 + T cells and yet no such effects occur in non-CD4 + T cells where the virus evades apoptosis and uses these cells as a reservoir. Indeed it is conceivable that if HIV gained the capacity to infect CD4 + T cells and not cause their demise, its principal effects in humans would be the cause of lymphomas and associated cancers.

These two examples of how lncRNAs can be co-opted by viral pathogens are highly contrasting. One directly implicates transcriptional regulation (HSV) and the

other (HIV) falls outside the 'traditional' targets of transcription regulation, but the effects are no less severe. Notwithstanding, their investigation is highly revelatory of host cell biology. They are also profound reminders of how viral pathogens have evolved abilities to exploit not only the coding or 'well illuminated' part of the genome, but also its dark matter or noncoding regions. It would be naive to believe that our understanding of transcription and its regulation is anywhere near complete. What generates broad enthusiasm for the study of noncoding RNA biology and its interface with pathogenesis is how such models can be used to uncover poorly understood aspects of host transcriptional regulation.

In bacterial infections the centrality of innate immune signaling and NF-Kappa signal transduction are evident by the large number of lncRNAs that have been implicated in innate immune signaling during bacterial pathogenesis or LPS stimulation. Several lncRNAs have been identified that target the NF-Kappa signal module directly and indirectly. Innate immune signaling is 'ground zero' of pathogenesis and evidently where a great deal of pathogenic ingenuity is expended! Pathogens in general and bacteria in particular devote coding potential in their genomes to either up or downregulate innate immune responses. Many coalesce these efforts around lncRNAs that regulate NF-Kappa. Potentially some of these lncRNAs that have been discovered to be pathogenic targets may also become therapeutic ones in inflammatory bowel disease, colitis, cancer and other maladies of the immune system.

This introduces the nascent field of 'drugging' lncRNA activity to abrogate or augment its activity. As this review has highlighted the numerous ways in which pathogens are able to hijack lncRNAs, so too has it exposed a number of new potential therapeutic targets. Thus techniques that are able to discretely target these lncRNAs may introduce an effective host-directed therapy, particularly as host cells are not mere 'orchestral bystanders' but rather retaliate to infection with a complex suite of responses. This promises to be an intensive area of research in the future with the technologies of RNA interference, genome editing and small molecule inhibitors all playing important roles. Circumventing the abilities of pathogens to control host transcription via lncRNAs may be a robust approach in combating pathogenic infection without the use of antibiotics. It is clear that discoveries in the field of lncRNA and pathogenesis will open new vistas in our understanding of transcriptional regulation of host biology.

References

Ackerman ME, Dugast AS, Alter G (2012) Emerging concepts on the role of innate immunity in the prevention and control of HIV infection. Annu Rev Med 63:113–130. doi:10.1146/annurev-med-050310-08522

Andersson R, Gebhard C, Miguel-Escalada I, Hoof I, Bornholdt J, Boyd M, Chen Y, Zhao X, Schmidl C, Suzuki T, Ntini E, Arner E, Valen E, Li K, Schwarzfischer L, Glatz D, Raithel J, Lilje B, Rapin N, Bagger FO, Jorgensen M, Andersen PR, Bertin N, Rackham O, Burroughs AM, Baillie JK, Ishizu Y, Shimizu Y, Furuhata E, Maeda S, Negishi Y,

Mungall CJ, Meehan TF, Lassmann T, Itoh M, Kawaji H, Kondo N, Kawai J, Lennartsson A, Daub CO, Heutink P, Hume DA, Jensen TH, Suzuki H, Hayashizaki Y, Muller F, Forrest AR, Carninci P, Rehli M, Sandelin A (2014) An atlas of active enhancers across human cell types and tissues. Nature 507(7493):455–461. doi:10.1038/nature1278

Arbibe L, Kim DW, Batsche E, Pedron T, Mateescu B, Muchardt C, Parsot C, Sansonetti PJ (2007) An injected bacterial effector targets chromatin access for transcription factor NF-kappaB to alter transcription of host genes involved in immune responses. Nat Immunol 8(1):47–56

Atianand MK, Fitzgerald KA (2014) Long non-coding RNAs and control of gene expression in the immune system. Trends Mol Med 20(11):623–631. doi:10.1016/j.molmed.2014.09.002

Aune TM, Spurlock CF 3rd (2016) Long non-coding RNAs in innate and adaptive immunity. Virus Res 212:146–160. doi:10.1016/j.virusres.2015.07.003

Banerji J, Rusconi S, Schaffner W (1981) Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. Cell 27(2 Pt 1):299–308

Barichievy S, Naidoo J, Mhlanga MM (2015) Non-coding RNAs and HIV: viral manipulation of host dark matter to shape the cellular environment. Front Genet 6:108. doi:10.3389/fgene.2015.00108

Brown JD, Lin CY, Duan Q, Griffin G, Federation AJ, Paranal RM, Bair S, Newton G, Lichtman AH, Kung AL, Yang T, Wang H, Luscinskas FW, Croce KJ, Bradner JE, Plutzky J (2014) NF-kappaB directs dynamic super enhancer formation in inflammation and atherogenesis. Mol Cell 56(2):219–231. doi:10.1016/j.molcel.2014.08.02

Buchler NE, Gerland U, Hwa T (2003) On schemes of combinatorial transcription logic. Proc Natl Acad Sci USA 100(9):5136–5141. doi:10.1073/pnas.0930314100

Bukong TN, Hall WW, Jacque JM (2010) Lentivirus-associated MAPK/ERK2 phosphorylates EMD and regulates infectivity. J Gen Virol 91(Pt 9):2381–2392. doi:10.1099/vir.0.019604-0

Bulger M, Groudine M (2010) Enhancers: the abundance and function of regulatory sequences beyond promoters. Dev Biol 339(2):250–257. doi:10.1016/j.ydbio.2009.11.035

Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, Rinn JL (2011) Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev 25(18):1915–1927. doi:10.1101/gad.17446611

Carpenter S (2016) Long noncoding RNA: novel links between gene expression and innate immunity. Virus Res 212:137–145. doi:10.1016/j.virusres.2015.08.019

Carpenter S, Aiello D, Atianand MK, Ricci EP, Gandhi P, Hall LL, Byron M, Monks B, Henry-Bezy M, Lawrence JB, O'Neill LA, Moore MJ, Caffrey DR, Fitzgerald KA (2013) A long noncoding RNA mediates both activation and repression of immune response genes. Science 341(6147):789–792. doi:10.1126/science.1240925

Cech TR, Steitz JA (2014) The noncoding RNA revolution-trashing old rules to forge new ones. Cell 157(1):77–94. doi:10.1016/j.cell.2014.03.008

Chan J, Atianand M, Jiang Z, Carpenter S, Aiello D, Elling R, Fitzgerald KA, Caffrey DR (2015) Cutting edge: a natural antisense transcript, AS-IL1alpha, controls inducible transcription of the proinflammatory Cytokine IL-1alpha. J Immunol 195(4):1359–1363. doi:10.4049/jimmunol.1500264

Chang CF, D'Souza WN, Ch'en IL, Pages G, Pouyssegur J, Hedrick SM (2012) Polar opposites: Erk direction of CD4 T cell subsets. J Immunol 189(2):721–731. doi:10.4049/jimmunol.1103015

Chang L, Karin M (2001) Mammalian MAP kinase signalling cascades. Nature 410(6824):37–40. doi:10.1038/35065000

Collins SE, Mossman KL (2014) Danger, diversity and priming in innate antiviral immunity. Cytokine Growth Factor Rev 25(5):525–531. doi:10.1016/j.cytogfr.2014.07.002

Collis P, Antoniou M, Grosveld F (1990) Definition of the minimal requirements within the human beta-globin gene and the dominant control region for high level expression. EMBO J 9(1):233–240

Cooper A, Garcia M, Petrovas C, Yamamoto T, Koup RA, Nabel GJ (2013) HIV-1 causes CD4 cell death through DNA-dependent protein kinase during viral integration. Nature 498(7454):376–379. doi:10.1038/nature12274

Cron RQ, Bartz SR, Clausell A, Bort SJ, Klebanoff SJ, Lewis DB (2000) NFAT1 enhances HIV-1 gene expression in primary human CD4 T cells. Clin Immunol 94(3):179–191. doi:10.1006/clim.1999.4831

Cui H, Xie N, Tan Z, Banerjee S, Thannickal VJ, Abraham E, Liu G (2014) The human long noncoding RNA lnc-IL7R regulates the inflammatory response. Eur J Immunol 44(7):2085–2095. doi:10.1002/eji.201344126

De Santa F, Barozzi I, Mietton F, Ghisletti S, Polletti S, Tusi BK, Muller H, Ragoussis J, Wei CL, Natoli G (2010) A large fraction of extragenic RNA pol II transcription sites overlap enhancers. PLoS Biol 8(5):e1000384. doi:10.1371/journal.pbio.1000384

Demeulemeester J, De Rijck J, Gijsbers R, Debyser Z (2015) Retroviral integration: site matters: mechanisms and consequences of retroviral integration site selection. BioEssays 37(11):1202–1214. doi:10.1002/bies.201500051

Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, Lagarde J, Veeravalli L, Ruan X, Ruan Y, Lassmann T, Carninci P, Brown JB, Lipovich L, Gonzalez JM, Thomas M, Davis CA, Shiekhattar R, Gingeras TR, Hubbard TJ, Notredame C, Harrow J, Guigo R (2012) The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res 22(9):1775–1789. doi:10.1101/gr.132159.111

Dimitrova N, Zamudio JR, Jong RM, Soukup D, Resnick R, Sarma K, Ward AJ, Raj A, Lee JT, Sharp PA, Jacks T (2014) LincRNA-p21 activates p21 in cis to promote Polycomb target gene expression and to enforce the G1/S checkpoint. Mol Cell 54(5):777–790. doi:10.1016/j.molcel.2014.04.025

Doitsh G, Cavrois M, Lassen KG, Zepeda O, Yang Z, Santiago ML, Hebbeler AM, Greene WC (2010) Abortive HIV infection mediates CD4 T cell depletion and inflammation in human lymphoid tissue. Cell 143(5):789–801. doi:10.1016/j.cell.2010.11.001

Enge M, Bao W, Hedstrom E, Jackson SP, Moumen A, Selivanova G (2009) MDM2-dependent downregulation of p21 and hnRNP K provides a switch between apoptosis and growth arrest induced by pharmacologically activated p53. Cancer Cell 15(3):171–183. doi:10.1016/j.ccr.2009.01.019

Fischer AM, Katayama CD, Pages G, Pouyssegur J, Hedrick SM (2005) The role of erk1 and erk2 in multiple stages of T cell development. Immunity 23(4):431–443. doi:10.1016/j.immuni.2005.08.013

Flynn RA, Chang HY (2012) Active chromatin and noncoding RNAs: an intimate relationship. Curr Opin Genet Dev 22(2):172–178. doi:10.1016/j.gde.2011.11.002

Gomez JA, Wapinski OL, Yang YW, Bureau JF, Gopinath S, Monack DM, Chang HY, Brahic M, Kirkegaard K (2013) The NeST long ncRNA controls microbial susceptibility and epigenetic activation of the interferon-gamma locus. Cell 152(4):743–754. doi:10.1016/j.cell.2013.01.015

Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zak O, Carey B.W, Cassady JP, Cavili MN, Jaenisch R, Mikkelsen TS, Jacks T, Hacohen N, Bernstein BE, Kellis M, Regev A, Rinn JL, Lander ES (2009) Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 458:223–227. doi:10.1038/nature07672

Guttman M, Rinn JL (2012) Modular regulatory principles of large non-coding RNAs. Nature 482(7385):339–346. doi:10.1038/nature10887

Habelhah H, Shah K, Huang L, Ostareck-Lederer A, Burlingame AL, Shokat KM, Hentze MW, Ronai Z (2001) ERK phosphorylation drives cytoplasmic accumulation of hnRNP-K and inhibition of mRNA translation. Nat Cell Biol 3(3):325–330. doi:10.1038/35060131

Hah N, Murakami S, Nagari A, Danko CG, Kraus WL (2013) Enhancer transcripts mark active estrogen receptor binding sites. Genome Res 23(8):1210–1223. doi:10.1101/gr.152306.112

Hamon MA, Cossart P (2008) Histone modifications and chromatin remodeling during bacterial infections. Cell Host Microbe 4(2):100–109. doi:10.1016/j.chom.2008.07.009

Hartlerode AJ, Scully R (2009) Mechanisms of double-strand break repair in somatic mammalian cells. Biochem J 423(2):157–168. doi:10.1042/BJ20090942

Hu W, Alvarez-Dominguez JR, Lodish HF (2012) Regulation of mammalian cell differentiation by long non-coding RNAs. EMBO Rep 13(11):971–983. doi:10.1038/embor.2012.145

Huarte M, Guttman M, Feldser D, Garber M, Koziol MJ, Kenzelmann-Broz D, Khalil AM, Zuk O, Amit I, Rabani M, Attardi LD, Regev A, Lander ES, Jacks T, Rinn JL (2010) A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. Cell 142(3):409–419. doi:10.1016/j.cell.2010.06.040

IIott NE, Heward JA, Roux B, Tsitsiou E, Fenwick PS, Lenzi L, Goodhead I, Hertz-Fowler C, Heger A, Hall N, Donnelly LE, Sims D, Lindsay MA (2014) Long non-coding RNAs and enhancer RNAs regulate the lipopolysaccharide-induced inflammatory response in human monocytes. Nat Commun 5:3979. doi:10.1038/ncomms4979

Imam H, Bano AS, Patel P, Holla P, Jameel S (2015) The lncRNA NRON modulates HIV-1 replication in a NFAT-dependent manner and is differentially regulated by early and late viral proteins. Sci Rep 5:8639. doi:10.1038/srep08639

Imamura K, Imamachi N, Akizuki G, Kumakura M, Kawaguchi A, Nagata K, Kato A, Kawaguchi Y, Sato H, Yoneda M, Kai C, Yada T, Suzuki Y, Yamada T, Ozawa T, Kaneki K, Inoue T, Kobayashi M, Kodama T, Wada Y, Sekimizu K, Akimitsu N (2014) Long noncoding RNA NEAT1-dependent SFPQ relocation from promoter region to paraspeckle mediates IL8 expression upon immune stimuli. Mol Cell 53(3):393–406. doi:10.1016/j.molcel.2014.01.009

Jackson SP, Bartek J (2009) The DNA-damage response in human biology and disease. Nature 461(7267):1071–1078. doi:10.1038/nature08467

Jacob F, Monod J (1961) Genetic regulatory mechanisms in the synthesis of proteins. J Mol Biol 3:318–356

Kaikkonen MU, Spann NJ, Heinz S, Romanoski CE, Allison KA, Stender JD, Chun HB, Tough DF, Prinjha RK, Benner C, Glass CK (2013) Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. Mol Cell 51(3):310–325. doi:10.1016/j.molcel.2013.07.010

Khanna KK, Jackson SP (2001) DNA double-strand breaks: signaling, repair and the cancer connection. Nat Genet 27(3):247–254. doi:10.1038/85798

Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S, Markenscoff-Papadimitriou E, Kuhl D, Bito H, Worley PF, Kreiman G, Greenberg ME (2010) Widespread transcription at neuronal activity-regulated enhancers. Nature 465(7295):182–187. doi:10.1038/nature09033

Kornienko AE, Guenzl PM, Barlow DP, Pauler FM (2013) Gene regulation by the act of long non-coding RNA transcription. BMC Biol 11:59. doi:10.1186/1741-7007-11-59

Krawczyk M, Emerson BM (2014) p50-associated COX-2 extragenic RNA (PACER) activates COX-2 gene expression by occluding repressive NF-kappaB complexes. Elife 3:e01776. doi:10.7554/eLife.01776

Lai F, Orom UA, Cesaroni M, Beringer M, Taatjes DJ, Blobel GA, Shiekhattar R (2013) Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. Nature 494(7438):497–501. doi:10.1038/nature11884

Lai F, Gardini A, Zhang A, Shiekhattar R (2015) Integrator mediates the biogenesis of enhancer RNAs. Nature 525(7569):399–403. doi:10.1038/nature14906

Lam MT, Cho H, Lesch HP, Gosselin D, Heinz S, Tanaka-Oishi Y, Benner C, Kaikkonen MU, Kim AS, Kosaka M, Lee CY, Watt A, Grossman TR, Rosenfeld MG, Evans RM, Glass CK (2013) Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. Nature 498(7455):511–515. doi:10.1038/nature12209

Lam MT, Li W, Rosenfeld MG, Glass CK (2014) Enhancer RNAs and regulated transcriptional programs. Trends Biochem Sci 39(4):170–182. doi:10.1016/j.tibs.2014.02.007

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKeRNAn K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann Y, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D,

Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blocker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kaspryzk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowki J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ (2001) Initial sequencing and analysis of the human genome. Nature 409(6822):860–921. doi:10.1038/35057062

Levine M, Tjian R (2003) Transcription regulation and animal diversity. Nature 424(6945):147–151. doi:10.1038/nature01763

Li W, Notani D, Ma Q, Tanasa B, Nunez E, Chen AY, Merkurjev D, Zhang J, Ohgi K, Song X, Oh S, Kim HS, Glass CK, Rosenfeld MG (2013) Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature 498(7455):516–520. doi:10.1038/nature12210

Li Z, Chao TC, Chang KY, Lin N, Patil VS, Shimizu C, Head SR, Burns JC, Rana TM (2014) The long noncoding RNA THRIL regulates TNFalpha expression through its interaction with hnRNPL. Proc Natl Acad Sci USA 111(3):1002–1007. doi:10.1073/pnas.1313768111

Lilley CE, Schwartz RA, Weitzman MD (2007) Using or abusing: viruses and the cellular DNA damage response. Trends Microbiol 15(3):119–126. doi:10.1016/j.tim.2007.01.003

Liu B, Sun L, Liu Q, Gong C, Yao Y, Lv X, Lin L, Yao H, Su F, Li D, Zeng M, Song E (2015) A cytoplasmic NF-kappaB interacting long noncoding RNA blocks IkappaB phosphorylation and suppresses breast cancer metastasis. Cancer Cell 27(3):370–381. doi:10.1016/j.ccell.2015.02.004

Lusic M, Giacca M (2014) Ground control to Major Tom: "prepare for HIV landing". Cell Host Microbe 16(5):557–559. doi:10.1016/j.chom.2014.10.012

Manganaro L, Lusic M, Gutierrez MI, Cereseto A, Del Sal G, Giacca M (2010) Concerted action of cellular JNK and Pin1 restricts HIV-1 genome integration to activated CD4 + T lymphocytes. Nat Med 16(3):329–333. doi:10.1038/nm.2102

Marini B, Kertesz-Farkas A, Ali H, Lucic B, Lisek K, Manganaro L, Pongor S, Luzzati R, Recchia A, Mavilio F, Giacca M, Lusic M (2015) Nuclear architecture dictates HIV-1 integration site selection. Nature 521(7551):227–231. doi:10.1038/nature14226

Mattick JS (2004) RNA regulation: a new genetics? Nat Rev Genet 5(4):316–323. doi:10.1038/nrg1321

Meek DW (2004) The p53 response to DNA damage. DNA Repair (Amst) 3(8–9):1049–1056. doi:10.1016/j.dnarep.2004.03.027

Melo CA, Drost J, Wijchers PJ, van de Werken H, de Wit E, Oude Vrielink JA, Elkon R, Melo SA, Leveille N, Kalluri R, de Laat W, Agami R (2013) eRNAs are required for p53-dependent enhancer activity and gene transcription. Mol Cell 49(3):524–535. doi:10.1016/j.molcel.2012.11.021

Mhlanga MM (2012) Overview of non-coding RNAs keystone symposia. Snowbird, Utah, 31 March 2012

Mogensen TH, Melchjorsen J, Larsen CS, Paludan SR (2010) Innate immune recognition and activation during HIV infection. Retrovirology 7:54. doi:10.1186/1742-4690-7-54

Moumen A, Masterson P, O'Connor MJ, Jackson SP (2005) hnRNP K: an HDM2 target and transcriptional coactivator of p53 in response to DNA damage. Cell 123(6):1065–1078

Mousavi K, Zare H, Dell'orso S, Grontved L, Gutierrez-Cruz G, Derfoul A, Hager GL, Sartorelli V (2013) eRNAs promote transcription by establishing chromatin accessibility at defined genomic loci. Mol Cell 51(5):606–617. doi:10.1016/j.molcel.2013.07.022

Nagano T, Fraser P (2011) No-nonsense functions for long noncoding RNAs. Cell 145(2):178–181. doi:10.1016/j.cell.2011.03.014

Nickerson JA, Krochmalnic G, Wan KM, Penman S (1989) Chromatin architecture and nuclear RNA. Proc Natl Acad Sci USA 86(1):177–181

Orom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, Lai F, Zytnicki M, Notredame C, Huang Q, Guigo R, Shiekhattar R (2010) Long noncoding RNAs with enhancer-like function in human cells. Cell 143(1):46–58. doi:10.1016/j.cell.2010.09.001

Ouyang J, Zhu X, Chen Y, Wei H, Chen Q, Chi X, Qi B, Zhang L, Zhao Y, Gao GF, Wang G, Chen JL (2014) NRAV, a long noncoding RNA, modulates antiviral responses through suppression of interferon-stimulated gene transcription. Cell Host Microbe 16(5):616–626. doi:10.1016/j.chom.2014.10.001

Pandey RR, Mondal T, Mohammad F, Enroth S, Redrup L, Komorowski J, Nagano T, Mancini-Dinardo D, Kanduri C (2008) Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. Mol Cell 32(2):232–246. doi:10.1016/j.molcel.2008.08.022

Paul J, Duerksen JD (1975) Chromatin-associated RNA content of heterochromatin and euchromatin. Mol Cell Biochem 9(1):9–16

Penny GD, Kay GF, Sheardown SA, Rastan S, Brockdorff N (1996) Requirement for Xist in X chromosome inactivation. Nature 379(6561):131–137

Ponjavic J, Oliver PL, Lunter G, Ponting CP (2009) Genomic and transcriptional co-localization of protein-coding and long non-coding RNA pairs in the developing brain. PLoS Genet 5(8):e1000617. doi:10.1371/journal.pgen.1000617

Ponting CP, Belgard TG (2010) Transcribed dark matter: meaning or myth? Hum Mol Genet 19(R2):R162–168. doi:10.1093/hmg/ddq362

Rapicavoli NA, Qu K, Zhang J, Mikhail M, Laberge RM, Chang HY (2013) A mammalian pseudogene lncRNA at the interface of inflammation and anti-inflammatory therapeutics. Elife 2:e00762. doi:10.7554/eLife.00762

Rich T, Allen RL, Wyllie AH (2000) Defying death after DNA damage. Nature 407(6805):777–783

Riley T, Sontag E, Chen P, Levine A (2008) Transcriptional control of human p53-regulated genes. Nat Rev Mol Cell Biol 9(5):402–412. doi:10.1038/nrm2395

Rinn JL, Chang HY (2012) Genome regulation by long noncoding RNAs. Annu Rev Biochem 81:145–166. doi:10.1146/annurev-biochem-051410-092902

Saccani S, Pantano S, Natoli G (2002) p38-Dependent marking of inflammatory genes for increased NF-kappa B recruitment. Nat Immunol 3(1):69–75

Salditt-Georgieff M, Darnell JE Jr (1982) Further evidence that the majority of primary nuclear RNA transcripts in mammalian cells do not contribute to mRNA. Mol Cell Biol 2(6):701–707

Schaukowitch K, Joo JY, Liu X, Watts JK, Martinez C, Kim TK (2014) Enhancer RNA facilitates NELF release from immediate early genes. Mol Cell 56(1):29–42

Sharma S, Findlay GM, Bandukwala HS, Oberdoerffer S, Baust B, Li Z, Schmidt V, Hogan PG, Sacks DB, Rao A (2011) Dephosphorylation of the nuclear factor of activated T cells (NFAT)

transcription factor is regulated by an RNA-protein scaffold complex. Proc Natl Acad Sci USA 108(28):11381–11386. doi:10.1073/pnas.1019711108

Sigova AA, Abraham BJ, Ji X, Molinie B, Hannett NM, Guo YE, Jangi M, Giallourakis CC, Sharp PA, Young RA (2015) Transcription factor trapping by RNA in gene regulatory elements. Science 350(6263):978–981. doi:10.1126/science.aad3346

Sleutels F, Zwart R, Barlow DP (2002) The non-coding air RNA is required for silencing autosomal imprinted genes. Nature 415(6873):810–813

Smale ST (2012) Transcriptional regulation in the innate immune system. Curr Opin Immunol 24 (1):51–57. doi:10.1016/j.coi.2011.12.008

Swingler S, Mann AM, Zhou J, Swingler C, Stevenson M (2007) Apoptotic killing of HIV-1-infected macrophages is subverted by the viral envelope glycoprotein. PLoS Pathog 3:1281–1290. doi:10.1371/journal.ppat.0030134

Tuan D, Kong S, Hu K (1992) Transcription of the hypersensitive site HS2 enhancer in erythroid cells. Proc Natl Acad Sci USA 89(23):11219–11223

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, Yao A, Ye J, Zhan M, Zhang W, Zhang H, Zhao Q, Zheng L, Zhong F, Zhong W, Zhu S, Zhao S, Gilbert D, Baumhueter S, Spier G, Carter C, Cravchik A, Woodage T, Ali F, An H, Awe A, Baldwin D, Baden H, Barnstead M, Barrow I, Beeson K, Busam D, Carver A, Center A, Cheng ML, Curry L, Danaher S, Davenport L, Desilets R, Dietz S, Dodson K, Doup L, Ferriera S, Garg N, Gluecksmann A, Hart B, Haynes J, Haynes C, Heiner C, Hladun S, Hostin D, Houck J, Howland T, Ibegwam C, Johnson J, Kalush F, Kline L, Koduru S, Love A, Mann F, May D, McCawley S, McIntosh T, McMullen I, Moy M, Moy L, Murphy B, Nelson K, Pfannkoch C, Pratts E, Puri V, Qureshi H, Reardon M, Rodriguez R, Rogers YH, Romblad D, Ruhfel B, Scott R, Sitter C, Smallwood M, Stewart E, Strong R, Suh E, Thomas R, Tint NN, Tse S, Vech C, Wang G, Wetter J, Williams S, Williams M, Windsor S, Winn-Deen E, Wolfe K, Zaveri J, Zaveri K, Abril JF, Guigo R, Campbell MJ, Sjolander KV, Karlak B, Kejariwal A, Mi H, Lazareva B, Hatton T, Narechania A, Diemer K, Muruganujan A, Guo N, Sato S, Bafna V, Istrail S, Lippert R, Schwartz R, Walenz B, Yooseph S, Allen D, Basu A, Baxendale J, Blick L, Caminha M, Carnes-Stine J, Caulk P, Chiang YH, Coyne M, Dahlke C, Mays A, Dombroski M, Donnelly M, Ely D, Esparham S, Fosler C, Gire H, Glanowski S, Glasser K, Glodek A, Gorokhov M, Graham K, Gropman B, Harris M, Heil J, Henderson S, Hoover J, Jennings D, Jordan C, Jordan J, Kasha J, Kagan L, Kraft C, Levitsky A, Lewis M, Liu X, Lopez J, Ma D, Majoros W, McDaniel J, Murphy S, Newman M, Nguyen T, Nguyen N, Nodell M, Pan S, Peck J, Peterson M, Rowe W, Sanders R, Scott J, Simpson M, Smith T, Sprague A, Stockwell T, Turner R, Venter E, Wang M, Wen M, Wu D, Wu M, Xia A, Zandieh A, Zhu X (2001) The sequence of the human genome. Science 291(5507):1304–1351

Vigneau S, Rohrlich PS, Brahic M, Bureau JF (2003) Tmevpg1, a candidate gene for the control of Theiler's virus persistence, could be implicated in the regulation of gamma interferon. J Virol 77(10):5632–5638

Wang KC, Chang HY (2011) Molecular mechanisms of long noncoding RNAs. Mol Cell 43 (6):904–914. doi:10.1016/j.molcel.2011.08.018

Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R, Chen Y, Lajoie BR, Protacio A, Flynn RA, Gupta RA, Wysocka J, Lei M, Dekker J, Helms JA, Chang HY (2011) A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature 472(7341):120–124. doi:10.1038/nature09819

Weinberg RA, Penman S (1968) Small molecular weight monodisperse nuclear RNA. J Mol Biol 38(3):289–304

Weinmann AS, Mitchell DM, Sanjabi S, Bradley MN, Hoffmann A, Liou HC, Smale ST (2001) Nucleosome remodeling at the IL-12 p40 promoter is a TLR-dependent. Rel-independent Event Nat Immunol 2(1):51–57

Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA (2013) Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell 153(2):307–319. doi:10.1016/j.cell.2013.03.035

Willingham AT, Orth AP, Batalov S, Peters EC, Wen BG, Aza-Blanc P, Hogenesch JB, Schultz PG (2005) A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. Science 309(5740):1570–1573

Winterling C, Koch M, Koeppel M, Garcia-Alcalde F, Karlas A, Meyer TF (2014) Evidence for a crucial role of a host non-coding RNA in influenza a virus replication. RNA Biol 11(1):66–75. doi:10.4161/rna.27504

Yamada K, Lim J, Dale JM, Chen H, Shinn P, Palm CJ, Southwick AM, Wu HC, Kim C, Nguyen M, Pham P, Cheuk R, Karlin-Newmann G, Liu SX, Lam B, Sakano H, Wu T, Yu G, Miranda M, Quach HL, Tripp M, Chang CH, Lee JM, Toriumi M, Chan MM, Tang CC, Onodera CS, Deng JM, Akiyama K, Ansari Y, Arakawa T, Banh J, Banno F, Bowser L, Brooks S, Carninci P, Chao Q, Choy N, Enju A, Goldsmith AD, Gurjal M, Hansen NF, Hayashizaki Y, Johnson-Hopson C, Hsuan VW, Iida K, Karnes M, Khan S, Koesema E, Ishida J, Jiang PX, Jones T, Kawai J, Kamiya A, Meyers C, Nakajima M, Narusaka M, Seki M, Sakurai T, Satou M, Tamse R, Vaysberg M, Wallender EK, Wong C, Yamamura Y, Yuan S, Shinozaki K, Davis RW, Theologis A, Ecker JR (2003) Empirical analysis of transcriptional activity in the Arabidopsis genome. Science 302(5646):842–846

Yan N, Chen ZJ (2012) Intrinsic antiviral immunity. Nat Immunol 13(3):214–222. doi:10.1038/ni.2229

Yoon JH, Abdelmohsen K, Srikantan S, Yang X, Martindale JL, De S, Huarte M, Zhan M, Becker KG, Gorospe M (2012) LincRNA-p21 suppresses target mRNA translation. Mol Cell 47(4):648–655. doi:10.1016/j.molcel.2012.06.027

Zhang Q, Chen CY, Yedavalli VS, Jeang KT (2013) NEAT1 long noncoding RNA and paraspeckle bodies modulate HIV-1 posttranscriptional expression. MBio 4(1):e00596–00512. doi:10.1128/mBio.00596-12

Zhang Y, Cao X (2015) Long noncoding RNAs in innate immunity. Cell Mol Immunol 13(2):138–147. doi:10.1038/cmi.2015.68

Zhou BB, Elledge SJ (2000) The DNA damage response: putting checkpoints in perspective. Nature 408(6811):433–439

Chapter 3
Interplays Between Gut Microbiota and Gene Expression Regulation by miRNAs: Towards a Symbiotic Vision of Host and Guest

Antonella Celluzzi and Andrea Masotti

Abstract Increasing evidences have emphasized the importance of gut microbiota and integrity of the intestinal epithelium to avoid the occurrence of many diseases. Recently, microRNAs have emerged as important gene expression regulators in many conditions. A dysregulated microRNA expression is a common feature of various human diseases, such as cancer, developmental abnormalities, muscular and cardiovascular disorders, and inflammatory diseases. Moreover, exosomal microRNAs have been recently reported to have a crucial role in modulating the bacterial gene expression. So far, the interplays between microRNAs expression and gut microbiota modulation have not been explored in details. To provide further insights into this interesting relationship, in this chapter we discussed some papers appeared in the literature in the last few years.

3.1 The Intestinal Epithelium and the Gut Microbiota

The human body contains a great variety of bacteria, collectively referred to as the human microbiota. The human intestinal tract harbors a diverse and complex microbial community, the gut microbiota, which plays a central role in human health. It has been estimated that our gut contains up to 100 trillion microbes, 1000 bacterial species and 100-fold more genes than those codified by the human genome (Ley et al. 2006b; Qin et al. 2010).

Humans have their first contact with bacteria during birth, when the baby passes through the mother's birth canal (Dethlefsen et al. 2007; Ley et al. 2006a). In the postnatal period, the human intestine is colonized rapidly by an array of microbes. The conditions known to influence the colonization process include the gestational age, the mode of delivery (vaginal birth vs. assisted delivery), diet (breast milk vs. formula),

A. Celluzzi · A. Masotti (✉)
Bambino Gesù Children's Hospital-IRCCS, Gene Expression – Microarrays Laboratory, Viale di San Paolo 15, 00146 Rome, Italy
e-mail: andrea.masotti@opbg.net

A.L. Leitão and F.J. Enguita (eds.), *Non-coding RNAs and Inter-kingdom Communication*, DOI 10.1007/978-3-319-39496-1_3

sanitation, and antibiotic treatment (Adlerberth and Wold 2009; Marques et al. 2010). By the end of the first year of life, infants possess an individually distinct microbial profile, converging toward the characteristic microbiota of an adult. By 2–5 years of age, the microbiota fully resembles that of an adult in terms of composition and diversity (Koenig et al. 2011; Yatsunenko et al. 2012). In the adult, the abundance and the composition of the gut microbial population is different between individuals and this variability is influenced by life style, weight, and overall metabolic state of the host (Tagliabue and Elli 2013; Tehrani et al. 2012). This life-long process of gut colonization led to the formation of a complex ecosystem where the host and its microbiome form an equilibrium that represents a remarkable example of reciprocal adaptation.

Disruptions to the normal balance between the gut microbiota and the host, that can occurs either by changes of the gut microbiota composition or by alterations of the host response, is associated with many pathological conditions such as obesity (Ley et al. 2006b; Turnbaugh et al. 2008), malnutrition (Kau et al. 2011), inflammatory bowel disease (IBD) (Dicksved et al. 2008; Frank et al. 2007), neurological disorders (Gonzalez et al. 2011) and cancer (Lupton 2004).

A coordinated interplay between commensal microbiota and mucosal immune responses occurs to maintain the host intestinal immune homeostasis. In fact, the immune system is the principal regulator of the gut microbiota homeostasis and acts mainly by maintaining the equilibrium between a correct defense against pathogens and tolerance to commensals. Environmental stimuli elicit continuously the intestinal epithelium and many gut cells are necessary to form a barrier against them. In fact, several intestinal diseases are caused by deregulation of the intestinal barrier function (Krogius-Kurikka et al. 2009). The intestinal epithelium is the largest mucosal surface of the body, covering $\sim 400\ m^2$. Its main function is to prevent infections and protect by invading pathogens (Johansson et al. 2011). The intestinal epithelium is organized into crypts and villi, and contains different cells: (i) pluripotent intestinal epithelial stem cells (pluripotent IESCs), that reside at the base of crypts and continuously renew the surface, (ii) enterocytes (for metabolic and digestive functions) and (iii) secretory IECs, including enteroendocrine cells, goblet cells and Paneth cells specialized for maintaining the digestive or barrier function of the epithelium. Enteroendocrine cells represent a link between the central and enteric neuroendocrine systems through the secretion of numerous hormones that regulate the digestive function. The luminal secretion of mucins and antimicrobial proteins (AMPs) by goblet cells and Paneth cells, respectively, establishes a physical and biochemical barrier to microbial contact with the epithelial surface and underlying immune cells (Gallo and Hooper 2012; Kim and Ho 2010). Many regulatory mechanisms control the equilibrium between microbiota and the host intestinal cell response (Coombes and Powrie 2008; Sartor 2008; Strober 2009). In fact, pathogens in commensal bacteria, abnormal microbial composition (i.e., decreased concentrations of protective bacteria) or defective host containment of commensal bacteria (i.e., reduced secretion of antimicrobial peptides to reduce mucosal bacterial overgrowth) may determine an imbalance of this delicate interplay (Sartor 2008). Moreover, this equilibrium is mainly determined

by mucosal dendritic cells, that have an important role in the regulation of intestinal immunity processes (Coombes and Powrie 2008; Strober 2009).

However, we cannot exclude that microRNAs as well may represent complementary molecular determinants potentially involved in these processes.

3.2 microRNAs Biogenesis and Processing

MicroRNAs (miRNAs) have emerged as major regulators of various biological processes and important mediators of immune development and virulence (Choi et al. 2014; O'Connell et al. 2010; Slaby et al. 2009). microRNAs (miRNAs) are short, highly conserved small noncoding RNA molecules naturally occurring in the genomes of plants and animals. miRNAs are 17–27 nucleotides long and regulate post-transcriptionally the mRNA expression, typically by binding to the 3′ untranslated region (3′UTR) of the complementary mRNA sequence, resulting in translational repression and gene silencing (Bartel 2004). microRNAs are transcribed by RNA polymerase II (Pol II) (Cai et al. 2004) and RNA polymerase III (Pol III) (Borchert et al. 2006) in primitive transcripts, named pri-miRNA. Pri-miRNAs are processed into fragments of ∼70-bp, the precursors (pre-miRNAs), in a two-step process catalyzed by the proteins Drosha and Dicer (Lee et al. 2003). The exportin-5 (Exp-5) recognizes the double-stranded pre-miR and transports it from the nucleus to the cytoplasm, irrespective of miRNA nucleotide sequence and the presence of diverse structural motifs (Lund et al. 2004; Okada et al. 2009). Once in the cytoplasm, the RNA III ribonuclease Dicer complex converts the pre-miRNA in a mature miRNA, producing a miRNA–miRNA* duplex (Cullen 2004), which displays a 2-nt 3′overhang at both ends. Only one miRNA strand (the guide strand, or -5p form) of the duplex is loaded into Argonaute protein (AGO) (O'Toole et al. 2006) to form the RISC complex (referred to as the miRISC) that is the effector of the reaction by recognizing the miRNA target in a sequence-specific manner and can mediate various type of gene silencing (Tijsterman and Plasterk 2004), mRNA degradation or translation inhibition (Djuranovic et al. 2012), whereas the inactive strand (the -3p form) is degraded (Kim 2005).

3.3 Interplays Between miRNAs and Microbiota

miRNAs have been also found to be implicated in gut microbiota-host interactions (Kaser et al. 2011). To investigate the mechanisms by which the host cell reprogram their transcription during colonization, germ-free mice were colonized with the microbiota from pathogen-free mice (Dalmasso et al. 2011). RNA extracted from ileum and colon of germ free and colonized mice, showed down- and up-regulated miRNAs: eight microRNAs were expressed in the ileum, whereas seven in the colon. The expression of host miRNAs is modulated in response to microbiota

colonization and this indicates that microbiota modulates host miRNAs expression suggesting an implication of miRNAs in microbiota-mediated host gene regulation. In particular, by intersecting the microarray-detected dysregulated genes with the potential targets of dysregulated miRNAs (predicted by at least two algorithms), the authors identified only one gene, *Abcc3*, potentially targeted by mmu-miR-665 in the colon, whereas no overlapping genes were found in the ileum (Dalmasso et al. 2011). Abcc3 belongs to the multidrug resistance-associated protein family, which mediates the metabolism of xenobiotics and endogenous toxins (Hooper et al. 2001). Therefore, mmu-miR-665 was identified as a microRNA potentially implicated in the colonization of microbiota through the direct targeting and inhibition of *Abcc3*.

Many authors found that different intestinal tracts have distinct miRNAs expression patterns. By using germ-free and conventionally raised mice, the impact of the endogenous microbiota on the global expression of caecal miRNAs in vivo has been investigated by Singh et al. (2012). The murine miRNA signature in the caecum is affected by the presence of the microbiota. Moreover, authors found that 34 putative miRNA target genes encode for proteins involved in the regulation of the intestinal barrier function (i.e., glycosylation enzymes, junctional proteins, proteins found in the mucus layers) and in the immune regulation (i.e., MHCI and II proteins). They found that the expression of miRNAs depends on the endogenous microbiota and that 16 unique miRNAs were deregulated between germ-free and conventional raised mice. By cross-matching the list of intestinal barrier genes predicted to be modulated by differentially expressed miRNAs, with genes already demonstrated to be deregulated in the mucosa of intestinal-specific Dicer knock-out mice (McKenna et al. 2010) the authors supported the hypothesis that gut commensals impact the intestinal barrier via miRNAs expression modulation.

3.4 Inflammatory Diseases

It is now apparent that a dysregulated miRNA expression is a common feature of various human diseases, such as cancer, developmental abnormalities, muscular and cardiovascular disorders, and inflammatory diseases such as inflammatory bowel diseases (IBD) (Takagi et al. 2010). In fact, a study by Xue et al. focused on the microbiota regulation of miRNAs expression and on the maintenance of intestinal homeostasis, and reported a connection between the expression of miR-10a and of its target IL-12/IL-23p40, a key molecule for innate immune responses to commensal bacteria (Xue et al. 2011). The authors found that commensal bacteria down-regulated dendritic cell miR-10a expression via TLR–TLR ligand interactions through a MyD88-dependent pathway and that mice with colitis expressed higher levels of IL-12/IL-23p40 and lower level of gut miR-10a, compared to control mice, opening new perspectives for the study of miRNAs regulation in intestinal diseases.

Intestinal inflammation is characterized by epithelial disruption, loss of barrier function, recruitment of immune cells, and host immune responses to gut microbiota. Recently, it has been observed that PepT1, a di/tripeptide transporter that uptakes bacterial products, is upregulated in inflamed colon tissue (Dai et al. 2015). This peptide has a role in bacterium-associated intestinal inflammation. The amount of this peptide is inversely correlated with the level of miR-193a-3p in inflamed colon tissues with active ulcerative colitis. Moreover, miR-193a-3p reduced PepT1 expression and activity as a target gene and subsequently suppressed the NF-κB pathway, suggesting that miR-193a-3p may have a crucial role to regulate the colonic inflammation process (through PepT1) and to maintain intestinal homeostasis.

Another example of microRNAs that regulate gut mucosal immunity has been reported by Biton et al. who studied miR-375 in mice with an inducible intestinal epithelial cell-specific deficiency in *Dicer1* (*Dicer1*$^{\Delta gut}$) (Biton et al. 2011) (Fig. 3.1).

Biton et al. reported that Dicer1 depletion in the mice gut leads to goblet-cell depletion and that the regulation of goblet-cell differentiation is dependent on the expression of miR-375 (Biton et al. 2011). The expression of this miRNA is able to inhibit the translation of KLF5, an antagonist of the goblet cell–differentiation factor KLF4, supporting the differentiation of goblet cells. Moreover, they observed a lower expression of IL-4, IL-5 and IL-13 in *Dicer1*$^{\Delta gut}$ mice and an enhanced susceptibility to infection by the helminth parasite *Trichuris muris*. IL-13, presumably supplied by T_H2 cells, induces miR-375 in intestinal epithelial cells in vitro and a downstream production of the T_H2-facilitating epithelial cytokine

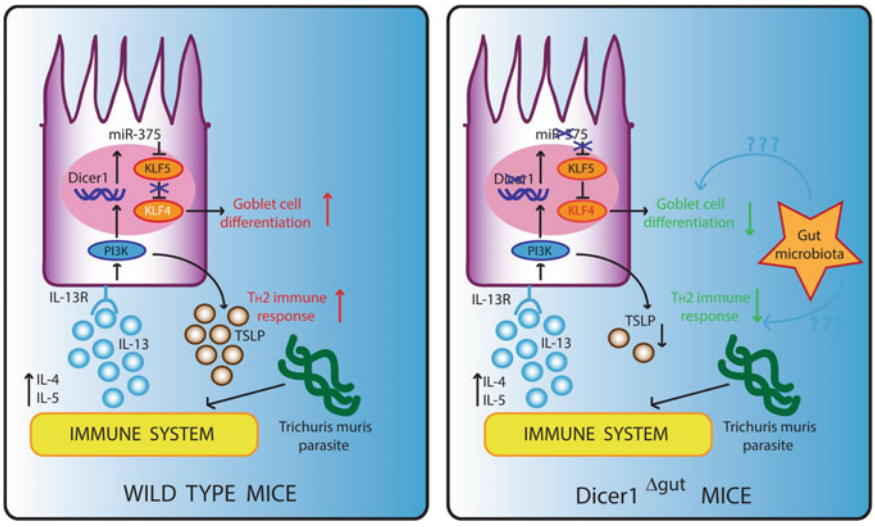


Fig. 3.1 Depletion of Dicer1 or miR-375 results in fewer goblet cells and diminished T_H2 responses

TSLP, indicating an appropriately balanced T_H2 feed-forward loop regulated by miR-375. Based on their results, the authors suggested that miR-375 directs the differentiation of goblet cells and the promotion of antiparasitic T_H2 immune responses. As the miR-375 expression is very high in the human intestine (Wu et al. 2010), mucosal expression of this particular miRNA might also be important in the regulation of intestinal homeostasis and protection against parasite infection in humans (Wu et al. 2010).

Previously (Masotti 2012) we reported a study by Chassin et al. who found that the TLR-4-mediated transcriptional activation of intestinal epithelial cells observed in mice immediately after birth, was induced by an oral ingestion of endotoxins from the environment and induced a post-transcriptional down-regulation of epithelial IRAK1 protein expression, which protected from secondary bacteria-induced epithelial damages (Chassin et al. 2010).

In a very recent paper, Runtsch et al. investigated the role of miR-146a in regulating intestinal immunity and barrier function and verified the miRNA expression in a variety of gut tissues in adult mice (Runtsch et al. 2015). By comparing intestinal gene expression in wild type (WT) and in miR-146a$^{-/-}$ mice, the authors demonstrated that miR-146a repressed a subset of immune-related signaling genes related to an increase of gut barrier and inflammation. Consistent with an enhanced intestinal barrier, Runtsch et al. found that miR-146a$^{-/-}$ mice, a model of Ulcerative Colitis (UC), are more resistant to the dextran sulphate-induced colitis compared to WT. The elevated expression of colonic miR-146a has been observed also in UC patients, therefore suggesting a crucial role for miR-146a in modulating the intestinal barrier function, which is a process that alters gut homeostasis and enhances some intestinal diseases. These results will constitute the basis of further research and will open new perspectives for therapeutic applications.

The same authors reviewed the literature and discussed the influence that miRNAs have on both immune and epithelial cell biology in the mammalian intestines and its impact on the microbiota. However, the authors emphasized the lack of studies aimed at deciphering the functions of specific miRNAs within the gut finalized to the understanding of the cellular mechanisms that promote intestinal homeostasis and the identification of potential molecular targets underlying intestinal diseases such as inflammatory bowel disease and colorectal cancer (Runtsch et al. 2014).

3.5 Symbiosis of Host and Guest

All of the papers discussed in the previous paragraphs described the interplay between the microbiota and the host. In particular, we discussed how microbiota modulates the gene expression of the host through miRNAs. So far, nothing has been know on how the host regulates the microbiota. This is a crucial point, because it represents the missing part in the big picture describing the symbiosis of the host and the guest (Fig. 3.2).

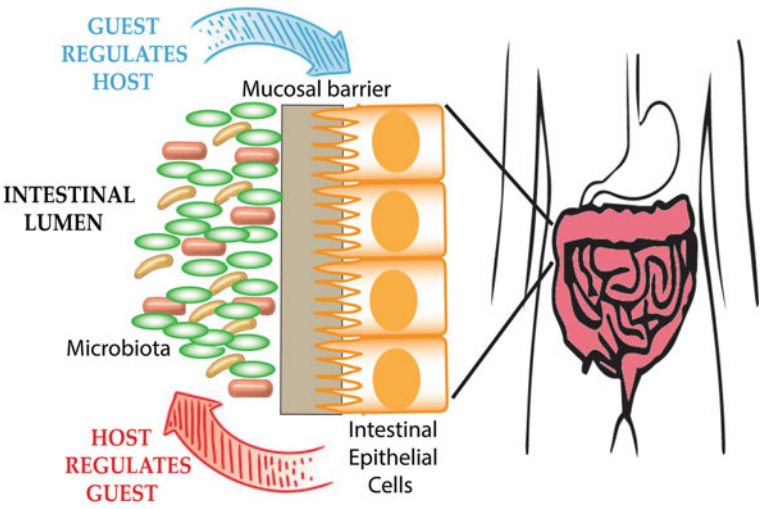


Fig. 3.2 The human intestinal lumen is populated by microorganisms (gut microbiota) that regulate the host gene expression through microRNAs. Similarly, the host produces extracellular vesicles containing microRNAs that regulate the expression of microbial genes. This 'symbiotic loop' is emerging as a powerful inter-kingdom communication system, although the precise molecular mechanisms underlying it are still not know. We have no doubt that this loop will be extensively explored in the next few years

To close this gap, a very recent work by Liu et al. described how the host selectively shapes the microbiota through miRNAs contained in extracellular vesicles (EVs) produced by the host itself (Liu et al. 2016). miRNAs, when contained in vesicles, are relatively stable compared to other RNAs (Jung et al. 2010). Fecal miRNAs can exist in EV-free forms, associated with high-density lipoproteins or argonaute protein (Creemers et al. 2012), or in a completely free form. Liu et al. reported that the miRNAs they have identified and characterized, can target specific bacterial genes after entering the bacteria, modulating their gene expression. In their work, Liu et al. used *Escherichia coli* and *Fusobacterium nucleatum,* two bacterial species that have been reported to promote colorectal cancer (Rubinstein et al. 2013). The authors demonstrated that different miRNAs have different ability to enter into bacteria and that miRNAs shapes bacteria with a temporal and spatial organization (Liu et al. 2016).

3.6 Conclusions

In this chapter, we discussed the papers that appeared in the literature in the last 5 years (Table 3.1), that studied the interplays between gut microbiota and gene expression modulation mediated by microRNAs.

Table 3.1 Papers dealing with the interplays between gut microbiota and gene expression modulation by microRNAs

Aim of the study	Experimental model	Investigated miRNAs	Target genes	Reference
To study whether miRNAs are involved in microbiota-mediated regulation of host gene expression	Germ-free mice colonized with the microbiota from pathogen-free mice	miR-298; miR-128; miR-200c*; miR-342-5p; miR-465c-5p; miR-466d-3p/5p; miR-665; miR-683	Abcc3	Dalmasso et al. (2011)
To study the impact of the endogenous microbiota on the global expression of caecal miRNAs in vivo	Germ-free and conventionally raised mice	miR-21*; miR-351; miR-487b; miR-467a; miR-27b; miR-148a; miR-145; miR-183; miR-133a; miR-150; miR-672; miR-181a; miR-664; miR-455; miR-138*; let-7 g*	34 genes among glycosylation enzymes, junctional proteins, proteins found in the mucus layers and in the immune regulation	Singh et al. (2012)
To study miRNAs affecting the intestinal epithelial monolayer	Mice with an inducible intestinal epithelial cell-specific deficiency in *Dicer1* (*Dicer1*$^{\Delta gut}$)	miR-375	KLF5	Biton et al. (2011)
To study the TLR-4-mediated transcriptional activation of intestinal epithelial cells (IECs)	Mice immediately after birth	miR-146a	IRAK-1	Chassin et al. (2010)
To study microbiota regulation of miRNA expression and intestinal homeostasis	C57BL/6 (B6), B6.IL-10$^{-/-}$, B6.MyD88$^{-/-}$ and B6.RAG$^{-/-}$ mice	miR-10a	IL-12/IL-23p40	Xue et al. (2011)
To study the role of miRNAs in the immune regulation of innate and adaptive responses to microbiota in Inflammatory bowel disease (IBD)	Inflamed ileal and/or colonic tissues of IBD patients and specific pathogen-free female C57BL/6 mice	miR-10a	(IL)-12/IL-23p40/NOD2	Wu et al. (2015)

(continued)

Table 3.1 (continued)

Aim of the study	Experimental model	Investigated miRNAs	Target genes	Reference
To study the role of miRNAs in regulating intestinal immunity and barrier functions	Intestines of germ-free (GF) and specific pathogen-free (SPF) mice and miR-146a$^{-/-}$mice	miR-146a	289 genes were upregulated and 77 genes were downregulate. Among these there are: members of the C-type lectin antimicrobial peptide family Reg3, genes that produce intestinal mucus, intestinal cell adhesion molecules	Runtsch et al. (2015)
To study the role of miRNAs as critical gene regulators and mediators of the activation of host responses to gut microbiota	Colonic tissue samples	miR-193a-3p	PepT1	Dai et al. (2015)
How the microbes are selected and whether the host specifically regulates microbial gene expression	Gut luminal contents from the distal ileum and colon and Dicer1 ΔIEC and Dicer1 fl/fl littermate mice	MiR-101, hsa-miR-515-5p, miR-876-5p, hsa-miR-325, and hsamiR-1253 could potentially target Fn nucleic acid sequences; hsa-miR-4747-3p, hsa-miR-1224-5p, hsa-miR-1226-5p, and hsa-miR-623 could potentially target E. coli nucleic acid sequences	16S rRNA/23S rRNA; E. coli yegHmRNA; RNaseP; rutA mRNA; fucO	Liu et al. (2016)

We wondered whether microRNAs could be exploited therapeutically to modulate an altered gut microbiota composition and ultimately restore a healthy condition. For example, it has been reported recently that the incidence of type 1 diabetes cannot be explained only by genetics, epigenetics and environmental factors only (Gulden et al. 2015). Lifestyle, diet and the use of antibiotics also should be taken into account. The diet supplementation with pre-/pro-biotics has emerged as a potential mean to improve gut integrity and avoid the occurrence of diseases. However, we think that the recent work by Liu and colleagues (Liu et al. 2016) is a clear demonstration that other bacterial modulatory mechanisms can be

elicited, as for example, the use of microRNAs for the dysregulation of bacterial gene expression. We still do not know if it will be possible to modulate gut bacterial composition by simply employing microRNAs (i.e., by modulating gut bacteria gene expression to activate cell death processes that could lead to a progressive enrichment or depletion of a given bacterial population). In any case, if validated, this kind of innovative 'therapeutic' intervention could be exploited also for other pathologies, and not be limited only to diabetes.

In the near future, many other works will surely prompt further research aimed at deciphering the existence of other types of interactions between the host microbiota and the guest itself. The interpretation of the complex 'inter-kingdom communication' system and all the ways and pathways by which these systems interact each other will be the next challenge that we are going to face

References

Adlerberth I, Wold AE (2009) Establishment of the gut microbiota in Western infants. Acta Paediatr 98(2):229–238. doi:10.1111/j.1651-2227.2008.01060.x

Bartel DP (2004) microRNAs: genomics, biogenesis, mechanism, and function. Cell 116(2):281–297

Biton M, Levin A, Slyper M, Alkalay I, Horwitz E, Mor H, Kredo-Russo S, Avnit-Sagi T, Cojocaru G, Zreik F, Bentwich Z, Poy MN, Artis D, Walker MD, Hornstein E, Pikarsky E, Ben-Neriah Y (2011) Epithelial microRNAs regulate gut mucosal immunity via epithelium-T cell crosstalk. Nat Immunol 12(3):239–246. doi:10.1038/ni.1994

Borchert GM, Lanier W, Davidson BL (2006) RNA polymerase III transcribes human microRNAs. Nat Struct Mol Biol 13(12):1097–1101. doi:10.1038/nsmb1167

Cai X, Hagedorn CH, Cullen BR (2004) Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. RNA 10(12):1957–1966. doi:10.1261/rna.7135204

Chassin C, Kocur M, Pott J, Duerr CU, Gutle D, Lotz M, Hornef MW (2010) miR-146a mediates protective innate immune tolerance in the neonate intestine. Cell Host Microbe 8(4):358–368. doi:10.1016/j.chom.2010.09.005

Choi EJ, Kim HB, Baek YH, Kim EH, Pascua PN, Park SJ, Kwon HI, Lim GJ, Kim S, Kim YI, Choi YK (2014) Differential microRNA expression following infection with a mouse-adapted, highly virulent avian H5N2 virus. BMC Microbiol 14:252. doi:10.1186/s12866-014-0252-0

Coombes JL, Powrie F (2008) Dendritic cells in intestinal immune regulation. Nat Rev Immunol 8(6):435–446. doi:10.1038/nri2335

Creemers EE, Tijsen AJ, Pinto YM (2012) Circulating microRNAs: novel biomarkers and extracellular communicators in cardiovascular disease? Circ Res 110(3):483–495. doi:10.1161/CIRCRESAHA.111.247452

Cullen BR (2004) Transcription and processing of human microRNA precursors. Mol Cell 16(6):861–865. doi:10.1016/j.molcel.2004.12.002

Dai X, Chen X, Chen Q, Shi L, Liang H, Zhou Z, Liu Q, Pang W, Hou D, Wang C, Zen K, Yuan Y, Zhang CY, Xia L (2015) MicroRNA-193a-3p Reduces Intestinal Inflammation in Response to Microbiota via Down-regulation of Colonic PepT1. J Biol Chem 290(26):16099–16115. doi:10.1074/jbc.M115.659318

Dalmasso G, Nguyen HT, Yan Y, Laroui H, Charania MA, Ayyadurai S, Sitaraman SV, Merlin D (2011) Microbiota modulate host gene expression via microRNAs. PLoS ONE 6(4):e19293. doi:10.1371/journal.pone.0019293

Dethlefsen L, McFall-Ngai M, Relman DA (2007) An ecological and evolutionary perspective on human-microbe mutualism and disease. Nature 449(7164):811–818. doi:10.1038/nature06245

Dicksved J, Halfvarson J, Rosenquist M, Jarnerot G, Tysk C, Apajalahti J, Engstrand L, Jansson JK (2008) Molecular analysis of the gut microbiota of identical twins with Crohn's disease. ISME J 2(7):716–727. doi:10.1038/ismej.2008.37

Djuranovic S, Nahvi A, Green R (2012) miRNA-mediated gene silencing by translational repression followed by mRNA deadenylation and decay. Science 336(6078):237–240. doi:10.1126/science.1215691

Frank DN, St Amand AL, Feldman RA, Boedeker EC, Harpaz N, Pace NR (2007) Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. Proc Natl Acad Sci USA 104(34):13780–13785. doi:10.1073/pnas.0706625104

Gallo RL, Hooper LV (2012) Epithelial antimicrobial defence of the skin and intestine. Nat Rev Immunol 12(7):503–516. doi:10.1038/nri3228

Gonzalez A, Stombaugh J, Lozupone C, Turnbaugh PJ, Gordon JI, Knight R (2011) The mind-body-microbial continuum. Dialogues Clin Neurosci 13(1):55–62

Gulden E, Wong FS, Wen L (2015) The gut microbiota and Type 1 Diabetes. Clin Immunol 159 (2):143–153. doi:10.1016/j.clim.2015.05.013

Hooper LV, Wong MH, Thelin A, Hansson L, Falk PG, Gordon JI (2001) Molecular analysis of commensal host-microbial relationships in the intestine. Science 291(5505):881–884. doi:10.1126/science.291.5505.881

Johansson ME, Larsson JM, Hansson GC (2011) The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions. Proc Natl Acad Sci USA 108(Suppl 1):4659–4665. doi:10.1073/pnas.1006451107

Jung M, Schaefer A, Steiner I, Kempkensteffen C, Stephan C, Erbersdobler A, Jung K (2010) Robust microRNA stability in degraded RNA preparations from human tissue and cell samples. Clin Chem 56(6):998–1006. doi:10.1373/clinchem.2009.141580

Kaser A, Niederreiter L, Blumberg RS (2011) Genetically determined epithelial dysfunction and its consequences for microflora-host interactions. Cell Mol Life Sci 68(22):3643–3649. doi:10.1007/s00018-011-0827-y

Kau AL, Ahern PP, Griffin NW, Goodman AL, Gordon JI (2011) Human nutrition, the gut microbiome and the immune system. Nature 474(7351):327–336. doi:10.1038/nature10213

Kim VN (2005) MicroRNA biogenesis: coordinated cropping and dicing. Nat Rev Mol Cell Biol 6 (5):376–385. doi:10.1038/nrm1644

Kim YS, Ho SB (2010) Intestinal goblet cells and mucins in health and disease: recent insights and progress. Curr Gastroenterol Rep 12(5):319–330. doi:10.1007/s11894-010-0131-2

Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, Angenent LT, Ley RE (2011) Succession of microbial consortia in the developing infant gut microbiome. Proc Natl Acad Sci USA 108(Suppl 1):4578–4585. doi:10.1073/pnas.1000081107

Krogius-Kurikka L, Lyra A, Malinen E, Aarnikunnas J, Tuimala J, Paulin L, Makivuokko H, Kajander K, Palva A (2009) Microbial community analysis reveals high level phylogenetic alterations in the overall gastrointestinal microbiota of diarrhoea-predominant irritable bowel syndrome sufferers. BMC Gastroenterol 9:95. doi:10.1186/1471-230X-9-95

Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Radmark O, Kim S, Kim VN (2003) The nuclear RNase III Drosha initiates microRNA processing. Nature 425(6956):415–419. doi:10.1038/nature01957

Ley RE, Peterson DA, Gordon JI (2006a) Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell 124(4):837–848. doi:10.1016/j.cell.2006.02.017

Ley RE, Turnbaugh PJ, Klein S, Gordon JI (2006b) Microbial ecology: human gut microbes associated with obesity. Nature 444(7122):1022–1023. doi:10.1038/4441022a

Liu S, da Cunha AP, Rezende RM, Cialic R, Wei Z, Bry L, Comstock LE, Gandhi R, Weiner HL (2016) The Host Shapes the Gut Microbiota via Fecal MicroRNA. Cell Host Microbe 19 (1):32–43. doi:10.1016/j.chom.2015.12.005

Lund E, Guttinger S, Calado A, Dahlberg JE, Kutay U (2004) Nuclear export of microRNA precursors. Science 303(5654):95–98. doi:10.1126/science.1090599

Lupton JR (2004) Microbial degradation products influence colon cancer risk: the butyrate controversy. J Nutr 134(2):479–482

Marques TM, Wall R, Ross RP, Fitzgerald GF, Ryan CA, Stanton C (2010) Programming infant gut microbiota: influence of dietary and environmental factors. Curr Opin Biotechnol 21 (2):149–156. doi:10.1016/j.copbio.2010.03.020

Masotti A (2012) Interplays between gut microbiota and gene expression regulation by miRNAs. Front Cell Infect Microbiol 2:137. doi:10.3389/fcimb.2012.00137

McKenna LB, Schug J, Vourekas A, McKenna JB, Bramswig NC, Friedman JR, Kaestner KH (2010) microRNAs control intestinal epithelial differentiation, architecture, and barrier function. Gastroenterology 139(5):1654–1664, 1664 e1651. doi:10.1053/j.gastro.2010.07.040

O'Connell RM, Rao DS, Chaudhuri AA, Baltimore D (2010) Physiological and pathological roles for microRNAs in the immune system. Nat Rev Immunol 10(2):111–122. doi:10.1038/nri2708

O'Toole AS, Miller S, Haines N, Zink MC, Serra MJ (2006) Comprehensive thermodynamic analysis of 3′ double-nucleotide overhangs neighboring Watson-Crick terminal base pairs. Nucleic Acids Res 34(11):3338–3344. doi:10.1093/nar/gkl428

Okada C, Yamashita E, Lee SJ, Shibata S, Katahira J, Nakagawa A, Yoneda Y, Tsukihara T (2009) A high-resolution structure of the pre-microRNA nuclear export machinery. Science 326(5957):1275–1279. doi:10.1126/science.1178705

Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, Mende DR, Li J, Xu J, Li S, Li D, Cao J, Wang B, Liang H, Zheng H, Xie Y, Tap J, Lepage P, Bertalan M, Batto JM, Hansen T, Le Paslier D, Linneberg A, Nielsen HB, Pelletier E, Renault P, Sicheritz-Ponten T, Turner K, Zhu H, Yu C, Jian M, Zhou Y, Li Y, Zhang X, Qin N, Yang H, Wang J, Brunak S, Dore J, Guarner F, Kristiansen K, Pedersen O, Parkhill J, Weissenbach J, Bork P, Ehrlich SD (2010) A human gut microbial gene catalogue established by metagenomic sequencing. Nature 464(7285):59–65. doi:10.1038/nature08821

Rubinstein MR, Wang X, Liu W, Hao Y, Cai G, Han YW (2013) Fusobacterium nucleatum promotes colorectal carcinogenesis by modulating E-cadherin/beta-catenin signaling via its FadA adhesin. Cell Host Microbe 14(2):195–206. doi:10.1016/j.chom.2013.07.012

Runtsch MC, Round JL, O'Connell RM (2014) microRNAs and the regulation of intestinal homeostasis. Front Genet 5:347. doi:10.3389/fgene.2014.00347

Runtsch MC, Hu R, Alexander M, Wallace J, Kagele D, Petersen C, Valentine JF, Welker NC, Bronner MP, Chen X, Smith DP, Ajami NJ, Petrosino JF, Round JL, O'Connell RM (2015) MicroRNA-146a constrains multiple parameters of intestinal immunity and increases susceptibility to DSS colitis. Oncotarget 6(30):28556–28572. doi:10.18632/oncotarget.5597

Sartor RB (2008) Microbial influences in inflammatory bowel diseases. Gastroenterology 134 (2):577–594. doi:10.1053/j.gastro.2007.11.059

Singh N, Shirdel EA, Waldron L, Zhang RH, Jurisica I, Comelli EM (2012) The murine caecal microRNA signature depends on the presence of the endogenous microbiota. Int J Biol Sci 8 (2):171–186

Slaby O, Svoboda M, Michalek J, Vyzula R (2009) microRNAs in colorectal cancer: translation of molecular biology into clinical application. Mol Cancer 8:102. doi:10.1186/1476-4598-8-102

Strober W (2009) The multifaceted influence of the mucosal microflora on mucosal dendritic cell responses. Immunity 31(3):377–388. doi:10.1016/j.immuni.2009.09.001

Tagliabue A, Elli M (2013) The role of gut microbiota in human obesity: recent findings and future perspectives. Nutr Metab Cardiovasc Dis 23(3):160–168. doi:10.1016/j.numecd.2012.09.002

Takagi T, Naito Y, Mizushima K, Hirata I, Yagi N, Tomatsuri N, Ando T, Oyamada Y, Isozaki Y, Hongo H, Uchiyama K, Handa O, Kokura S, Ichikawa H, Yoshikawa T (2010) Increased expression of microRNA in the inflamed colonic mucosa of patients with active ulcerative colitis. J Gastroenterol Hepatol 25(Suppl 1):S129–S133. doi:10.1111/j.1440-1746.2009.06216.x

Tehrani AB, Nezami BG, Gewirtz A, Srinivasan S (2012) Obesity and its associated disease: a role for microbiota? Neurogastroenterol Motil 24(4):305–311. doi:10.1111/j.1365-2982.2012.01895.x

Tijsterman M, Plasterk RH (2004) Dicers at RISC; the mechanism of RNAi. Cell 117(1):1–3

Turnbaugh PJ, Backhed F, Fulton L, Gordon JI (2008) Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. Cell Host Microbe 3(4):213–223. doi:10.1016/j.chom.2008.02.015

Wu W, He C, Liu C, Cao AT, Xue X, Evans-Marin HL, Sun M, Fang L, Yao S, Pinchuk IV, Powell DW, Liu Z, Cong Y (2015) miR-10a inhibits dendritic cell activation and Th1/Th17 cell immune responses in IBD. Gut 64(11):1755–1764. doi:10.1136/gutjnl-2014-307980

Wu F, Zhang S, Dassopoulos T, Harris ML, Bayless TM, Meltzer SJ, Brant SR, Kwon JH (2010) Identification of microRNAs associated with ileal and colonic Crohn's disease. Inflamm Bowel Dis 16(10):1729–1738. doi:10.1002/ibd.21267

Xue X, Feng T, Yao S, Wolf KJ, Liu CG, Liu X, Elson CO, Cong Y (2011) Microbiota downregulates dendritic cell expression of miR-10a, which targets IL-12/IL-23p40. J Immunol 187(11):5879–5886. doi:10.4049/jimmunol.1100535

Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, Magris M, Hidalgo G, Baldassano RN, Anokhin AP, Heath AC, Warner B, Reeder J, Kuczynski J, Caporaso JG, Lozupone CA, Lauber C, Clemente JC, Knights D, Knight R, Gordon JI (2012) Human gut microbiome viewed across age and geography. Nature 486(7402):222–227. doi:10.1038/nature11053

Chapter 4
Azotobacter vinelandii Small RNAs: Their Roles in the Formation of Cysts and Other Processes

Miguel Castañeda, Liliana López-Pliego and Guadalupe Espín

Abstract *Azotobacter vinelandii* is a soil bacterium that undergoes differentiation to form cysts resistant to desiccation. Alginate, polyhydroxybutyrate (PHB) and alkylresorcinols (AR) are structural components of the cysts. The synthesis of these compounds has been shown to be under control of the Gac/Rsm signal transduction pathway. This pathway includes eight small non-coding RNAs (RsmZ1-7 and RsmY), and the translational repressor protein RsmA. Binding of RsmA to Rsm-RNAs inhibits its repressor activity. Here, we review and discuss the roles of sRNAs in the formation of cysts and other processes in *A. vinelandii.* We also report a search for genes encoding small RNAs in the *A. vinelandii* genome. Two new sRNAs potentially related to the control of cyst components synthesis were identified in this search.

4.1 Introduction

Azotobacter vinelandii is a soil nitrogen fixing soil bacterium related to *Pseudomonas* species to the extent, that it has been proposed to belong to this genus (Ozen and Ussery 2012; Rediers et al. 2004). Additionally 50 % of *A. vinelandii* genes encode proteins that show a similarity of 67 % or higher with a *Pseudomonas* homologues (Gonzalez-Casanova et al. 2014). In spite of close relationship, *A. vinelandii* shows important phenotypic features that are not observed in most *Pseudomonas* species, including the ability to fix nitrogen under aerobic conditions,

M. Castañeda · L. López-Pliego
Centro de Investigaciones en Ciencias Microbiologicas, Instituto de Ciencias, Benemerita Universidad Autónoma de Puebla, Apartado Postal 1622, 72000 Puebla, Mexico

G. Espín (✉)
Department of Molecular Microbiology, Instituto de Biotecnología, Universidad Nacional Autónoma de Mexico, Avenida Universidad 2001, 62210 Cuernavaca, Morelos, Mexico
e-mail: espin@ibt.unam.mx

A.L. Leitão and F.J. Enguita (eds.), *Non-coding RNAs and Inter-kingdom Communication*, DOI 10.1007/978-3-319-39496-1_4

a high respiration rate, and to undergo a differentiation process to form desiccation resistant cysts. *A. vinelandii* produce the exo-polysaccharide alginate, the intracellular polyester polyhydroxybutyrate (PHB), and the phenolic lipids alkylresorcinols (AR) and alkylpyrones (AP). Alginate, PHB, AR, and AP are structural components of cysts. Mature cysts consist of a cell known as central body, covered by a capsule composed of two layers named the intine and exine. Numerous PHB granules are always observed within the central body, alginate is a major component of the capsule, and AR and AP that are synthesized upon encystment induction replace the phospholipids of the cyst membranes and are components of the exine (Segura et al. 2014).

The expression of genes involved in the synthesis of cyst components mentioned above has been shown to be under the control of small non-coding RNAs (sRNAs) (Hernandez-Eligio et al. 2012; Manzo et al. 2011; Muriel-Millan et al. 2014; Romero et al. 2016). sRNAs are present in diverse organisms (i.e. bacteria, archaea and eukaryotes). In prokaryotes sRNAs have been linked to specialized housekeeping functions, to several stress situations, to pathogenicity and in catabolic carbon repression. The roles played by most sRNAs in the regulation of bacterial gene expression are exerted at a post-transcriptional level, where they activate or repress translation, or by stabilizing or destabilizing mRNAs (Waters and Storz 2009).

Four classical sRNAs, named Ffs, SsrA, RnpB SsrS play essential housekeeping roles and the genes encoding these RNAs have been identified in most bacterial genomes (Sonnleitner and Haas 2011). Their presence in *A. vinelandii* has not been reported.

Most sRNAs regulate gene expression at the level of translation mainly by two general mechanisms: The most common mechanism involves short, imperfect base-pairing interactions with the 5' end of mRNAs (Brantl 2002; Storz et al. 2011). The sRNA-mRNA base pairing mechanism requires a chaperon protein called Hfq. One example is the *Escherichia coli* RyhB RNA and *Pseudomonas* PrrF1-PrrF2 sRNAs (Masse and Gottesman 2002; Wilderman et al. 2004), These RNAs are involved in iron homeostasis, repressing the expression of some genes encoding nonessential proteins that use iron (Salvail and Masse 2012). RyhB regulates its targets by the formation of sRNA-mRNA complexes stabilized by the chaperone Hfq. A RhyB/PrrF homolog, known as ArrF (**A**zotobacter **r**egulatory RNA involving Fe) is present in the *A. vinelandii* genome, and is involved in iron homeostasis and also in the control of PHB synthesis (Jung and Kwon 2008; Muriel-Millan et al. 2014). Other sRNA that interacts with Hfq is RgsA (Park et al. 2013). Recently a homolog of RgsA has been characterized in *A. vinelandii*, in this bacterium RgsA is an important factor to resist oxidative stress, and plays a role in biofilm formation (Huerta et al. 2016).

The other regulatory mechanism involves binding of the sRNAs to a small protein that acts as translational repressor. One example involves the regulatory sRNAs named CsrB and CsrC in *E. coli* or RsmZ, RsmY and RsmX in *Pseudomonas* species. These sRNAs bind to the CsrA and RsmA proteins respectively. These proteins repress translation by binding their target mRNAs. The

proteins encoded in mRNAs that are targets the CsrA/RsmA, are involved in numerous and diverse functions (Lapouge et al. 2008). Seven *rsmZ* genes and one *rsmY* as well as the gene encoding the RsmA protein are present in the *A. vinelandii* genome. Although two RsmY RNAs: named RsmY1 and RsmY2 were first reported (Manzo et al. 2011; Setubal et al. 2009), they were later shown to constitute a single gene (Hernandez-Eligio et al. 2012). In *A. vinelandii* the Rsm-RNAs are involved in regulating the synthesis of the cyst components alginate, PHB and AR (Hernandez-Eligio et al. 2012; Manzo et al. 2011; Romero et al. 2016).

Other example of a regulatory mechanism involving sRNAs and a translation repressor protein has been reported in *Pseudomonas* species, where the sRNA named CrcZ is involved in carbon catabolic repression. Here the small protein Crc binds the CrcZ RNAs, and also to the 5' leader of mRNAs having the sequence GGG to repress their translation. Hfq is required for the binding of the small Crc protein to its target mRNAs or to CrcZ sRNA (for a Review see (Sonnleitner and Haas 2011)). Two copies of CrcZ are present in the *A. vinelandii* genome (Filiatrault et al. 2013).

Here, we present a review and a discussion of the mechanisms by which sRNAs control the synthesis of cyst components and report the results of a search for sRNAs encoded in the *A. vinelandii* genome.

4.2 Search of Genes Encoding sRNAs in the *A. vinelandii* Genome

In order to identify genes for sRNAs in the *A. vinelandii* DJ genome, we generated a database file of its intergenic regions. The intergenic sequences were analyzed with the INFERNAL algorithm using the Rfam database as comparison pattern (http://rfam.xfam.org/), using the tool "batch sequence search" of this server. The query intergenic sequences were obtained from the *A. vinelandii* DJ genome sequence (http://img.jgi.doe.gov/). The analysis identified 30 potential sRNAs. This search also revealed the presence of 10 riboswitches or *in cis* RNA element, 15 self-catalytic ribozymes, a CRISPR element, and 3 RNA thermometers (ROSE) element 3 (data not shown). The 30 sRNAs, their Rfam identifiers, and their position in the chromosome are listed in Table 4.1, and include the four sRNAs ubiquitous in bacteria: SsrS, SsrA, Ffs and RnpB (Sonnleitner and Haas 2011). SsrS or 6S is an RNA that specifically associates with RNA polymerase holoenzyme containing the sigma70 factor. This interaction represses expression from sigma70-dependent promoters during stationary phase (Wassarman 2007); SsrA, also known as 10Sa and tmRNA (transfer-messenger RNA), participates in a mechanism known as trans-translation. The tmRNA forms a ribonucleoprotein complex (tmRNP) together with Small Protein B (SmpB), Elongation Factor Tu (EF-Tu), and ribosomal protein S1. The tmRNA and its associated proteins bind to bacterial ribosomes, which have stalled in the middle of protein biosynthesis; it

recycles the stalled ribosome, adds a proteolysis-inducing tag to the unfinished polypeptide, and facilitates the degradation of the aberrant messenger RNA (Keiler 2015); Ffs is a 4.5S RNA which is a component of a signal recognition particle, that promotes the insertion of membrane proteins into the cytoplasmic membranes (Ulbrandt et al. 1997); RnpB is the 10Sb RNA subunit of the catalytic ribonuclease RNase P that is present in almost all living cells. Transfer RNA (tRNAs) genes are transcribed as precursors with extra residues at their 5' and 3' ends that have to be removed to generate functional tRNAs. In bacteria the RNAse P is responsible for removing the extra 5' residues (for a review see Kirsebom 2007). Due to its essential nature, it is expected that the functions of these RNAs in *A. vinelandii* is similar to those described above.

As expected, the previously identified and characterized RsmZ1-7, RsmY (Hernandez-Eligio et al. 2012; Manzo et al. 2011), the ArrF (Jung and Kwon 2008), the RgsA (Huerta et al. 2016), and the two CrcZ sRNAs were found in this search and are listed in Table 4.1. Interestingly a new copy of *rsmZ*, here named *rsmZ8* was also identified, raising to nine the number of Rsm-RNAs present in *A. vinelandii*.

Homologs to the *P. aeruginosa* P11, P14, P15, P20 (PhrS), P26, P27, P31 sRNAs, that were predicted by the bioinformatic program SRNApredict2, and confirmed by Northern blot analysis in this bacterium (Livny et al. 2006), were identified in the *A. vinelandii* genome Table 4.1. P20 (PhrS) promotes the expression of *pqsR*, an important regulator to establish the quorum sensing mediated by quinones in *P. aeruginosa* (Sonnleitner and Haas 2011). As *A. vinelandii* does not possess quorum-sensing systems related with quinones, PhrS must have a different function in this bacterium. The functions of P11, P14, P15, P26, P27, P31 are unknown.

An sRNA homologue to Spot 42 sRNA of *E. coli* (Beisel and Storz 2011) is present in *P. aeruginosa* where it was named Pseudomon 1. Recently it was renamed as ErsA and was shown to regulate the expression of the alginate, biosynthetic gene *algC* at a posttranscriptional level (Ferrara et al. 2015).

IsrK is present in *Salmonella typhimurium*, the gene encoding this sRNA is located within a virulence island and is a host-induced gene (Padalon-Brauch et al. 2008).

Finally, the other sRNAs found in *A. vinelandii* are; a homologue of t44 of unknown function, found in enteric bacteria and in *Pseudomonas* (Tjaden et al. 2002), and two C4 RNAs. C4 antisense molecules are related with the presence of P1 and P7 temperate phages in bacteria (Citron and Schuster 1990).

Several of the sRNAs listed in the Table 1 such as IsrK, RgsA, Pseudomon-1 (EsrA), CsrZ and PhrS have been shown to interact and mediate its function with the chaperone Hfq (Ferrara et al. 2015; Park et al. 2013; Sonnleitner and Blasi 2014; Sonnleitner and Haas 2011; Zhang et al. 2003). Therefore, in *A. vinelandii* Hfq the predicted product of gene Avin7540 could mediate the activity of the same RNAs.

4.3 RsmZ and RsmY SRNAs Involved in Synthesis of Cyst Components and Cyst Resistance to Desiccation

In many γ-proteobacteria the conserved two-component GacS-GacA system (named BarA-UvrY in *E. coli*) controls the expression of one or more genes encoding small RNAs. These RNAs bind small proteins named RsmA or CsrA that act as translational repressors. Thus the GacS-GacA system controls the post-transcriptional regulatory system that involves the translation repressor protein RsmA (CsrA) and the sRNAs RsmZ, RsmY (CsrB) (Lapouge et al. 2008). As in other γ-proteobacteria in *A. vinelandii* GacA the protein acting as a transcriptional activator is required for the transcription of the genes encoding the eight Rsm-sRNAs (RsmZ1, RsmZ2, RsmZ3, RsmZ4, RsmZ5, RsmZ6, RsmZ7 and RsmY (Hernandez-Eligio et al. 2012; Manzo et al. 2011). In agreement with this requirement, the Rsm-sRNAs genes possess a GacA binding box within their regulatory region (Manzo et al. 2011). In the search carried out in this study, we found an additional *rsmZ* gene (*rsmZ8*). Although the nucleotide sequence of RsmZ8 is the least conserved among the RsmZs, its secondary structure is well conserved (data not shown). An inspection of the regulatory region of *rsmZ8* revealed the presence of GacA binding sites (Table 4.2).

4.3.1 Rsm-RNAs and Alginate Production

In *A. vinelandii* as in *Pseudomonas* species the genes encoding the enzymes that participate in alginate biosynthesis (with the exception of *algC* which encodes for the second enzyme of the pathway) constitute a cluster headed by *algD*, which encodes GDP-mannose dehydrogenase, the key enzyme of the pathway (Campos et al. 1996). Alginate is essential for cyst formation as mutants unable to produce this polysaccharide are impaired in cyst development (Campos et al. 1996; Mejia-Ruiz et al. 1997). Early, we reported that inactivation of the *gacS* or *gacA* gene impaired alginate synthesis (Castañeda et al. 2000, 2001).

The RsmZ1 and RsmZ2 RNAs were shown to interact with the RsmA protein by RNA gel mobility shift experiments, and RsmA was shown to bind to a 200 nt transcript, corresponding to the 5' leader of the *algD* mRNA that includes a putative RsmA binding site (Manzo et al. 2011). Thus, It was concluded that in the *gacA* mutant the absence of Rsm RNAs allows, the free RsmA protein to bind the *algD* mRNA inhibiting its translation. Indeed, transcription of a single Rsm RNA (RsmZ1) from a GacA-independent promoter, restored alginate synthesis to the *gacA* mutant (Manzo et al. 2011). These studies clearly indicated the important role of the Rsm-RNAs acting as RsmA anti-repressors to allow translation of *algD* and the synthesis of alginate an essential component of the cyst capsule.

4.3.2 Rsm-RNAs and Alkylresorcinols Production

In *A. vinelandii* alkylresorcinols replace the phospholipids in the cysts membranes and are components of the exine, the outer layer of the cyst envelope. Mutants impaired in AR synthesis produce cysts with a defective exine (Segura et al. 2009). The biosynthesis of alkylresorcinols is also under the control of the Gac-Rsm system. This control is exerted at the level of *arpR* expression. ArpR is the transcriptional activator of the alkylresorcinol biosynthetic operon *arsABCD* (Romero et al. 2013). Inactivation of *gacA* impaired the synthesis of AR, and mutations in *rsmZ1* diminished the synthesis of these phenolic lipids. Additionally the RsmA protein was shown to bind to a transcript corresponding to the 5' leader of the *arpR* mRNA (Romero et al. 2016). Taken together these data indicated the essentiality of Rsm RNAs to release the repression of *arpR* mRNA translation exerted by RsmA.

4.3.3 Rsm-RNAs and PHB Production

Numerous PHB granules are present in mature cysts, and mutants unable to synthesize PHB produce cysts that lack PHB granules but resist desiccation (Segura et al. 2003) indicating that PHB is not essential for cyst formation or germination. However, this result does not rule out the possibility that under natural conditions PHB plays an important role in the formation and/or survival of cysts. Encystment in laboratory conditions is induced in Burk's medium containing 0.2 % of γ-hydroxybutyrate as sole carbon source, thus it is likely that in soil, this compound is provided by the degradation of PHB. Similar to alginate and AR genes, the expression of the PHB biosynthetic and regulatory genes is controlled by the Gac-Rsm pathway, as inactivation of *gacA* impaired PHB synthesis. The PHB biosynthetic operon *phbBAC* is activated by the transcriptional activator PhbR. RsmA was shown to interact with RNAs corresponding to the *phbB* and *phbR* mRNA leaders. Additionally the stability of *phbB* and *phbR* transcripts is increased in a *rsmA* mutant strain (Hernandez-Eligio et al. 2012). Therefore as in the case of alginate, and AR the Rsm RNAs are also required to antagonize the repressor activity of RsmA on translation of *phbR* and *phbB* mRNAs.

4.3.4 Expression of Hsp20 an Essential Protein for Cyst Resistance to Desiccation Is Regulated at a Posttranscriptional Level

The small heat shock protein Hsp20 is one of the most abundant proteins in *A. vinelandii* cysts and is present in low levels in vegetative cells (Cocotl-Yañez et al. 2014). However, transcription of *hsp20* is higher in vegetative cells than in cysts,

Table 4.1 Genes encoding sRNA found in silico in *A. vinelandii*

sRNA	Rfam identifiers	Localization	Reference
Classical sRNAs			
6S (ssrS)	RF00013	Intergenic 4809934..4810217	Wassarman (2007)
tmRNA (10Sa, *ssrA*)	RF00023	Intergenic 4355922..4356427	Keiler (2015)
Bacteria_small_SRP (4.5Sm *ffs*)	RF00169	Intergenic 1969172..1970650	Siu et al. (2007)
RNase P RNA (P28, 10Sb, *rnpB*)	RF00010	Intergenic 1278334..1279056	Kirsebom (2007)
sRNAs that sequester RNA-binding proteins			
PrrB_RsmZ (RsmZ1)	RF00166	Intergenic 385392..385847	Manzo et al. (2011)
PrrB_RsmZ (RsmZ2)		Intergenic 1671372..1672144	Manzo et al. (2011)
PrrB_RsmZ (RsmZ3)		Intergenic 396386..397732	Manzo et al. (2011)
PrrB_RsmZ (RsmZ4)		Intergenic 399749..401572	Manzo et al. (2011)
PrrB_RsmZ (RsmZ5)		Intergenic 422104..422515	Manzo et al. (2011)
PrrB_RsmZ (RsmZ6)		Intergenic 1692051..1692868	Manzo et al. (2011)
PrrB_RsmZ (RsmZ7)		Intergenic 843721..844218	Manzo et al. (2011)
PrrB_RsmZ (RsmZ8)		Intergenic 116768..117043	This work
RsmY	RF00195	Intergenic 444526..445166	Manzo et al. (2011)
CrcZ (CrcZ1)	RF01675	Intergenic 4303112..4303775	Filiatrault (2013)
CrcZ (CrcZ2)		Intergenic 452629..452762	Filiatrault (2013)
Base pairing sRNAs			
Pseudomon-1 (Spot 42, ErsA)	RF01719	Intergenic 64850..65144	Ferrara et al. (2015)
P20 (PhrS)	RF01673	Intergenic 3126790..3126965	Sonnleitner et al. (2011)
Prrf1 (ArrF)	RF00444	Avin 42580	Muriel-Millan et al. (2014)
Isrk	RF01394	Intergenic 3302323..3303960	Padalon-Brauch et al. (2008)
C4	RF01695	Intergenic 3302323..3303960	Citron and Schuster (1990)

(continued)

Table 4.1 (continued)

sRNA	Rfam identifiers	Localization	Reference
C4	–	Intergenic 4107181..4107534	–
RgsA (P16)	RF00628	Intergenic 1479161..1479523	Huerta et al. 2016
Others sRNAs			
T44	RF00127	Intergenic 3946329..3946592	Tjaden et al. (2002)
P11	RF00625	Intergenic 2336465..2336727	Livny et al. (2006)
P14	RF01669	Intergenic 2306571..2306897	Livny et al. (2006)
P15	RF00627	Avin 33220	Livny et al. (2006)
P26	RF00630	Intergenic 593666..593879	Livny et al. (2006)
P27	RF01674	Intergenic 592518..592714	Livny et al. (2006)
P31	RF01676	Intergenic 4335150..4335335	Livny et al. (2006)

but translation is much higher under encystment conditions than in vegetative cells, indicating a post-transcriptional regulation mechanism. Indeed the *hsp20* Shine-Dalgarno possesses a potential RsmA binding site that matched the RsmA binding sites present in the *algD*, *phbR,* and *arpR* transcripts (Cocotl-Yañez et al. 2014), suggesting that expression of Hsp20 is under Rsm-RNAs control. Mutations of *hsp20* did not affect the synthesis of cyst components nor the morphology of the cyst, however wild type-like cysts formed by the *hsp20* mutant were unable to resist desiccation (Cocotl-Yañez et al. 2014).

By the mean of allowing the expression of genes involved in the synthesis of cyst component alginate, AR and PHB as well as the *hsp20* gene essential for cyst desiccation resistance, the Rsm-RNAs play an important role in the differentiation process to produce cysts. A model for the control of cyst formation by the Rsm-RNAs is shown in Fig. 4.1.

4.3.5 Role of Rsm-RNAs Reiterated Genes

The presence of multiple copies of Rsm-RNAs in *A. vinelandii* raises the question about their biological role. One possibility is to provide with a range of levels of these regulatory RNAs (doses). This range could be achieved by differential expression under different conditions. This in turn will result in a wide range of levels of the free RsmA repressor. Indeed although *rsmZ1-7* and *rsmY* genes and

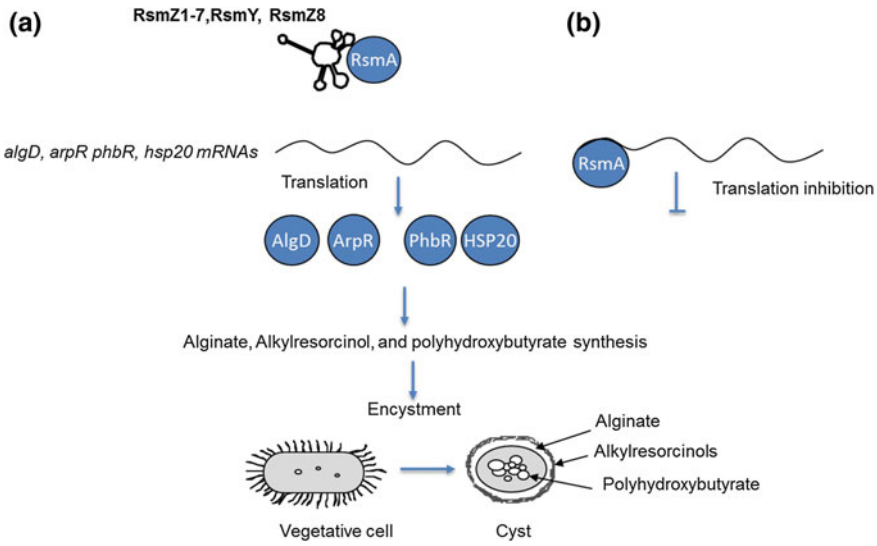


Fig. 4.1 Model for the control of cyst formation by the Rsm-RNAs. **a** The binding of RsmA to the Rsm-RNAs inhibits its translation repression function, allowing expression of the AlgD, ArpR PhbR and Hsp20 proteins, the synthesis of alginate, alkylresorcinol and PHB proteins, and the formation of cysts. **b** Binding of RsmA to the 5' UTRs of the *algD*, *arpR*, *phbR* and *hsp20* mRNAs inhibits their translation

probably *rsmZ8*, are all activated by GacA, their expression levels varies among them (Hernandez-Eligio et al. 2012). This could be related to the position of the GacA binding sequences, respect to the transcription start site, as well as the % of AT (Table 4.2). In fact, regulators with affinity to AT rich sequences like IHF and H-NS have been related with differential expression of Rsm-RNAs genes in *Pseudomonas* species (Brencic et al. 2009; Humair et al. 2010).

The effect of the RNA structure-affinity on the binding to the protein RsmA can also provide the cells with an even higher range of free RsmA levels. In a model of RsmA (CsrA)-RNAs interaction, the greater number of stem-loop structures, the more RsmA molecules sequestered is expected (Babitzke and Romeo 2007; Heroven et al. 2012). Indeed in the *A. vinelandii* Rsm-RNAs the numbers of stem-loops structures vary from four to nine (Fig. 4.2, Table 4.2). Differences in their ΔG, and therefore their stability are also predicted (Table 4.2). Other factor affecting the level of Rsm RNAs is the turn over, as their enzymatic degradation is related to their structure (Seyll and Van Melderen 2013).

In summary, we propose that the redundancy of Rsm-RNAs allows the cell to have a wide range of RsmA levels in response to different environmental or metabolic signals including those that induce differentiation. In fact, redundancy in *A. vinelandii* is higher than in any *Pseudomonas* species, where the number of Rsm-RNAs varies from two to five (Moll et al. 2010). Thus, the presence of nine

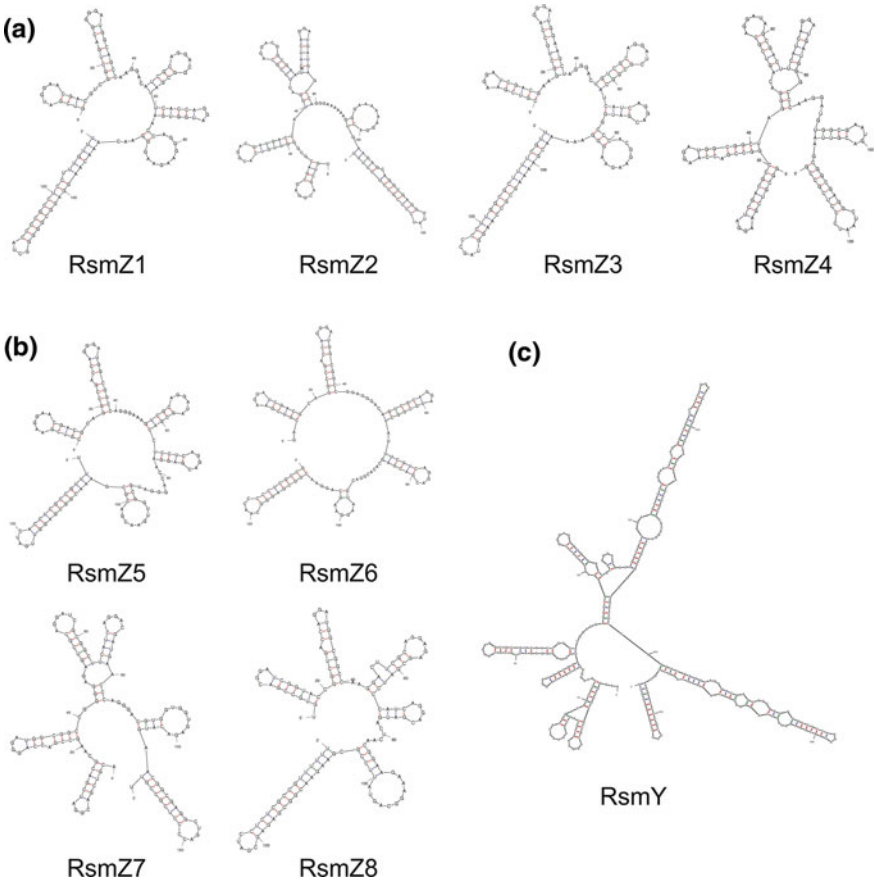


Fig. 4.2 Stem-loop structures of *A. vinelandii* Rsm small RNAs. **a** RsmZ1, RsmZ2, RsmZ3 and RsmZ4. **b** RsmZ5, RsmZ6, RsmZ7 and RsmZ8. **c** RsmY

Table 4.2 *A. vinelandii* Rsm-RNAs properties and characteristics

Rsm sRNA	Number of stem-loops with GGA motif	ΔG (kcal/mol)	GacA box position (pb)	AT content (%) of rsm genes regulatory regions
RsmZ1	5	−61.85	−175, −155	45
RsmZ2	5	−50.36	−181, −163	50
RsmZ3	5	−68.04	−82, −64	40
RsmZ4	4	−65.40	−80, −62	42
RsmZ5	4	−66.05	−80, −62	36
RsmZ6	4	−60.34	−81, −63	39
RsmZ7	5	−63.26	−80, −62	33
RsmZ8	4	−65.27	−75, −57	44
RsmY	9	−148.90	−75, −57	37

Rsm-RNAs in *A. vinelandii* could be related to their involvement in cyst formation, a feature not observed in its close relatives, the *Pseudomonas*.

4.4 ArrF Controls Iron Homeostasis and PHB Synthesis in *A. vinelandii*

Iron acquisition is controlled by the Fur (**F**erric **u**ptake **r**epressor) protein, which represses the transcription of many genes, including those associated with iron uptake, like siderophores and receptors (Escolar et al. 1999). Under high concentrations of iron, Fur forms a complex with Fe^{2+}, and binds to "iron boxes" present in its target genes, repressing their transcription. When iron is limited, Fur is unable to bind iron boxes. Fur also represses the transcription of RyhB in *E. coli*, and of ArrF in *A. vinelandii* (Jung and Kwon 2008; Masse and Gottesman 2002). RhyB represses the expression of genes encoding non-essential proteins that use iron (Salvail and Masse 2012). RhyB sRNA forms sRNA-mRNA complexes stabilized by Hfq. RhyB also exerts a positive regulation on the expression of *shiA*. In this case the rhyB-mRNA pairing disrupt an inhibitory structure allowing translation (Prevost et al. 2007).

Iron is essential for different biological processes in *A. vinelandii* including nitrogen fixation, respiration, and nitrogenase protection (Page and Huyer 1984; Page and von Tigerstrom 1982). ArrF was shown to negatively regulate the expression of *sodB* gene encoding Fe-containing superoxide dismutase and FeSII protein (Jung and Kwon 2008), which is known to mediate the conformational protection of nitrogenase enzyme against oxygen inactivation. Iron also affects PHB accumulation in this bacterium (Pyla et al. 2009). Under Iron limitation transcription of *arrF* increased significantly. The expression of PhbR, the transcriptional activator of PHB synthesis also increased post- transcriptionally in iron limitation, suggesting that iron limitation increases translation of *phbR* through ArrF (Muriel-Millan et al. 2014). The study of the role of ArrF in the control of PHB was carried out in vegetative cells. Thus its role in encystment remains to be investigated.

4.5 Roles of EsrA sRNA in *Pseudomonas*

In *P. aeruginosa* and *A. vinelandii* the alginate biosynthetic, gene *algC* is transcribed from a promoter dependent on AlgU, a homologue of the RpoE alternative sigma factor (Gaona et al. 2004; Zielinski et al. 1991). Therefore, inactivation of *algU* in *A. vinelandii* abrogates alginate synthesis and encystment (Moreno et al. 1998). In *P. aeruginosa* the EsrA RNA regulates expression of AlgC at a post-transcriptional level and, *esrA* as *algC* is transcribed from a AlgU promoter (Ferrara et al. 2015). As shown in Table 1 an EsrA homolog sharing 78 % of identity with *Pseudomonas esrA* is present in *A. vinelandii*. Interestingly conserved −10 and −35

boxes that correspond to an AlgU promoter are present in the regulatory region of the *A. vinelandii esrA*. These observations suggest that in *A. vinelandii* ErsA sRNA may be involved in the control of alginate synthesis.

4.6 CrcZ and Its Role in Carbon Catabolic Repression in *Pseudomonas* Species

Carbon catabolic repression is a mechanism that allows bacteria to catabolize preferred carbon sources for most efficient growth. In enteric bacteria this feature is controlled by the phosphoenolpyruvate (PEP): carbohydrate phosphotransferase system (carbohydrate PTS) by the mean of controlling cAMP-dependent transcriptional activation (for a review see Lengeler and Jahreis 2009).

In *Pseudomonas* and in *A. vinelandii* PTS plays no role in catabolic repression (Filiatrault et al. 2013; Moreno et al. 2015). Instead, in *Pseudomonas* species *aeruginosa*, *putida*, and *syringae* the control of carbon catabolic repression involves the two-component system CbrBA, the sRNAs CrcX, CrcY or CrcZ, and the Crc protein. The expression of the genes encoding Crc sRNAs is dependent on the alternative sigma factor RpoN and the CbrB transcriptional activator. Crc is a small protein that recognizes and bind mRNAs encoding proteins involved in catabolic pathways of less preferred carbon sources. This binding results in repression of translation. In the presence of Hfq, CrcZ sRNA binds the translation repressor Crc. When *Pseudomonas* grows on a less preferred carbon source, CbrB activates transcription of CrcZ that sequesters the Crc protein allowing translation of mRNAs that are Crc targets. On the other side, in the presence of a preferred source, CbrAB activity is inhibited, resulting in low levels of CrcZ sRNA and repression of mRNA by Crc.

The *crbAB*, *crc* genes, and two copies encoding sRNA CrcZ are present in the *A. vinelandii* genome (Filiatrault et al. 2013; Setubal et al. 2009). The high conservation of the catabolic repression CbrAB/Crc system genes between *Pseudomonas* and *A. vinelandii* suggest the same function in this bacterium. If this was the case, The CbrAB/Crc system is likely to have an effect on the synthesis of alginate and PHB as the precursors for their synthesis (fructose 6P and acetyl-CoA respectively) are originated from the catabolism of the carbon source provided.

4.7 Outlook

Several important aspects of the encystment process that undergoes *A. vinelandii*, such as the control of gene expression that leads to the synthesis of important cyst components have been shown to be under Rsm-RNAs control. However, more details about their individual expression and their interaction with RsmA remain to be investigated. We do not rule out that in addition to the Rsm-RNAs other RNAs

like EsrA a regulator of the alginate biosynthetic, gene *algC* in *P. aeruginosa* also participate in the control of alginate synthesis, an essential component of mature cysts.

Another important subject that remains to be investigated is the behavior of *A. vinelandii* cysts or vegetative cells in natural soil conditions. It is well known that bacteria are killed by viruses (bacteriophages), and are predated by other bacteria and eukaryotic organisms as protozoan, amoebas and fungi. Bacteria has developed anti-depredator strategies, one of them is the synthesis of exo-polysaccharides, as mutant strains impaired in exo-polysaccharide production are more susceptible to bacteriovores (Matz et al. 2004). Whether alginate producing, cells of *A. vinelandii* are more resistant to predators than alginate minus mutants remains to be investigated. However, It has been reported that *A. vinelandii* cysts (that possess an alginate capsule), are more resistant to predation by *Agromyces ramosus* than vegetative cells (Casida 1983). Whether cysts are also more resistant to eukaryotic predators like protozoan remains to be tested. Thus, in *A. vinelandii* sRNAs are essential for the production of exo-polysaccharide and for cyst formation, two features that may play essential roles in the survival of this bacterium in soil.

References

Babitzke P, Romeo T (2007) CsrB sRNA family: sequestration of RNA-binding regulatory proteins. Curr Opin Microbiol 10(2):156–163. doi:10.1016/j.mib.2007.03.007

Beisel CL, Storz G (2011) The base-pairing RNA spot 42 participates in a multioutput feedforward loop to help enact catabolite repression in *Escherichia coli*. Mol Cell 41(3):286–297. doi:10.1016/j.molcel.2010.12.027

Brantl S (2002) Antisense-RNA regulation and RNA interference. Biochim Biophys Acta 1575 (1–3):15–25

Brencic A, McFarland KA, McManus HR, Castang S, Mogno I, Dove SL, Lory S (2009) The GacS/GacA signal transduction system of *Pseudomonas aeruginosa* acts exclusively through its control over the transcription of the RsmY and RsmZ regulatory small RNAs. Mol Microbiol 73(3):434–445. doi:10.1111/j.1365-2958.2009.06782.x

Campos M, Martinez-Salazar JM, Lloret L, Moreno S, Nunez C, Espin G, Soberon-Chavez G (1996) Characterization of the gene coding for GDP-mannose dehydrogenase (algD) from *Azotobacter vinelandii*. J Bacteriol 178(7):1793–1799

Casida LE (1983) Interaction of *Agromyces ramosus* with Other Bacteria in Soil. Appl Environ Microbiol 46(4):881–888

Castañeda M, Guzman J, Moreno S, Espin G (2000) The GacS sensor kinase regulates alginate and poly-beta-hydroxybutyrate production in *Azotobacter vinelandii*. J Bacteriol 182(9):2624–2628

Castañeda M, Sanchez J, Moreno S, Nunez C, Espin G (2001) The global regulators GacA and sigma(S) form part of a cascade that controls alginate production in *Azotobacter vinelandii*. J Bacteriol 183(23):6787–6793. doi:10.1128/JB.183.23.6787-6793.2001

Citron M, Schuster H (1990) The c4 repressors of bacteriophages P1 and P7 are antisense RNAs. Cell 62(3):591–598

Cocotl-Yañez M, Moreno S, Encarnacion S, Lopez-Pliego L, Castañeda M, Espin G (2014) A small heat-shock protein (Hsp20) regulated by RpoS is essential for cyst desiccation resistance in *Azotobacter vinelandii*. Microbiology 160(Pt 3):479–487. doi:10.1099/mic.0.073353-0

Escolar L, Perez-Martin J, de Lorenzo V (1999) Opening the iron box: transcriptional metalloregulation by the Fur protein. J Bacteriol 181(20):6223–6229

Ferrara S, Carloni S, Fulco R, Falcone M, Macchi R, Bertoni G (2015) Post-transcriptional regulation of the virulence-associated enzyme AlgC by the sigma(22)-dependent small RNA ErsA of *Pseudomonas aeruginosa*. Environ Microbiol 17(1):199–214. doi:10.1111/1462-2920.12590

Filiatrault MJ, Stodghill PV, Wilson J, Butcher BG, Chen H, Myers CR, Cartinhour SW (2013) CrcZ and CrcX regulate carbon source utilization in *Pseudomonas syringae* pathovar tomato strain DC3000. RNA Biol 10(2):245–255. doi:10.4161/rna.23019

Gaona G, Nunez C, Goldberg JB, Linford AS, Najera R, Castañeda M, Guzman J, Espin G, Soberon-Chavez G (2004) Characterization of the *Azotobacter vinelandii* algC gene involved in alginate and lipopolysaccharide production. FEMS Microbiol Lett 238(1):199–206. doi:10.1016/j.femsle.2004.07.044

Gonzalez-Casanova A, Aguirre-von-Wobeser E, Espin G, Servin-Gonzalez L, Kurt N, Spano D, Blath J, Soberon-Chavez G (2014) Strong seed-bank effects in bacterial evolution. J Theor Biol 356:62–70. doi:10.1016/j.jtbi.2014.04.009

Hernandez-Eligio A, Moreno S, Castellanos M, Castañeda M, Nunez C, Muriel-Millan LF, Espin G (2012) RsmA post-transcriptionally controls PhbR expression and polyhydroxybutyrate biosynthesis in *Azotobacter vinelandii*. Microbiology 158(Pt 8):1953–1963. doi:10.1099/mic.0.059329-0

Heroven AK, Bohme K, Dersch P (2012) Regulation of virulence gene expression by regulatory RNA elements in *Yersinia pseudotuberculosis*. Adv Exp Med Biol 954:315–323. doi:10.1007/978-1-4614-3561-7_39

Huerta JM, Aguilar I, Lopez-Pliego L, Fuentes-Ramirez LE, Castañeda M (2016) The Role of the ncRNA RgsA in the Oxidative Stress Response and Biofilm Formation in *Azotobacter vinelandii*. Curr Microbiol. doi:10.1007/s00284-016-1003-2

Humair B, Wackwitz B, Haas D (2010) GacA-controlled activation of promoters for small RNA genes in *Pseudomonas fluorescens*. Appl Environ Microbiol 76(5):1497–1506. doi:10.1128/AEM.02014-09

Jung YS, Kwon YM (2008) Small RNA ArrF regulates the expression of sodB and feSII genes in *Azotobacter vinelandii*. Curr Microbiol 57(6):593–597. doi:10.1007/s00284-008-9248-z

Keiler KC (2015) Mechanisms of ribosome rescue in bacteria. Nat Rev Microbiol 13(5):285–297. doi:10.1038/nrmicro3438

Kirsebom LA (2007) RNase P RNA mediated cleavage: substrate recognition and catalysis. Biochimie 89(10):1183–1194. doi:10.1016/j.biochi.2007.05.009

Lapouge K, Schubert M, Allain FH, Haas D (2008) Gac/Rsm signal transduction pathway of gamma-proteobacteria: from RNA recognition to regulation of social behaviour. Mol Microbiol 67(2):241–253. doi:10.1111/j.1365-2958.2007.06042.x

Lengeler JW, Jahreis K (2009) Bacterial PEP-dependent carbohydrate: phosphotransferase systems couple sensing and global control mechanisms. Contrib Microbiol 16:65–87. doi:10.1159/000219373

Livny J, Brencic A, Lory S, Waldor MK (2006) Identification of 17 *Pseudomonas aeruginosa* sRNAs and prediction of sRNA-encoding genes in 10 diverse pathogens using the bioinformatic tool sRNAPredict2. Nucleic Acids Res 34(12):3484–3493. doi:10.1093/nar/gkl453

Manzo J, Cocotl-Yañez M, Tzontecomani T, Martinez VM, Bustillos R, Velasquez C, Goiz Y, Solis Y, Lopez L, Fuentes LE, Nunez C, Segura D, Espin G, Castañeda M (2011) Post-transcriptional regulation of the alginate biosynthetic gene algD by the Gac/Rsm system in *Azotobacter vinelandii*. J Mol Microbiol Biotechnol 21(3–4):147–159. doi:10.1159/000334244

Masse E, Gottesman S (2002) A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. Proc Natl Acad Sci USA 99(7):4620–4625. doi:10.1073/pnas.032066599

Matz C, Bergfeld T, Rice SA, Kjelleberg S (2004) Microcolonies, quorum sensing and cytotoxicity determine the survival of *Pseudomonas aeruginosa* biofilms exposed to protozoan grazing. Environ Microbiol 6(3):218–226

Mejia-Ruiz H, Moreno S, Guzman J, Najera R, Leon R, Soberon-Chavez G, Espin G (1997) Isolation and characterization of an *Azotobacter vinelandii* algK mutant. FEMS Microbiol Lett 156(1):101–106

Moll S, Schneider DJ, Stodghill P, Myers CR, Cartinhour SW, Filiatrault MJ (2010) Construction of an rsmX co-variance model and identification of five rsmX non-coding RNAs in *Pseudomonas syringae* pv. tomato DC3000. RNA Biol 7(5):508–516

Moreno R, Hernandez-Arranz S, La Rosa R, Yuste L, Madhushani A, Shingler V, Rojo F (2015) The Crc and Hfq proteins of *Pseudomonas putida* cooperate in catabolite repression and formation of ribonucleic acid complexes with specific target motifs. Environ Microbiol 17 (1):105–118. doi:10.1111/1462-2920.12499

Moreno S, Najera R, Guzman J, Soberon-Chavez G, Espin G (1998) Role of alteRNAtive sigma factor algU in encystment of *Azotobacter vinelandii*. J Bacteriol 180(10):2766–2769

Muriel-Millan LF, Castellanos M, Hernandez-Eligio JA, Moreno S, Espin G (2014) Posttranscriptional regulation of PhbR, the transcriptional activator of polyhydroxybutyrate synthesis, by iron and the sRNA ArrF in *Azotobacter vinelandii*. Appl Microbiol Biotechnol 98 (5):2173–2182. doi:10.1007/s00253-013-5407-7

Ozen AI, Ussery DW (2012) Defining the *Pseudomonas genus*: where do we draw the line with Azotobacter? Microb Ecol 63(2):239–248. doi:10.1007/s00248-011-9914-8

Padalon-Brauch G, Hershberg R, Elgrably-Weiss M, Baruch K, Rosenshine I, Margalit H, Altuvia S (2008) Small RNAs encoded within genetic islands of *Salmonella typhimurium* show host-induced expression and role in virulence. Nucleic Acids Res 36(6):1913–1927. doi:10.1093/nar/gkn050

Page WJ, Huyer M (1984) Derepression of the *Azotobacter vinelandii* siderophore system, using iron-containing minerals to limit iron repletion. J Bacteriol 158(2):496–502

Page WJ, von Tigerstrom M (1982) Iron- and molybdenum-repressible outer membrane proteins in competent *Azotobacter vinelandii*. J Bacteriol 151(1):237–242

Park SH, Butcher BG, Anderson Z, Pellegrini N, Bao Z, D'Amico K, Filiatrault MJ (2013) Analysis of the small RNA P16/RgsA in the plant pathogen *Pseudomonas syringae* pv. tomato strain DC3000. Microbiology 159 (Pt 2):296–306. doi:10.1099/mic.0.063826-0

Prevost K, Salvail H, Desnoyers G, Jacques JF, Phaneuf E, Masse E (2007) The small RNA RyhB activates the translation of shiA mRNA encoding a permease of shikimate, a compound involved in siderophore synthesis. Mol Microbiol 64(5):1260–1273. doi:10.1111/j.1365-2958.2007.05733.x

Pyla R, Kim TJ, Silva JL, Jung YS (2009) Overproduction of poly-beta-hydroxybutyrate in the *Azotobacter vinelandii* mutant that does not express small RNA ArrF. Appl Microbiol Biotechnol 84(4):717–724. doi:10.1007/s00253-009-2002-z

Rediers H, Vanderleyden J, De Mot R (2004) *Azotobacter vinelandii*: a Pseudomonas in disguise? Microbiology 150(Pt 5):1117–1119. doi:10.1099/mic.0.27096-0

Romero Y, Guzman J, Moreno S, Cocotl M, Espin G, Castañeda M, Vences M, Segura D (2016) The GacS/A-RsmA signal transduction pathway controls the synthesis of alkylresorcinol lipids that replace membrane phospholipids during encystment of *Azotobacter vinelandii* SW136. PLoS ONE 11, e0153266. doi:10.1371/journal.pone.0153266

Romero Y, Moreno S, Guzman J, Espin G, Segura D (2013) Sigma factor RpoS controls alkylresorcinol synthesis through ArpR, a LysR-type regulatory protein, during encystment of *Azotobacter vinelandii*. J Bacteriol 195(8):1834–1844. doi:10.1128/JB.01946-12

Salvail H, Masse E (2012) Regulating iron storage and metabolism with RNA: an overview of posttranscriptional controls of intracellular iron homeostasis. Wiley Interdiscip Rev RNA 3 (1):26–36. doi:10.1002/wrna.102

Segura D, Cruz T, Espin G (2003) Encystment and alkylresorcinol production by *Azotobacter vinelandii* strains impaired in poly-beta-hydroxybutyrate synthesis. Arch Microbiol 179 (6):437–443. doi:10.1007/s00203-003-0553-4

Segura D, Nunez C, Espín G (2014) *Azotobacter* Cysts. In: Ltd JWS (ed) Encyclopedia of life sciences. Chichester, West Sussex, UK. doi:10.1002/9780470015902.a0000295.pub2

Segura D, Vite O, Romero Y, Moreno S, Castañeda M, Espin G (2009) Isolation and characterization of *Azotobacter vinelandii* mutants impaired in alkylresorcinol synthesis: alkylresorcinols are not essential for cyst desiccation resistance. J Bacteriol 191(9):3142–3148. doi:10.1128/JB.01575-08

Setubal JC, dos Santos P, Goldman BS, Ertesvag H, Espin G, Rubio LM, Valla S, Almeida NF, Balasubramanian D, Cromes L, Curatti L, Du Z, Godsy E, Goodner B, Hellner-Burris K, HeRNAndez JA, Houmiel K, Imperial J, Kennedy C, Larson TJ, Latreille P, Ligon LS, Lu J, Maerk M, Miller NM, Norton S, O'Carroll IP, Paulsen I, Raulfs EC, Roemer R, Rosser J, Segura D, Slater S, Stricklin SL, Studholme DJ, Sun J, Viana CJ, Wallin E, Wang B, Wheeler C, Zhu H, Dean DR, Dixon R, Wood D (2009) Genome sequence of *Azotobacter vinelandii*, an obligate aerobe specialized to support diverse anaerobic metabolic processes. J Bacteriol 191(14):4534–4545. doi:10.1128/JB.00504-09

Siu FY, Spanggord RJ, Doudna JA (2007) SRP RNA provides the physiologically essential GTPase activation function in cotranslational protein targeting. RNA 13(2):240-250. doi:10.1261/rna.13540

Seyll E, Van Melderen L (2013) The ribonucleoprotein Csr network. Int J Mol Sci 14(11):22117–22131. doi:10.3390/ijms141122117

Sonnleitner E, Blasi U (2014) Regulation of Hfq by the RNA CrcZ in *Pseudomonas aeruginosa* carbon catabolite repression. PLoS Genet 10(6):e1004440. doi:10.1371/journal.pgen.1004440

Sonnleitner E, Haas D (2011) Small RNAs as regulators of primary and secondary metabolism in *Pseudomonas* species. Appl Microbiol Biotechnol 91(1):63–79. doi:10.1007/s00253-011-3332-1

Storz G, Vogel J, Wassarman KM (2011) Regulation by small RNAs in bacteria: expanding frontiers. Mol Cell 43(6):880–891. doi:10.1016/j.molcel.2011.08.022

Tjaden B, Saxena RM, Stolyar S, Haynor DR, Kolker E, Rosenow C (2002) Transcriptome analysis of *Escherichia coli* using high-density oligonucleotide probe arrays. Nucleic Acids Res 30(17):3732–3738

Ulbrandt ND, Newitt JA, Bernstein HD (1997) The *E. coli* signal recognition particle is required for the insertion of a subset of inner membrane proteins. Cell 88(2):187–196

Wassarman KM (2007) 6S RNA: a regulator of transcription. Mol Microbiol 65(6):1425–1431. doi:10.1111/j.1365-2958.2007.05894.x

Waters LS, Storz G (2009) Regulatory RNAs in bacteria. Cell 136(4):615–628. doi:10.1016/j.cell.2009.01.043

Wilderman PJ, Sowa NA, FitzGerald DJ, FitzGerald PC, Gottesman S, Ochsner UA, Vasil ML (2004) Identification of tandem duplicate regulatory small RNAs in *Pseudomonas aeruginosa* involved in iron homeostasis. Proc Natl Acad Sci USA 101(26):9792–9797. doi:10.1073/pnas.0403423101

Zhang A, Wassarman KM, Rosenow C, Tjaden BC, Storz G, Gottesman S (2003) Global analysis of small RNA and mRNA targets of Hfq. Mol Microbiol 50(4):1111–1124

Zielinski NA, Chakrabarty AM, Berry A (1991) Characterization and regulation of the *Pseudomonas aeruginosa* algC gene encoding phosphomannomutase. J Biol Chem 266 (15):9754–9763

Chapter 5
Streptomyces Bacteria: Specialized Metabolism, Inter-species Interations and Non-coding RNAs

Matthew J. Moody, Stephanie E. Jones, David A. Crisante and Marie A. Elliot

Abstract *Streptomyces* bacteria are abundant in soil environments, where they have an unusual multicellular life cycle that involves filamentous growth and spore formation. They also produce an extraordinary range of compounds known as specialized metabolites, and it is through these compounds that they interact with many of their terrestrial neighbours. Specialized metabolite production is subject to a wide range of regulatory inputs, and it is predicted that non-coding RNAs are amongst the many regulators governing metabolic output. RNA-sequencing experiments have revealed many non-coding RNAs expressed within specialized metabolic clusters of diverse *Streptomyces* species, with antisense RNAs featuring prominently. Here, we highlight a number of specialized metabolites whose gene clusters contain known non-coding RNAs, and consider possible roles for these RNA regulators in influencing *Streptomyces* interactions with other organisms in the environment.

5.1 Introduction

The soil is home to a diverse community of organisms that encompasses everything from the microscopic (e.g. bacteria and fungi) to the macroscopic (e.g. plants and insects). Among these organisms, the soil-dwelling actinobacteria, notably the streptomycetes, are the most ubiquitous of the soil bacteria. Streptomycetes are best known for their ability to produce a multitude of medically useful compounds, and for their complex developmental cycle. Unlike most bacteria, *Streptomyces* have a multicellular life cycle, which involves a series of defined, differentiated stages (Fig. 5.1). *Streptomyces* are sporulating bacteria, and following spore germination,

Matthew J. Moody—Stephanie E. Jones, these two authors contributed equally to this work.

M.J. Moody · S.E. Jones · D.A. Crisante · M.A. Elliot (✉)
Department of Biology and Michael G. DeGroote Institute for Infectious Disease Research, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4K1, Canada
e-mail: melliot@mcmaster.ca

A.L. Leitão and F.J. Enguita (eds.), *Non-coding RNAs and Inter-kingdom Communication*, DOI 10.1007/978-3-319-39496-1_5

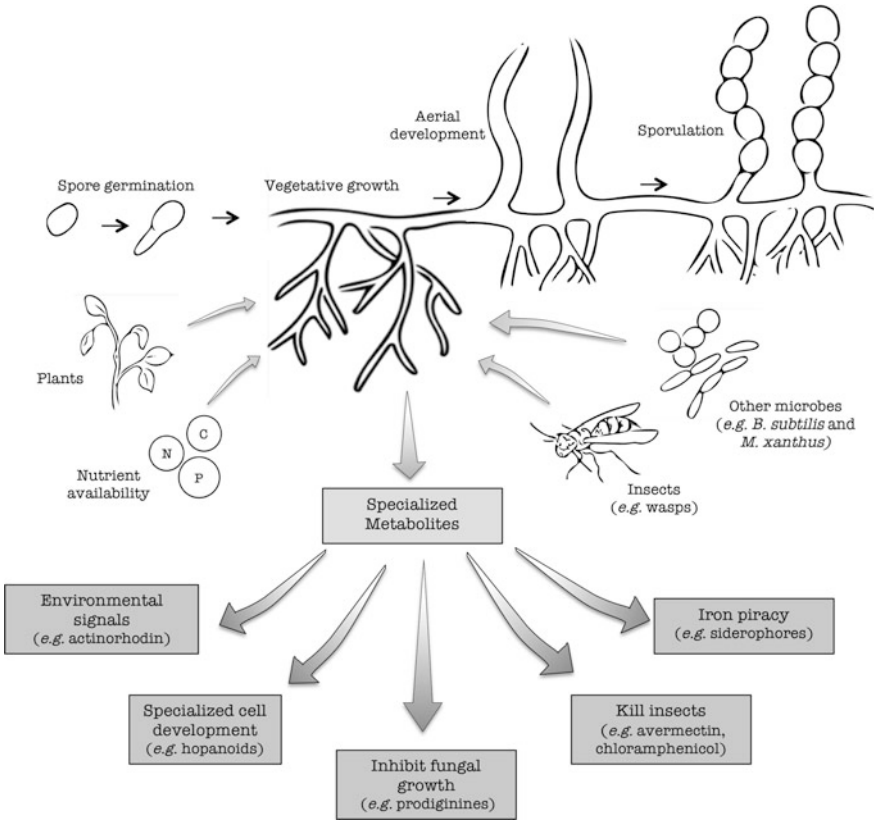


Fig. 5.1 ***Streptomyces*** **life cycle, and interactions that affect specialized metabolite production.** When *Streptomyces* spores encounter favourable growth conditions, they germinate. Germ tubes grow by hyphal tip extension and branching, generating a dense vegetative mycelial network. In response to environmental cues (such as those indicated below the life cycle), aerial hyphae extend upwards and eventually septate into spores. These spores can then be dispersed, and the cycle can begin anew. Specialized metabolite production typically initiates during the transition from vegetative growth to aerial development, and these can both mediate inter-species (and inter-kingdom) interactions, and confer growth benefits to *Streptomyces* by helping to obtain new nutrients or promote developmental progression

they grow via hyphal tip extension and branching to form a filamentous network of cells known as the vegetative mycelium. When nutrients become limiting, or in response to as yet unknown environmental cues, reproductive growth is initiated, with unbranched filaments first extending from the vegetative cells, rising into the air. These aerial filaments then undergo a synchronous round of septation to form chains of unigenomic spores (Flärdh and Buttner 2009; McCormick and Flärdh 2012).

During the *Streptomyces* life cycle, the transition from vegetative to aerial growth coincides with the onset of specialized (or secondary) metabolite production (Fig. 5.1). These metabolites differ from primary metabolites in that they are not

required for *Streptomyces* vegetative growth. Instead, their production is expected to confer a fitness benefit in their natural environment, and based on the coordination of their regulation with aerial hyphae formation, they are expected to be most important during the reproductive (sporulation) phase of growth. Virtually all *Streptomyces* species examined to date have the genetic ability to produce at least 20–30 distinct specialized metabolites (Traxler et al. 2013). These metabolites have been broadly co-opted for use in medicine and agriculture: many are employed as antibiotics, whilst others have utility as anti-fungals, anti-parasitic compounds and chemotherapeutic agents (Nett et al. 2009). Despite the importance of these compounds to clinical and veterinary medicine, we know little about their ecological roles.

Interactions between different organisms are inevitable in the diverse soil environment, and these can take the form of cooperative or competitive associations. For *Streptomyces* bacteria, specialized metabolites are predicted to mediate interactions with other organisms in the environment (Fig. 5.1). It has long been thought that these metabolites are agents of inter-species competition, whereby *Streptomyces* release chemical weapons to defend their territory in response to nutrient limitation. Recent findings—driven by the realization that antibiotics rarely reach inhibitory levels in the soil, yet can still induce differential transcription in nearby species—have prompted speculation that these specialized metabolites may serve as signaling molecules (Davies 2013). Whether *Streptomyces* specialized metabolites act as weapons or signals, they play a central role in mediating a myriad of interactions in the heterogeneous soil environment. Recent work has revealed that associations with other bacteria can alter the secreted metabolome of *Streptomyces* (Abrudan et al. 2015; Traxler et al. 2013), supporting the notion that these specialized metabolites mediate interactions with other bacteria, and likely other soil organisms including fungi, insects and plants. Although several remarkable examples of streptomycete-eukaryotic interactions have been described (Barke et al. 2010; Bignell et al. 2010; Currie et al. 1999; Kaltenpoth et al. 2005), much remains to be elucidated about the dynamics and consequences of these inter-kingdom interactions. Here, we consider how *Streptomyces* non-coding RNAs—key regulatory elements that play multifaceted roles in bacteria—could affect the activity/secretion of specialized metabolites, and how their effects may influence *Streptomyces* interactions with organisms ranging from other bacteria to higher order eukaryotes.

5.2 Regulation of Specialized Metabolite Production

The production of any given specialized metabolite is driven by gene products encoded from a discrete gene cluster in the *Streptomyces* chromosome (Fig. 5.2). These metabolic gene clusters direct the expression of the enzymes needed to synthesize the core metabolite, along with those that modify and tailor the

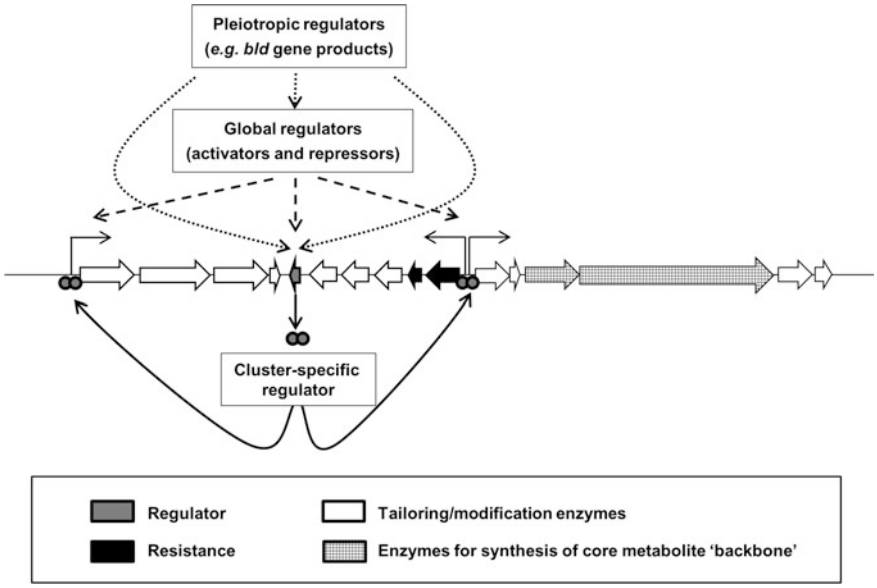
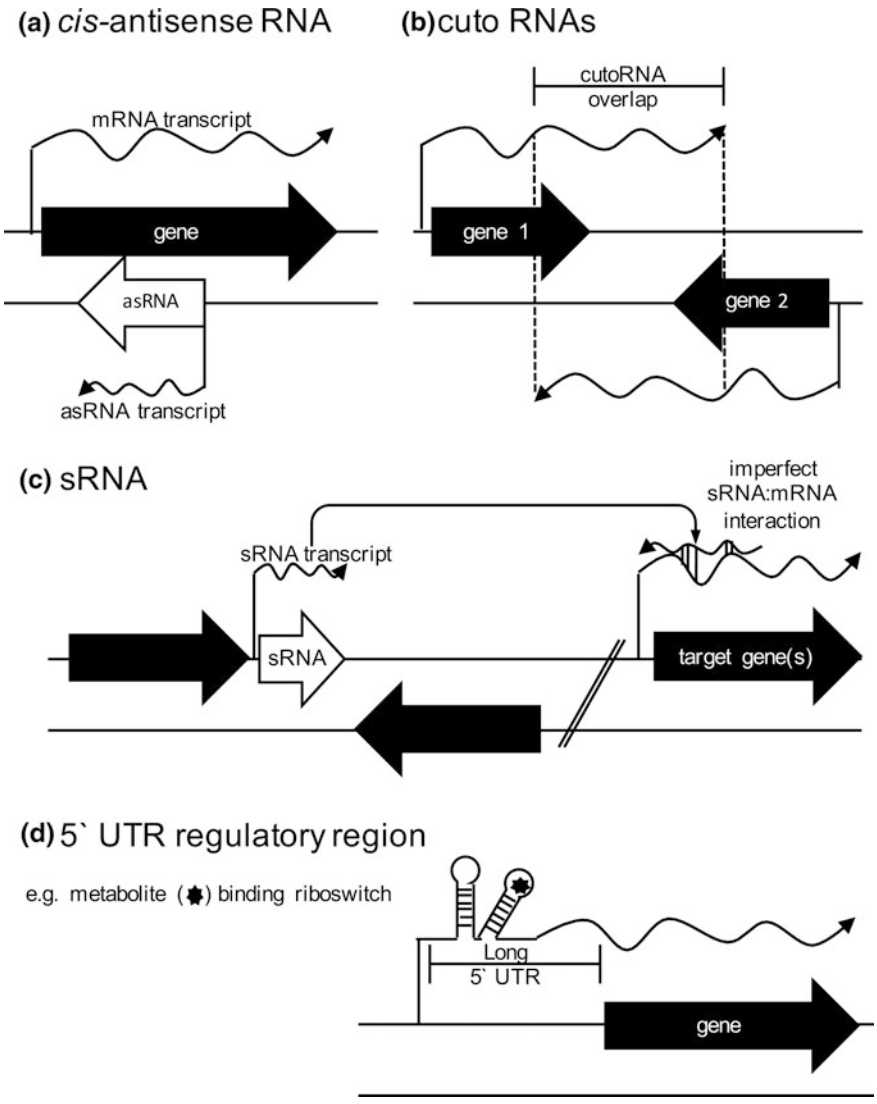


Fig. 5.2 Hierarchy of protein-based regulation of gene expression within a specialized metabolite biosynthetic cluster. Specialized metabolite biosynthetic clusters typically contain all the genes whose products are needed to synthesize the metabolite ('core' and 'tailoring' enzymes), as well as those encoding regulators and resistance/efflux determinants. These clusters are subject to extensive regulation. The most basic level of control is exerted by the cluster-specific regulators, which are encoded within the biosynthetic cluster whose expression they control. Global regulators are expressed from locations outside these biosynthetic clusters, and often affect the expression of more than one cluster; they can directly influence the expression of the biosynthetic genes, or, more commonly, the cluster-specific regulator. Pleiotropic regulators control both development and specialized metabolite production, and the best characterized of these are the *bld* gene products (Flärdh and Buttner 2009)

molecule. In addition, these clusters also direct the expression of regulatory proteins and resistance/secretion determinants (e.g. efflux pumps). Perhaps not surprisingly, specialized metabolite production is a highly regulated process, integrating diverse regulatory inputs including a wide range of nutritional cues (e.g. carbon, nitrogen and phosphorus availability) and the proximity of other organisms (Fig. 5.1). Specialized metabolic gene clusters are subject to control by different mechanisms, including pleiotropic regulators that couple metabolite production with aerial development, global metabolic regulators that influence the production of multiple metabolites, and cluster-specific regulators that activate/repress the expression of individual metabolites (Fig. 5.2) (reviewed in Bibb 2005). More recently, there have been reports suggesting that non-coding RNAs (ncRNAs) may also play key roles in governing the production of specialized metabolites (Moody et al. 2013).

5.3 Regulation by Non-coding RNAs

The regulatory roles ascribed to ncRNAs have grown tremendously over the last decade, and many classes of ncRNAs have now been identified in bacteria. These include antisense RNAs (asRNAs), *trans*-encoded small RNAs (sRNAs) and riboswitches (Fig. 5.3). All but the sRNAs have been associated with specialized



◀ **Fig. 5.3 Schematic diagram of the different classes of bacterial non-coding RNAs**. Protein-coding genes are depicted as *black arrows* and non-coding RNAs are shown as *white arrows*. RNA transcripts are shown above their corresponding gene, with transcription initiating at the *vertical line*, and terminating at the *small arrowhead*. **a** *cis*-antisense RNAs (referred to here as simply antisense RNAs or asRNAs) are transcribed on the strand, opposite a protein-coding gene. **b** cutoRNAs are a subset of asRNA, and arise when the 3′ UTR of an mRNA overlaps with that of a convergently-oriented downstream gene. The region of overlap is indicated with a bracketed line. **c** Intergenic *trans*-encoded small RNA (sRNA) genes are found in the intergenic regions between genes, and typically target (by imperfect complementary base-pairing) one or more mRNAs expressed from disparate chromosomal locations. **d** Long 5′ UTRs can be subject to a wide range of regulatory effects, with riboswitches being amongst the best-studied. Riboswitches adopt structures that permit the specific recognition and binding of small metabolites, leading to a change in the transcriptional read-through or translation of the associated mRNA. Other long 5′ UTRs can be targeted by sRNAs or RNA binding proteins, which alter the stability or translatability of the associated mRNA

metabolic clusters, and the current lack of sRNA-mediated regulation is almost certain to change as our understanding of these regulators grows.

In the case of asRNAs, these are most commonly expressed from the strand opposite a protein-coding gene, and consequently, they share perfect complementarity with their corresponding mRNA target (Georg and Hess 2011) (Fig. 5.3). In binding to their target mRNAs, they can affect mRNA stability (positively or negatively), and/or influence translation by modulating ribosome accessibility. Similar regulatory outcomes (transcript stability and mRNA translatability) have been ascribed to the sRNAs, which are so named for their relatively small sizes (~50–300 nt). asRNAs can also arise from the transcriptional read-through of convergently oriented genes that lack transcriptional terminators, such that there is considerable overlap of their 3′ untranslated region (and in some instances, overlap extending into the coding sequence) (Fig. 5.3, and below). We have termed this subset of asRNAs 'cutoRNAs' (for convergent untranslated overlapping RNAs) (Moody et al. 2013). While it is possible to envision a multitude of regulatory outcomes stemming from such overlap (e.g. coordinated degradation; coordinated translation), there is not yet any experimental evidence supporting such a role for these ncRNAs.

Unlike asRNAs, sRNAs are expressed from intergenic regions, and act in *trans* to target one or more mRNAs through imperfect base-pairing (Fig. 5.3). Often the secondary structures of sRNAs and mRNAs play an important role in facilitating the interaction. In many bacterial species, productive sRNA-mRNA interactions require the activity of the RNA chaperone Hfq; however, there is no such chaperone found in the streptomycetes (Jousselin et al. 2009). A handful of sRNAs have been shown to target proteins in place of mRNAs, where they act as molecular decoys, sequestering proteins from their native targets, thus inhibiting their activity (Waters and Storz 2009).

asRNAs and sRNAs are typically expressed as discrete genetic entities. In contrast, riboswitches are structured RNA elements usually found at the 5′ end of an mRNA. They adopt different conformations in the presence/absence of a ligand,

and serve to modify translation, transcription or stability of the downstream *cis*-encoded mRNA (Fig. 5.3) (reviewed in Montange and Batey 2008).

To date, hundreds of ncRNAs have been identified in model *Streptomyces* species, and they appear to be abundant in specialized metabolic gene clusters (Moody et al. 2013; Swiercz et al. 2008; Vockenhuber et al. 2011). Given the potential for specialized metabolites to serve as communication signals and/or chemical weapons, and given the role of ncRNAs in rapidly altering genetic programs in response to environmental changes, it is easy to imagine situations whereby ncRNAs would modulate inter-species and inter-kingdom interactions. Below, we highlight a number of specialized metabolites whose gene clusters contain known ncRNAs, and consider possible roles for these RNA regulators in affecting *Streptomyces* associations with other organisms in the environment.

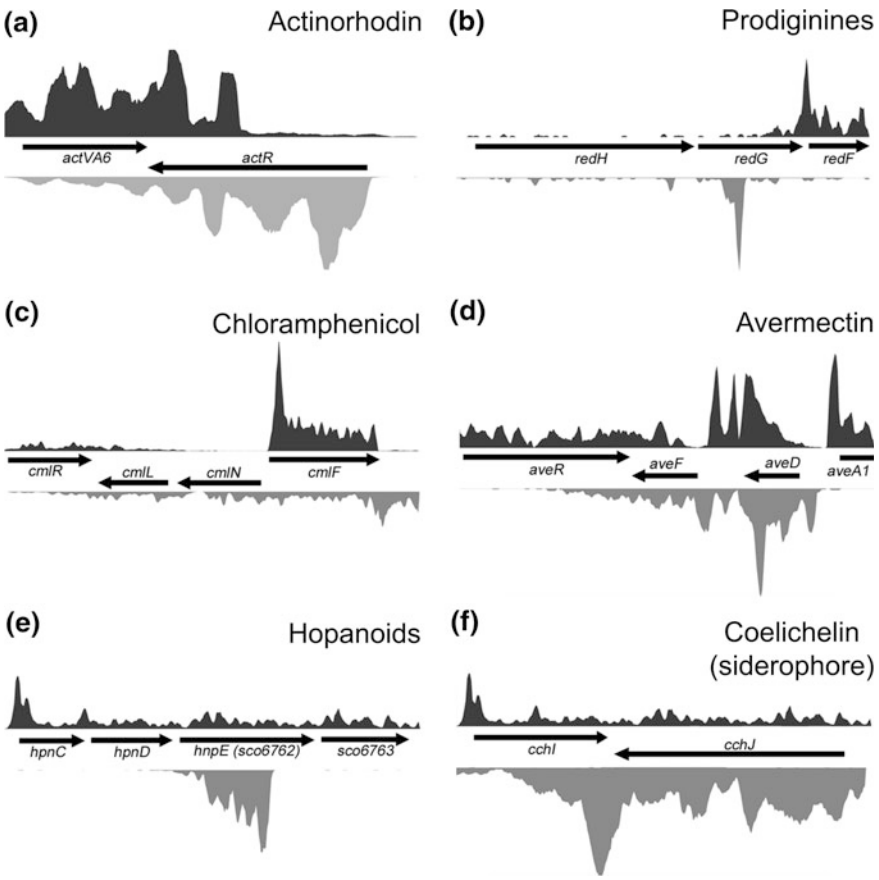
5.4 Actinorhodin—A Redox-Active Pigmented Antibiotic Produced by *Streptomyces coelicolor*

The best studied streptomycete, *Streptomyces coelicolor*, is so named for its characteristic blue colouration (coelicolor = sky blue). This is due to the production of a blue-pigmented specialized metabolite known as actinorhodin. In the environment, actinorhodin is secreted from the cell, where it can either diffuse freely, or be packaged into extracellular vesicles (Schrempf et al. 2011). It has antibiotic activity, specifically inhibiting the growth of nearby Gram-positive bacteria, and can also serve as a signalling molecule, where it is recognized and bound by receptors in other bacteria, altering the transcription profiles of these organisms (Xu et al. 2010). Actinorhodin production is further stimulated by the nearby growth of other microbes, including *Myxococcus xanthus*, *Bacillus subtilis*, and *Serratia* species. *M. xanthus* is a predatory soil microbe, and would inhabit the same environmental niche as *Streptomyces,* but as a Gram-negative bacterium, it would be impervious to the antibiotic activity of actinorhodin. Intriguingly, actinorhodin curtails *M. xanthus* predation of *S. coelicolor*, suggesting that in addition to its antibiotic and signalling properties, actinorhodin may also modulate the behaviour of other microbes through its action as a repellent (Perez et al. 2011).

The regulatory networks governing actinorhodin biosynthesis are the best understood of any specialized metabolic cluster studied to date. Expression of its cluster-specific regulator (ActII-ORF4) is directly controlled by multiple pleiotropic and global regulators (Fernandez-Moreno et al. 1991; Gao et al. 2012; McKenzie and Nodwell 2007; Rigali et al. 2008; Uguru et al. 2005; Wang et al. 2013) (e.g. Fig. 5.2). A second regulator within the actinorhodin biosynthetic cluster, ActR, directs the expression of several transporter genes, whose products are thought to confer self-resistance to actinorhodin by exporting it from the cytoplasm. Remarkably, ActR coordinates both actinorhodin synthesis and efflux: as a TetR-like regulator, ActR represses the expression of its divergently expressed

target genes (Tahlan et al. 2007) until it binds either actinorhodin or an actinorhodin precursor. Ligand binding results in a conformational change in ActR that inhibits its DNA binding ability, ultimately leading to efflux gene expression (Xu et al. 2012).

While there are currently no *trans*-encoded sRNAs known to affect actinorhodin production, nor any independently-encoded asRNAs identified within the biosynthetic cluster, there is a striking cutoRNA shared by the *actR* gene and the convergently expressed *actVA-5/6* operon (Fig. 5.4a). The *actR* and *actVA-6* genes share 34 nucleotides of sequence overlap at their 3′ ends, and transcription from *actR* and the *actVA-6*-containing operon extends the full length of their respective downstream genes (Fig. 5.4a). A similar organization is found in other specialized metabolic clusters [e.g. the simocyclinone (DNA gyrase-targeting antibiotic) biosynthetic cluster in *Streptomyces antibioticus* (Le et al. 2009)]. It is tantalizing to speculate that this genetic overlap, and corresponding transcriptional overlap, provides a means of physically coupling the final steps in antibiotic production



◀ **Fig. 5.4 RNA-seq expression profiles of select non-coding RNAs associated with specialized metabolite biosynthetic clusters**. Genes are shown schematically as *black arrows*. Above these, the *dark grey* coloured graphs represent relative read coverage for the forward strand, while the *bottom*, *light grey* graphs represent read coverage for the reverse strand. **a** Within the actinorhodin biosynthetic gene cluster in *S. coelicolor*, the lack of transcriptional termination for both the *actR* gene (encoding an efflux pump regulator) and the convergently expressed *actVA-6* operon (encoding tailoring enzymes), results in the formation of a cutoRNA. **b** An asRNA is expressed opposite *redG* (encoding a Rieske oxygenase-like enzyme) in the prodiginine biosynthetic gene cluster of *S. coelicolor*. **c** The chloramphenicol biosynthetic cluster of *S. venezuelae* contains many non-coding RNAs. There is an asRNA expressed opposite *cmlF* (efflux pump), and cutoRNAs are expressed from the convergently arranged *cmlR* (cluster-specific activator) and the *cmlLN* operon (likely biosynthetic enzyme and efflux pump, respectively). **d** Many non-coding RNAs are located within the avermectin biosynthetic gene cluster of *S. avermitilis*. cutoRNAs arise from the convergently transcribed *aveR* (cluster-specific regulator) and *aveF* (tailoring enzyme). Numerous asRNAs, are expressed opposite the *aveDF* operon, where both these genes specify tailoring enzymes. And the highly expressed 5′ UTR of *aveA1*, which encodes a polyketide synthase (core biosynthetic enzyme), may have a regulatory function, given the differential expression of the UTR relative to the coding sequence. **e** Within the hopanoid biosynthetic gene cluster of *S. coelicolor* an asRNA is expressed opposite *sco6762* (product of unknown function). **f** A cutoRNA is created by the convergent transcription of two genes (*cchJ*/*cchI*) encoding a transporter/biosynthetic enzyme, respectively, in the coelichelin siderophore biosynthetic gene cluster

with antibiotic export. With actinorhodin, this would directly connect the ActVA-6-mediated production of an actinorhodin precursor and dimerization of this intermediate molecule by ActVA-5 (Sciara et al. 2003; Valton et al. 2008), with the regulator that senses both precursor molecules and the final actinorhodin product. The complementarity shared by these mRNAs may ensure that the resulting translated products are coordinately synthesized in space and time, or may alter the stability and/or translation of either transcript. The cutoRNA connection with actinorhodin export would link it directly to the secretion of this antibiotic into the environment, where it would be available to exert its effects on any nearby organisms.

5.5 Prodiginines—Fungal Activity Modulators

In addition to actinorhodin, *S. coelicolor* also produces a family of red-pigmented specialized metabolites known collectively as the prodiginines. These compounds have antibacterial, antifungal, antimalarial, immunosuppressive and anticancer activities. *S. coelicolor* produces a mixture of prodiginines, including undecyl-prodigiosin (Red) and its cyclic derivative butyl-meta-cycloheptylprodiginine (streptorubin B) (Williamson et al. 2006). The wide range of bioactive properties ascribed to the prodiginines has prompted considerable speculation about the ecological function of these molecules. Similar to actinorhodin, prodiginine

production by *S. coelicolor* is stimulated upon interactions with other bacteria, including association with live *Bacillus subtilis* (Traxler et al. 2013; Yang et al. 2009), heat-killed *B. subtilis* or *Staphylococcus aureus* (Luti and Mavituna 2011), and live mycolic acid-containing bacteria (Onaka et al. 2011). A recent study revealed antibiotics produced by other streptomycetes (specifically, jadomycin B produced by *Streptomyces venezuelae*) can induce *S. coelicolor* prodiginine production (Wang et al. 2014). In each of these cases, it is not known how or why prodiginine production is induced. It has been proposed that prodiginines are produced in response to competition for nutrients with other bacteria; it is conceivable that fungi sharing streptomycete territory could similarly induce prodiginine synthesis. The induction of prodiginine production by jadomycin B further suggests other antibiotics—including those produced by fungi—could stimulate prodiginine release (Traxler and Kolter 2015).

Meschke et al. (2012) recently demonstrated the first direct role for prodiginines in an inter-kingdom interaction. The authors showed prodiginines produced by *Streptomyces lividans* could suppress the growth of the fungus *Verticillium dahlia*, the causative agent of vascular wilt in over 200 plant species including *Arabidopsis thaliana*. Co-cultivation of *S. coelicolor* and *V. dahlia* led to enhanced prodiginine expression, and a corresponding reduction in *V. dahlia* infection of *A. thaliana* (Meschke et al. 2012). While the benefit of prodiginine production for *S. lividans* in this case is unclear, the co-existence of microbes and plants in the rhizosphere suggests that a complex network of prodiginine-fungal-plant interactions could influence the growth and behaviour of multiple species.

Prodiginine production in *S. coelicolor* is directed by the *red* biosynthetic cluster, and yields a 2:1 mixture of undecylprodigiosin:streptorubin B (Williamson et al. 2006). The *red* cluster consists of 22 genes, organized into four transcriptional units, with the *redG* product, a Rieske oxygenase-like enzyme, catalyzing the conversion of undecylprodigiosin to streptorubin B. An asRNA is encoded on the strand opposite *redG* (Fig. 5.4b), and this ncRNA could conceivably provide a means of modulating the undecylprodigiosin:streptorubin B ratio, without affecting the expression of downstream genes whose products are required at earlier stages in the undecylprodigiosin biosynthetic pathway (Moody et al. 2013). *S. lividans* has a near identical cluster to that of *S. coelicolor,* raising the possibility that an equivalent asRNA may control undecylprodigiosin:streptorubin B levels in this species as well. In *S. lividans,* as in *S. coelicolor,* undecylprodigiosin levels are typically higher than those of streptorubin B, and levels of both molecules were increased by $\sim$4× during co-cultivation with *V. dahlia* (Meschke et al. 2012). It would be interesting to determine whether interactions with different fungi (or bacteria) could alter the levels of the *redG* asRNA, thus skewing the undecylprodigiosin:streptorubin B ratio.

5.6 Chloramphenicol

As mentioned briefly above, *S. venezuelae* can stimulate prodiginine production through its secretion of the antibiotic jadomycin B. *S. venezuelae* is, however, better known for its production of chloramphenicol, which was the first antibiotic to be synthesized on a large scale for clinical use. Although the therapeutic use of chloramphenicol has declined due to its adverse side-effects, its broad spectrum activity (against both Gram-positive and Gram-negative bacteria) means it is still prescribed in some circumstances (Nitzan et al. 2010). A recent study revealed that *S. venezuelae* produces appreciable amounts of chloramphenicol in its natural soil environment (Berendsen et al. 2013). Despite its high natural occurrence in soils, the role played by chloramphenicol in inter-bacterial and inter-kingdom interactions has not been directly explored.

Chloramphenicol inhibits protein synthesis by binding the 50S ribosomal subunit in bacteria and mitochondria. Blocks in mitochondrial protein synthesis, such as those associated with antibiotic inhibition of the 50S subunit, can result in decreased ATP biosynthesis and generalized mitochondrial stress, and this in turn can be associated with accelerated human cancer progression (Li et al. 2010). Given the relative abundance of chloramphenicol in the soil, it seems likely that *S. venezuelae*, and other producing bacteria, could use chloramphenicol as a chemical weapon to induce mitochondrial stress in eukaryotes. This could provide a fitness benefit to *S. venezuelae* during competition with fungi for nutrients, and indeed, supporting this proposal is the observation that chloramphenicol can inhibit the growth of the soil fungus *Spocullariopsis* (now *Microascus*) *brevicaulis* (Broadbent and Terry 1958).

The chloramphenicol biosynthetic cluster in *S. venezuelae* comprises 15 protein-coding genes and includes several prominent asRNAs (Fig. 5.4c). One of these is expressed opposite the gene encoding the chloramphenicol efflux pump CmlF. It is thought that CmlF directs the secretion of chloramphenicol into the environment, and modulating levels of its asRNA might provide a way of altering chloramphenicol release in response to the presence of bacteria or fungi (Moody et al. 2013). The other asRNAs are cutoRNAs expressed from a region containing three protein-coding genes that were recently shown to be part of the chloramphenicol cluster: *cmlR*, a transcriptional activator of the cluster; *cmlL*, a phosphopantetheinyl transferase; and *cmlN*, a Na^+/H^+ antiporter (Fernandez-Martinez et al. 2014). The transcriptional organization in this region is strikingly complex, and is analogous to that of *actR* and the *actVA-6* operon in *S. coelicolor*. *cmlL* and *cmlN* are transcribed as an operon, and this operon is convergently oriented relative to *cmlR*. The 3′ UTR of the *cmlLN* transcript extends throughout the entire *cmlR* encoded region, and likewise, the 3′ UTR of the *cmlR* transcript extends throughout the majority of the *cmlLN* operon. Given that CmlR activates the cluster and is required for chloramphenicol production, while CmlN is thought to be involved in chloramphenicol export, both asRNAs might be able to rapidly modulate chloramphenicol production and export in response to changes in environmental conditions (Moody et al. 2013).

5.7 The Avermectin Cluster in *Streptomyces avermitilis*

The anti-eukaryotic properties of chloramphenicol are not unique to this metabolite. *Streptomyces avermitilis* produces a group of macrocyclic lactones known as avermectins, which are highly toxic to insects and nematodes. These toxic properties have led to the pervasive use of avermectins as anti-parasitic agents in human and veterinary medicine (Omura and Crump 2004). Avermectins act to paralyze—and ultimately kill—insects, by disrupting glutamate-gated chloride channels that are critical for neurotransmission (Arena et al. 1991; Cully et al. 1996; Kane et al. 2000). Why *S. avermitilis* produces avermectins in its natural environment is not known, but *Streptomyces*-insect associations have been well documented. These interactions are, however, more typically in the form of symbiotic relationships with e.g. wasps and termites, where the insect provides *Streptomyces* with nutrients, and *Streptomyces* in turn, supplies an arsenal of specialized metabolic weapons to ward off microbial pathogens (Seipke et al. 2012). It is conceivable that similar relationships have evolved in which *S. avermitilis* produces avermectins to protect a host from enemy insects.

In *S. avermitilis,* the avermectin biosynthetic gene cluster spans ~81 kb and includes 19 protein-coding genes. While ncRNAs abound within this cluster, there is an 18 kb region in which they are particularly enriched (Fig. 5.4d). This region contains genes that contribute to the regulation, synthesis and tailoring of the avermectins (Ikeda et al. 2003; Omura et al. 2001), and includes genes coding for the cluster-specific activator AveR, two polyketide tailoring enzymes (AveF and AveD), and the first of four polyketide synthases (AveA1) [polyketides are a class of specialized metabolites that are produced by enzyme complexes having similarity to fatty acid synthases]. In addition to these protein-coding genes, this region encompasses multiple asRNAs, a significant cutoRNA, and a long 5′ UTR that may have regulatory potential (Fig. 5.4d).

aveR and *aveF* are convergently transcribed but lack intergenic transcriptional terminators. This organization provides an opportunity to generate a long cutoRNA, as a result of read-through transcription for each gene. In the absence of temporal data, it is not yet clear whether these genes are expressed at the same time, and if they are, what benefit may arise from their transcriptional overlap. Upstream of *aveF* and its co-transcribed *aveD* gene, there exist several highly expressed asRNAs, which may function to alter the relative amounts of AveD and AveF accumulating in the cell. Finally, the 5′ UTR of *aveA1* is unusually long (130 nt), suggesting that there may be regulatory elements (e.g. riboswitch or sRNA/protein regulator binding sites) contained within this region (Moody et al. 2013). It has yet to be determined how these ncRNAs impact avermectin production; however, the transcriptional complexity in this small region suggest that ncRNA-mediated regulation likely has profound effects on avermectin production.

5.8 Antisense RNA Control of Hopanoid Production—A Role in Altering Membrane Properties

While many specialized metabolites are produced in a largely species-specific manner, including actinorhodin and avermectin, there are others that are produced by many bacterial species, like the hopanoids and siderophores (see below). Hopanoids are pentacyclic compounds produced by bacteria, and are considered to be functionally equivalent to eukaryotic sterols (Saenz et al. 2012). Like sterols, hopanoids alter membrane function, specifically affecting membrane fluidity and permeability (Wu et al. 2015). In other environmental bacteria (specifically *Burkholderia*), hopanoid production enhances resistance to antibiotics and detergents, and promotes survival in the face of acidic pH, temperature fluctuations and changes to osmotic pressure (Malott et al. 2012; Schmerk et al. 2011). In *Streptomyces*, hopanoids do not seem to promote resistance to exogenous stresses (Seipke et al. 2012), although the hopanoid gene cluster in *S. coelicolor* is under the control of the thiol/oxidative stress-responsive sigma factor SigR, suggesting a role in mitigating oxidative or redox stress (Kim et al. 2012). The primary function of hopanoids in *Streptomyces,* however, seems to be during sporulation, where they are proposed to minimize water loss, and thus promote spore viability (Poralla et al. 2000). This is consistent with the observation that these molecules are produced predominantly during sporulation on solid media (Moody et al. 2013; Poralla et al. 2000). This role in helping promote the formation of specialized cell types (like spores) seems to be conserved throughout the actinobacteria. In the streptomycetes, spore formation permits survival during adverse growth conditions, and this dormant state offers immunity to many of the chemical insults levied against them by other bacteria and fungi.

The hopanoid biosynthetic cluster in *S. coelicolor* is predicted to encompass 12 genes (Bentley et al. 2002; Liu et al. 2014; Pan et al. 2015; Seipke et al. 2012) (Fig. 5.4e). In examining the transcriptional profile of the hopanoid cluster in several streptomycetes, a strong asRNA was observed to be expressed opposite *sco6762* in *S. coelicolor* and its orthologues in other species. In *Frankia*, an actinobacterium and nitrogen-fixing plant symbiont, hopanoids contribute to the development of nitrogen-fixing nodules within plant roots, facilitating the formation of an oxygen-impermeable barrier around the nitrogenase enzyme (Berry et al. 1993). This gene (*sco6762*) is conserved in *Frankia*, and perhaps with it, the asRNA. The product of SCO6762 is most similar to that of HpnE, an enzyme predicted to convert hydroxysqualene to squalene (Pan et al. 2015), where squalene is an intermediate in hopanoid biosynthesis. Based on what is known about asRNA activity in other systems, we predict this asRNA may function to negatively regulate the levels of SCO6762, and possibly decouple the expression of this gene (and those downstream in the operon) from the hydroxysqualene biosynthetic genes upstream. This may allow the cells to direct the metabolites produced by the upstream HpnC/HpnD gene products into other pathways, thereby downregulating hopanoid production. It will be interesting to determine whether expression of the

asRNA is limited to cells in which hopanoid production is less important, such as in *Frankia* cells not associated with the nodule/nitrogenase production, and in *Streptomyces* vegetative (non-sporulating) cells.

5.9 Siderophores—Production and Piracy in the Soil

Iron is essential for the growth of virtually all organisms, but in the environment, soluble iron is a scarce commodity. Consequently, there is fierce competition between bacteria and fungi for the available iron. Microbes acquire iron through the release of small molecule siderophores. These specialized metabolites are used to scavenge iron: they bind Fe(III), and are recognized by receptors on the bacterial/fungal surface. The Fe-siderophore complex may then be actively transported into the cell (most common in bacteria), or the iron molecule may be selectively taken up (more common in fungi) (Saha et al. 2013). Many soil bacteria can recognize and take-up iron-bound siderophores produced by other species in a process known as siderophore-piracy.

When plated next the bacterium *Amycolatopsis sp*. AA4*, S. coelicolor* fails to develop aerial hyphae. This change in development is due to siderophore-piracy: *Amycolatopsis sp.* AA4 preferentially takes up the siderophore desferrioxamine produced by *S. coelicolor,* and consequently *S. coelicolor* fails to complete its aerial developmental program, as this is a highly iron-dependent process (Traxler et al. 2012). Notably, *Amycolatopsis* downregulates the expression of its native siderophore biosynthetic genes when growing next to *S. coelicolor*, which likely provides a strong competitive advantage, given that siderophore biosynthesis is an energetically expensive process (Traxler et al. 2012). In addition to desferrioxamine, *S. coelicolor* produces two other siderophores—coelichelin and coelibactin (Barona-Gomez et al. 2006). It is likely that different species have evolved the ability to use a variety of siderophores, and thus have the capacity to switch their iron-uptake strategies to minimize the effects of siderophore piracy by their neighbours (Barona-Gomez et al. 2006). More recent work has expanded the idea of siderophore-piracy to inter-kingdom interactions. Growth of a mutant *S. coelicolor* strain lacking the ability to produce desferrioxamine can be restored when grown near two different fungi—a *Penicillium* species and *Engyodontium album*, perhaps through its own act of stealing fungal siderophores (Arias et al. 2015).

How cells determine which iron-acquisition systems to activate is not fully understood. However, in multiple *Streptomyces* species, there are asRNAs expressed opposite both siderophore biosynthetic- and uptake-associated genes (Moody et al. 2013). For example, within the coelichelin biosynthetic cluster of *S. coelicolor,* a cutoRNA arises through the convergent transcription of *cchJ*, a putative coelichelin transporter, and *cchI* a coelichelin biosynthetic enzyme (Fig. 5.4f) (Lautru et al. 2005), similar to the situation in the actinorhodin cluster, where synthesis and transport are connected (see Fig. 5.4a). Further work is required to determine if these asRNAs are modulated in response to signals from other species.

5.10 Conclusions and Outlook

Although we are only beginning to understand how ncRNAs affect *Streptomyces* specialized metabolite production, it is clear that these compounds can mediate a wide range of interactions with other bacteria, fungi, insects and plants. Given the heterogeneous nature of the soil, it is not surprising that specialized metabolites appear to be central to the evolution of relationships between streptomycetes and other organisms. We have described a number of examples of *Streptomyces* inter-kingdom interactions mediated by specialized metabolites; in each case, asRNAs are associated with the corresponding biosynthetic pathways. These asRNAs have the potential to contribute to the regulation/production of the associated compound, possibly providing a more effective means of modulating gene expression than proteins, or perhaps simply providing a means of fine-tuning gene expression. Several described asRNAs are associated with specialized metabolite exporters, and in these cases, it is conceivable that asRNAs could modulate the levels of specialized metabolite secretion, and in doing so, could dictate whether the compound is secreted at low levels, when it might function as a signalling molecule, or at higher levels, where it may act as a weapon.

Since the majority of *Streptomyces* biosynthetic clusters are not expressed under laboratory conditions, it is impossible to predict how broadly asRNAs, and ncRNAs in general, will feature in the regulation of specialized metabolite production and inter-kingdom interactions. We anticipate that further studies of *Streptomyces* under more natural conditions will reveal widespread roles for specialized metabolites in the cooperation and competition with other organisms in its environment, and will reveal novel roles for asRNAs and other types of ncRNAs in the activation and regulation of specialized metabolite production.

Acknowledgments This work was supported by an NSERC Discovery Grant (No. 04681), and an NSERC Discovery Accelerator Supplement. M.J.M. was supported by an NSERC Doctoral Scholarship, and S.E.J was supported by a Vanier Canada Graduate Scholarship.

References

Abrudan MI, Smakman F, Grimbergen AJ, Westhoff S, Miller EL, van Wezel GP, Rozen DE (2015) Socially mediated induction and suppression of antibiosis during bacterial coexistence. Proc Natl Acad Sci USA 112(35):11054–11059. doi:10.1073/pnas.1504076112

Arena JP, Liu KK, Paress PS, Cully DF (1991) Avermectin-sensitive chloride currents induced by *Caenorhabditis elegans* RNA in *Xenopus* oocytes. Mol Pharmacol 40(3):368–374

Arias AA, Lambert S, Martinet L, Adam D, Tenconi E, Hayette MP, Ongena M, Rigali S (2015) Growth of desferrioxamine-deficient *Streptomyces* mutants through xenosiderophore piracy of airborne fungal contaminations. FEMS Microbiol Ecol 91(7). doi:10.1093/femsec/fiv080

Barke J, Seipke RF, Gruschow S, Heavens D, Drou N, Bibb MJ, Goss RJ, Yu DW, Hutchings MI (2010) A mixed community of actinomycetes produce multiple antibiotics for the fungus farming ant *Acromyrmex octospinosus*. BMC Biol 8:109. doi:10.1186/1741-7007-8-109

Barona-Gomez F, Lautru S, Francou FX, Leblond P, Pernodet JL, Challis GL (2006) Multiple biosynthetic and uptake systems mediate siderophore-dependent iron acquisition in *Streptomyces coelicolor* A3(2) and *Streptomyces ambofaciens* ATCC 23877. Microbiology 152(Pt 11):3355–3366. doi:10.1099/mic.0.29161-0

Bentley SD, Chater KF, Cerdeno-Tarraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D, Bateman A, Brown S, Chandra G, Chen CW, Collins M, Cronin A, Fraser A, Goble A, Hidalgo J, Hornsby T, Howarth S, Huang CH, Kieser T, Larke L, Murphy L, Oliver K, O'Neil S, Rabbinowitsch E, Rajandream MA, Rutherford K, Rutter S, Seeger K, Saunders D, Sharp S, Squares R, Squares S, Taylor K, Warren T, Wietzorrek A, Woodward J, Barrell BG, Parkhill J, Hopwood DA (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). Nature 417(6885):141–147. doi:10.1038/417141a

Berendsen B, Pikkemaat M, Romkens P, Wegh R, van Sisseren M, Stolker L, Nielen M (2013) Occurrence of chloramphenicol in crops through natural production by bacteria in soil. J Agric Food Chem 61(17):4004–4010. doi:10.1021/jf400570c

Berry AM, Harriott OT, Moreau RA, Osman SF, Benson DR, Jones AD (1993) Hopanoid lipids compose the *Frankia* vesicle envelope, presumptive barrier of oxygen diffusion to nitrogenase. Proc Natl Acad Sci USA 90(13):6091–6094

Bibb MJ (2005) Regulation of secondary metabolism in streptomycetes. Curr Opin Microbiol 8 (2):208–215. doi:10.1016/j.mib.2005.02.016

Bignell DR, Huguet-Tapia JC, Joshi MV, Pettis GS, Loria R (2010) What does it take to be a plant pathogen: genomic insights from *Streptomyces* species. Antonie Van Leeuwenhoek 98(2):179–194. doi:10.1007/s10482-010-9429-1

Broadbent D, Terry DA (1958) Effect of chloramphenicol on a fungus. Nature 182(4642):1107–1108

Cully DF, Paress PS, Liu KK, Schaeffer JM, Arena JP (1996) Identification of a *Drosophila melanogaster* glutamate-gated chloride channel sensitive to the antiparasitic agent avermectin. J Biol Chem 271(33):20187–20191

Currie CR, Scot JA, Summerbell RC, Malloch D (1999) Fungus-growing ants use antibiotic-producing bacteria to control garden parasites. Nature 398:701–704. doi:10.1038/19519

Davies J (2013) Specialized microbial metabolites: functions and origins. J Antibiot (Tokyo) 66 (7):361–364. doi:10.1038/ja.2013.61

Fernandez-Martinez LT, Borsetto C, Gomez-Escribano JP, Bibb MJ, Al-Bassam MM, Chandra G (2014) New insights into chloramphenicol biosynthesis in *Streptomyces venezuelae* ATCC 10712. Antimicrob Agents Chemother 58(12):7441–7450. doi:10.1128/AAC.04272-14

Fernandez-Moreno MA, Caballero JL, Hopwood DA, Malpartida F (1991) The *act* cluster contains regulatory and antibiotic export genes, direct targets for translational control by the *bldA* tRNA gene of *Streptomyces*. Cell 66(4):769–780

Flärdh K, Buttner MJ (2009) *Streptomyces* morphogenetics: dissecting differentiation in a filamentous bacterium. Nat Rev Microbiol 7(1):36–49. doi:10.1038/nrmicro1968

Gao C, Hindra, Mulder D, Yin C, Elliot MA (2012) Crp is a global regulator of antibiotic production in *Streptomyces*. MBio 3(6). doi:10.1128/mBio.00407-12

Georg J, Hess WR (2011) cis-antisense RNA, another level of gene regulation in bacteria. Microbiol Mol Biol Rev 75(2):286–300. doi:10.1128/MMBR.00032-10

Ikeda H, Ishikawa J, Hanamoto A, Shinose M, Kikuchi H, Shiba T, Sakaki Y, Hattori M, Omura S (2003) Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. Nat Biotechnol 21(5):526–531. doi:10.1038/nbt820

Jousselin A, Metzinger L, Felden B (2009) On the facultative requirement of the bacterial RNA chaperone. Hfq. Trends Microbiol 17(9):399–405. doi:10.1016/j.tim.2009.06.003

Kaltenpoth M, Gottler W, Herzner G, Strohm E (2005) Symbiotic bacteria protect wasp larvae from fungal infestation. Curr Biol 15(5):475–479. doi:10.1016/j.cub.2004.12.084

Kane NS, Hirschberg B, Qian S, Hunt D, Thomas B, Brochu R, Ludmerer SW, Zheng Y, Smith M, Arena JP, Cohen CJ, Schmatz D, Warmke J, Cully DF (2000) Drug-resistant

Drosophila indicate glutamate-gated chloride channels are targets for the antiparasitics nodulisporic acid and ivermectin. Proc Natl Acad Sci USA 97(25):13949–13954. doi:10.1073/pnas.240464697

Kim MS, Dufour YS, Yoo JS, Cho YB, Park JH, Nam GB, Kim HM, Lee KL, Donohue TJ, Roe JH (2012) Conservation of thiol-oxidative stress responses regulated by SigR orthologues in actinomycetes. Mol Microbiol 85(2):326–344. doi:10.1111/j.1365-2958.2012.08115.x

Lautru S, Deeth RJ, Bailey LM, Challis GL (2005) Discovery of a new peptide natural product by *Streptomyces coelicolor* genome mining. Nat Chem Biol 1(5):265–269. doi:10.1038/nchembio731

Le TB, Fiedler HP, den Hengst CD, Ahn SK, Maxwell A, Buttner MJ (2009) Coupling of the biosynthesis and export of the DNA gyrase inhibitor simocyclinone in *Streptomyces antibioticus*. Mol Microbiol 72(6):1462–1474. doi:10.1111/j.1365-2958.2009.06735.x

Li CH, Cheng YW, Liao PL, Yang YT, Kang JJ (2010) Chloramphenicol causes mitochondrial stress, decreases ATP biosynthesis, induces matrix metalloproteinase-13 expression, and solid-tumor cell invasion. Toxicol Sci 116(1):140–150. doi:10.1093/toxsci/kfq085

Liu W, Sakr E, Schaeffer P, Talbot HM, Donisi J, Hartner T, Kannenberg E, Takano E, Rohmer M (2014) Ribosylhopane, a novel bacterial hopanoid, as precursor of C35 bacteriohopanepolyols in *Streptomyces coelicolor* A3(2). ChemBioChem 15(14):2156–2161. doi:10.1002/cbic.201402261

Luti KJ, Mavituna F (2011) Elicitation of *Streptomyces coelicolor* with dead cells of *Bacillus subtilis* and *Staphylococcus aureus* in a bioreactor increases production of undecylprodigiosin. Appl Microbiol Biotechnol 90(2):461–466. doi:10.1007/s00253-010-3032-2

Malott RJ, Steen-Kinnaird BR, Lee TD, Speert DP (2012) Identification of hopanoid biosynthesis genes involved in polymyxin resistance in *Burkholderia multivorans*. Antimicrob Agents Chemother 56(1):464–471. doi:10.1128/AAC.00602-11

McCormick JR, Flärdh K (2012) Signals and regulators that govern *Streptomyces* development. FEMS Microbiol Rev 36(1):206–231. doi:10.1111/j.1574-6976.2011.00317.x

McKenzie NL, Nodwell JR (2007) Phosphorylated AbsA2 negatively regulates antibiotic production in *Streptomyces coelicolor* through interactions with pathway-specific regulatory gene promoters. J Bacteriol 189(14):5284–5292. doi:10.1128/JB.00305-07

Meschke H, Walter S, Schrempf H (2012) Characterization and localization of prodiginines from *Streptomyces lividans* suppressing *Verticillium dahliae* in the absence or presence of *Arabidopsis thaliana*. Environ Microbiol 14(4):940–952. doi:10.1111/j.1462-2920.2011.02665.x

Montange RK, Batey RT (2008) Riboswitches: emerging themes in RNA structure and function. Annu Rev Biophys 37:117–133. doi:10.1146/annurev.biophys.37.032807.130000

Moody MJ, Young RA, Jones SE, Elliot MA (2013) Comparative analysis of non-coding RNAs in the antibiotic-producing *Streptomyces* bacteria. BMC Genom 14:558. doi:10.1186/1471-2164-14-558

Nett M, Ikeda H, Moore BS (2009) Genomic basis for natural product biosynthetic diversity in the actinomycetes. Nat Prod Rep 26(11):1362–1384. doi:10.1039/b817069j

Nitzan O, Suponitzky U, Kennes Y, Chazan B, Raul R, Colodner R (2010) Is chloramphenicol making a comeback? Isr Med Assoc J 12(6):371–374

Omura S, Crump A (2004) The life and times of ivermectin—a success story. Nat Rev Microbiol 2(12):984–989. doi:10.1038/nrmicro1048

Omura S, Ikeda H, Ishikawa J, Hanamoto A, Takahashi C, Shinose M, Takahashi Y, Horikawa H, Nakazawa H, Osonoe T, Kikuchi H, Shiba T, Sakaki Y, Hattori M (2001) Genome sequence of an industrial microorganism *Streptomyces avermitilis*: deducing the ability of producing secondary metabolites. Proc Natl Acad Sci USA 98(21):12215–12220. doi:10.1073/pnas.211433198

Onaka H, Mori Y, Igarashi Y, Furumai T (2011) Mycolic acid-containing bacteria induce natural-product biosynthesis in *Streptomyces* species. Appl Environ Microbiol 77(2):400–406. doi:10.1128/AEM.01337-10

Pan JJ, Solbiati JO, Ramamoorthy G, Hillerich BS, Seidel RD, Cronan JE, Almo SC, Poulter CD (2015) Biosynthesis of squalene from farnesyl diphosphate in bacteria: three steps catalyzed by three enzymes. ACS Cent Sci 1(2):77–82. doi:10.1021/acscentsci.5b00115

Perez J, Munoz-Dorado J, Brana AF, Shimkets LJ, Sevillano L, Santamaria RI (2011) *Myxococcus xanthus* induces actinorhodin overproduction and aerial mycelium formation by *Streptomyces coelicolor*. Microb Biotechnol 4(2):175–183. doi:10.1111/j.1751-7915.2010.00208.x

Poralla K, Muth G, Hartner T (2000) Hopanoids are formed during transition from substrate to aerial hyphae in *Streptomyces coelicolor* A3(2). FEMS Microbiol Lett 189(1):93–95

Rigali S, Titgemeyer F, Barends S, Mulder S, Thomae AW, Hopwood DA, van Wezel GP (2008) Feast or famine: the global regulator DasR links nutrient stress to antibiotic production by *Streptomyces*. EMBO Rep 9(7):670–675. doi:10.1038/embor.2008.83

Saenz JP, Sezgin E, Schwille P, Simons K (2012) Functional convergence of hopanoids and sterols in membrane ordering. Proc Natl Acad Sci USA 109(35):14236–14240. doi:10.1073/pnas.1212141109

Saha R, Saha N, Donofrio RS, Bestervelt LL (2013) Microbial siderophores: a mini review. J Basic Microbiol 53(4):303–317. doi:10.1002/jobm.201100552

Schmerk CL, Bernards MA, Valvano MA (2011) Hopanoid production is required for low-pH tolerance, antimicrobial resistance, and motility in *Burkholderia cenocepacia*. J Bacteriol 193(23):6712–6723. doi:10.1128/JB.05979-11

Schrempf H, Koebsch I, Walter S, Engelhardt H, Meschke H (2011) Extracellular *Streptomyces* vesicles: amphorae for survival and defence. Microb Biotechnol 4(2):286–299. doi:10.1111/j.1751-7915.2011.00251.x

Sciara G, Kendrew SG, Miele AE, Marsh NG, Federici L, Malatesta F, SchimpeRNA G, Savino C, Vallone B (2003) The structure of ActVA-Orf6, a novel type of monooxygenase involved in actinorhodin biosynthesis. EMBO J 22(2):205–215. doi:10.1093/emboj/cdg031

Seipke RF, Kaltenpoth M, Hutchings MI (2012) *Streptomyces* as symbionts: an emerging and widespread theme? FEMS Microbiol Rev 36(4):862–876. doi:10.1111/j.1574-6976.2011.00313.x

Swiercz JP, Hindra Bobek J, Haiser HJ, Di Berardo C, Tjaden B, Elliot MA (2008) Small non-coding RNAs in *Streptomyces coelicolor*. Nucleic Acids Res 36(22):7240–7251. doi:10.1093/nar/gkn898

Tahlan K, Ahn SK, Sing A, Bodnaruk TD, Willems AR, Davidson AR, Nodwell JR (2007) Initiation of actinorhodin export in *Streptomyces coelicolor*. Mol Microbiol 63(4):951–961

Traxler MF, Kolter R (2015) Natural products in soil microbe interactions and evolution. Nat Prod Rep 32(7):956–970. doi:10.1039/c5np00013k

Traxler MF, Seyedsayamdost MR, Clardy J, Kolter R (2012) Interspecies modulation of bacterial development through iron competition and siderophore piracy. Mol Microbiol 86(3):628–644. doi:10.1111/mmi.12008

Traxler MF, Watrous JD, Alexandrov T, Dorrestein PC, Kolter R (2013) Interspecies interactions stimulate diversification of the *Streptomyces coelicolor* secreted metabolome. MBio 4(4). doi:10.1128/mBio.00459-13

Uguru GC, Stephens KE, Stead JA, Towle JE, Baumberg S, McDowall KJ (2005) Transcriptional activation of the pathway-specific regulator of the actinorhodin biosynthetic genes in *Streptomyces coelicolor*. Mol Microbiol 58(1):131–150. doi:10.1111/j.1365-2958.2005.04817.x

Valton J, Mathevon C, Fontecave M, Niviere V, Ballou DP (2008) Mechanism and regulation of the two-component FMN-dependent monooxygenase ActVA-ActVB from *Streptomyces coelicolor*. J Biol Chem 283(16):10287–10296. doi:10.1074/jbc.M709730200

Vockenhuber MP, Sharma CM, Statt MG, Schmidt D, Xu Z, Dietrich S, Liesegang H, Mathews DH, Suess B (2011) Deep sequencing-based identification of small non-coding RNAs in *Streptomyces coelicolor*. RNA Biol 8(3):468–477

Wang R, Mast Y, Wang J, Zhang W, Zhao G, Wohlleben W, Lu Y, Jiang W (2013) Identification of two-component system AfsQ1/Q2 regulon and its cross-regulation with GlnR in *Streptomyces coelicolor*. Mol Microbiol 87(1):30–48. doi:10.1111/mmi.12080

Wang W, Ji J, Li X, Wang J, Li S, Pan G, Fan K, Yang K (2014) Angucyclines as signals modulate the behaviors of *Streptomyces coelicolor*. Proc Natl Acad Sci USA 111(15):5688–5693. doi:10.1073/pnas.1324253111

Waters LS, Storz G (2009) Regulatory RNAs in bacteria. Cell 136(4):615–628. doi:10.1016/j.cell.2009.01.043

Williamson NR, Fineran PC, Leeper FJ, Salmond GP (2006) The biosynthesis and regulation of bacterial prodiginines. Nat Rev Microbiol 4(12):887–899. doi:10.1038/nrmicro1531

Wu CH, Kong L, Bialecka-Fornal M, Park S, Thompson AL, Kulkarni G, Conway SJ, Newman DK (2015) Quantitative hopanoid analysis enables robust pattern detection and comparison between laboratories. Geobiology 13(4):391–407. doi:10.1111/gbi.12132

Xu G, Wang J, Wang L, Tian X, Yang H, Fan K, Yang K, Tan H (2010) "Pseudo" gamma-butyrolactone receptors respond to antibiotic signals to coordinate antibiotic biosynthesis. J Biol Chem 285(35):27440–27448. doi:10.1074/jbc.M110.143081

Xu Y, Willems A, Au-Yeung C, Tahlan K, Nodwell JR (2012) A two-step mechanism for the activation of actinorhodin export and resistance in *Streptomyces coelicolor*. MBio 3(5):e00191–00112. doi:10.1128/mBio.00191-12

Yang YL, Xu Y, Straight P, Dorrestein PC (2009) Translating metabolic exchange with imaging mass spectrometry. Nat Chem Biol 5(12):885–887. doi:10.1038/nchembio.252

Chapter 6
Role of Small RNAs in *Wolbachia*-Mosquito Interactions

Sassan Asgari

Abstract *Wolbachia* are endosymbiotic bacteria prevalent in many arthropods, in particular insect species, and nematodes. While they are mostly known for reproductive manipulations of their host, some strains may confer fitness advantages to their host, including protection from virus infection. As a consequence, utilization of *Wolbachia* to suppress transmission of vector-borne viruses and other pathogens has attracted immense interest in recent years. In particular, transinfection of *Wolbachia* strains with strong anti-viral properties into mosquitoes has proven successful in inhibition of a number of mosquito-borne pathogens. While the effects of *Wolbachia* on their hosts have been known for decades, little is known about the molecular mechanisms underlying these interactions. Study of small non-coding RNAs as key regulatory molecules involved in many cellular pathways may provide leads to unravel these molecular mechanisms. Here, recent findings on the role of small RNAs in mediating *Wolbachia*-mosquito interactions are reviewed and discussed.

6.1 Introduction

Mosquitoes are the most important group of insects in regards to human health. Female mosquitoes require a blood meal for egg production. As a consequence, they inflict biting related nuisance and allergic responses, but most importantly can potentially transmit pathogens among their vertebrate hosts. Mosquitoes are vectors of a number of important diseases such as malaria, dengue fever, yellow fever, chikungunya, various virus-induced encephalitis and filariasis. In the 21st century, we are still struggling to control mosquito-borne diseases. This failure is generally attributed to lack of effective vaccines or drugs against the majority of vector-borne pathogens, development of resistance of both vectors and pathogens to insecticides

S. Asgari (✉)
Australian Infectious Disease Research Centre, School of Biological Sciences, The University of Queensland, Brisbane, QLD 4072, Australia
e-mail: s.asgari@uq.edu.au

A.L. Leitão and F.J. Enguita (eds.), *Non-coding RNAs and Inter-kingdom Communication*, DOI 10.1007/978-3-319-39496-1_6

and anti-pathogen drugs, respectively, lack of clear understanding of pathogen-host interactions, and socio-economic issues.

While attempts to develop vaccines and drugs against vector-borne pathogens are continuing, most of the current control measures focus on management of vector populations or blocking replication of pathogens in mosquitoes (McGraw and O'Neill 2013). Besides chemical/biological control of mosquito populations aimed at adult and larval stages, genetic modification of mosquitoes to produce non-viable progeny through sterile insect technique (e.g. Alphey et al. 2010; Carvalho et al. 2015), paratransgenesis (genetic modification of symbiotic bacteria reintroduced into mosquitoes expressing anti-microbial molecules) (Wilke and Marrelli 2015), and utilization of symbiotic bacteria, such as *Wolbachia* (reviewed in Hoffman et al. 2015), have been considered among others.

Wolbachia pipientis is an endosymbiotic alphaproteobacterium with huge strain diversity in the field and a range of different effects on the host. It is estimated that 40–65 % of insect species are infected with *Wolbachia* (Hilgenboecker et al. 2008; Zug and Hammerstein 2012). They are mostly known for reproductive manipulations of their hosts, including cytoplasmic incompatibility (CI), male-killing, induction of parthenogenesis, and feminization (O'Neill et al. 1997). CI is considered as one of the key effects of *Wolbachia* on its hosts, which facilitates its rapid spread and establishment in insect populations due to conferring selective advantage to *Wolbachia*-infected females (Hoffman and Turelli 1997). There are two types of CI: (1) uni-directional CI, in which crosses between uninfected females with infected males lead to the production of non-viable eggs; (2) bi-directional CI takes place when crosses occur between individuals that carry different strains of *Wolbachia* resulting in sterile eggs (Bourtzis et al. 2003) (Fig. 6.1). The exact mechanism underlying CI, which occurs post-fertilization, is not currently well understood; however, it appears that malfunctioning of the parental genomes during embryogenesis could be responsible (Bourtzis et al. 2003; Landmann et al. 2009).

Besides reproductive manipulations, other effects of *Wolbachia* on the natural or transinfected hosts have been observed. For example, a major effect of *Wolbachia* that was discovered first in *Drosophila* and then in other insects is conferring protection against microbial infections, in particular RNA viruses; although this effect appears to be strain and host specific (reviewed in Johnson 2015). In transinfected *Aedes aegypti* with *Wolbachia*, in addition to the CI effect, increase in blood ingestion time, and reduction in the lifespan and number of eggs laid by the mosquitoes as compared with uninfected mosquitoes was observed (Sylvestre et al. 2013). The virus protection discovery in *Drosophila* led to the examination of transinfected *Ae. aegypti* mosquitoes (with *w*MelPop strain of *Wolbachia*) for the same effect. These mosquitoes were originally produced to reduce their life span (McMeniman et al. 2009). Interestingly, it was found that the mosquitoes blocked replication of arboviruses such as dengue and chikungunya viruses, but also a number of other pathogens transmitted by mosquitoes (such as *Plasmodium* and filarial nematodes) (Moreira et al. 2009). Following this, other strains of *Wolbachia* were stably or transiently transinfected into mosquitoes in order to investigate pathogen blocking (e.g. Bian et al. 2010, 2013; Hughes et al. 2011). Success in

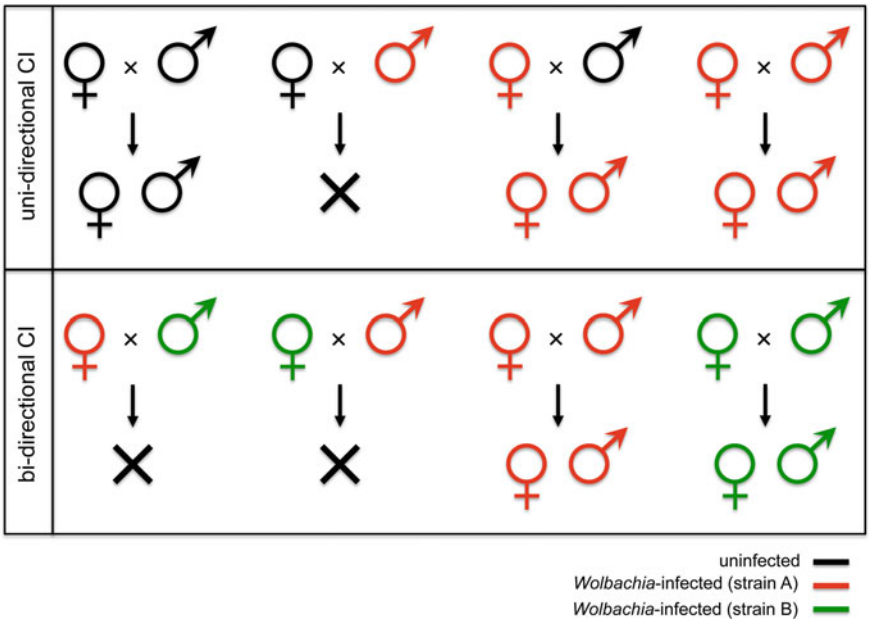


Fig. 6.1 Diagram depicting uni-directional and bi-directional cytoplasmic incompatibility (CI) in insects

laboratory and semi-field trials showing refractoriness of stable *Wolbachia*-transinfected *Ae. aegypti* mosquitoes with *w*Mel and *w*MelPop strains to dengue virus (Walker et al. 2011) led to the release of *Wolbachia*-infected mosquitoes in north Queensland, Australia. Field studies have indicated successful establishment of the mosquitoes (Hoffmann et al. 2011), in particular *w*Mel-transinfected mosquitoes that have remained refractory to dengue virus replication (Hoffmann et al. 2014; Frentiu et al. 2014).

After decades of studies on *Wolbachia* in insects, the underlying molecular mechanisms behind host manipulations remain largely elusive. Exploring host gene regulation in *Wolbachia*-infected hosts provides an approach towards elucidating these underlying mechanisms. microRNAs (miRNAs) as small non-coding RNAs of $\sim$22 nt have proven to play important roles in the regulation of gene expression in most eukaryotes and involved in many cellular processes, including host-pathogen interactions (Asgari 2013; Hussain and Asgari 2014). This chapter aims to review the latest findings with regards to the role of small RNAs, in particular miRNAs, in *Wolbachia*-mosquito interactions.

6.2 Effect of *Wolbachia* on the Host miRNA and piRNA Profiles

Similar to several other insect host-microorganism interactions (Hussain and Asgari 2014), infection with the *w*MelPop strain of *Wolbachia* alters the miRNA profile of *Ae. aegypti* mosquitoes (Hussain et al. 2011). Overall, the abundance of 13 *Ae. aegypti* miRNAs was changed in microarray analysis, nine of which were confirmed with northern hybridization. Among these, aae-miR-2940-5p was highly induced in *Wolbachia*-infected mosquitoes suggesting its possible involvement in *Wolbachia*-mosquito interactions. Interestingly, based on reported miRNAs on miRBase, aae-miR-2940-5p appears to be only present in mosquitoes and not in other insects including other members of the order Diptera to which mosquitoes belong. Furthermore, this miRNA was downregulated in *Ae. albopictus* mosquitoes infected with chikungunya virus (Shrinet et al. 2014), and dengue virus (DENV) (Liu et al. 2015) and *Ae. albopictus* C6/36 cells infected with West Nile virus (Slonchak et al. 2014). While the genes regulated by miR-2940 are not known in most of these interactions, altogether, evidence suggests that miR-2940 may play a role in mosquito-pathogen interactions (see below).

In addition to changes in the abundance of host miRNAs demonstrated by microarray analysis in whole mosquitoes, deep sequencing of small RNAs from *w*MelPop-CLA infected Aag2 mosquito cells showed that modifications of mature miRNAs representing isomiRs of miRNAs increase as compared to non-infected cells (Mayoral et al. 2014a). These modifications include 5′ and 3′ extensions, 3′ addition (non-template addition of nucleotides) and 5′ and 3′ trimming as well as nucleotide substitutions in the middle of the mature miRNA. Among these, the 3′ extension is more common in *Wolbachia*-infected cells. While the impact of these changes in *Wolbachia*-infected cells on the target repertoire of the miRNAs is not known, bioinformatics analysis suggests significant shifts in the predicated targets of isomiRs. For example, in DENV-infected *Ae. aegypti* mosquitoes (Etebari et al. 2015) and human dendritic cells infected with *Mycobacterium tuberculosis* (Siddle et al. 2015), a large number of predicted targets of miRNA sequences were not predicted targets of the alternative 5′ shift isomiRs. This change in the target profile may have a profound impact on miRNA-mediated functions.

As well as miRNAs, comparison of deep sequencing reads in *Wolbachia*-infected and non-infected mosquito cells revealed an overall increase in the abundance of piRNAs in infected cells, with few exceptions (Mayoral et al. 2014a). piRNAs or PIWI-interacting RNAs are generally around 27 nt and mostly involved in the protection of the genome from transposable elements, in particular in germ cells (Chambeyron and Seitz 2014; Ishizu et al. 2012). Further, they have been implicated in functions outside of the germ cells, e.g. regulating memory storage in the brain, dendritic spine development (reviewed in Ishizu et al. 2012) and in host-virus interactions as knock-down of the pathway leads to increased virus replication (Miesen et al. 2015; Morazzani et al. 2012; Schnettler et al. 2013). The role of piRNAs, in particular those differentially abundant in the presence of *Wolbachia* remain to be explored.

6.3 Role of Mosquito miRNAs in *Wolbachia* Maintenance

Collected evidence suggests that differentially abundant miRNAs in *Wolbachia*-infected mosquitoes affect *Wolbachia* replication/maintenance. The highly induced aae-miR-2940-5p targets at least three genes that all seem to contribute to *Wolbachia* maintenance since manipulations of the transcript levels affect the endosymbiont's density in the host tissues. The first target of aae-miR-2940-5p identified was the mosquito's *metalloprotease FtsH* (*MetP*), which was found highly upregulated in *Wolbachia*-infected mosquitoes (Hussain et al. 2011). The target sequences reside in the 3′ UTR of *MetP* mRNA and although the mechanism of the positive interaction is not currently known, it is postulated that the interaction most likely results in the stabilization of the mRNA as shown for other instances (Bruno et al. 2011; Vasudevan et al. 2007). Cloning the MetP target sequences downstream of the *GFP* reporter gene also caused increases in the *GFP* transcript levels. Interestingly, when either the *MetP* or aae-miR-2940 were knocked down in mosquitoes by injection of *MetP* dsRNA or inhibitor of aae-miR-2940-5p, *Wolbachia* density significantly declined in the mosquitoes (Hussain et al. 2011). These observations demonstrated that upregulation of MetP via induction of aae-miR-2940-5p is essential for *Wolbachia* maintenance. The function of this protein in insects and the mechanism by which the protein may affect *Wolbachia* density is not known. However, in mammalian cells, FtsH metalloproteases are membrane-bound and ATP-dependent and are mainly responsible for controlling the quality of integral membrane proteins in mitochondria by degrading incorrectly folded proteins (Ito and Akiyama 2005). Knockdown of the gene in mammalian cells leads to increase in nonassembled and oxidized proteins in the mitochondrial membrane, impaired cell proliferation and susceptibility to apoptotic stimuli (Stiburek et al. 2012).

DNA methyltransferase 2 (*Dnmt2*) was identified as another target of aae-miR-2940-5p. In contrast to *MetP*, *AeDnmt2* is negatively regulated by the miRNA. As a consequence, in *Wolbachia*-infected mosquitoes in which the miRNA is highly induced, AeDnmt2 levels are suppressed. The downregulation of the target gene was found beneficial to *Wolbachia* since its overexpression by ectopic expression in mosquito cells infected with *Wolbachia* was followed by reductions in the endosymbiont's density. Similarly, overexpression of Dnmt2 in *Drosophila melanogater* also led to reduction in *Wolbachia* density (LePage et al. 2014). Interestingly, a general reduction in the methylation of *Ae. aegypti* genome was observed when infected with *Wolbachia w*MelPop strain as compared with uninfected mosquitoes (Ye et al. 2013), which coincides with the downregulation of AeDnmt2 in *Wolbachia*-infected mosquitoes. In *Ae. aegypti*, similar to *D. melanogaster*, there is only one DNA methyltransferase gene annotated in the gnome, which is *Dnmt2*. However, reports suggest that this gene may not be involved in DNA methylation but rather RNA methylation (Schaefer et al. 2010; Tuorto et al. 2012). In *D. melanogaster Dnmt2*-deficient flies, the level of genome methylation was not different to wild-type flies (Raddatz et al. 2013) further confirming that

Dnmt2 is not involved in genome methylation. However, while genome methylation occurs in *Ae. aegypti* (Ye et al. 2013), it is not known whether the Dnmt2 homolog (AeDnmt2) is involved in DNA methylation or another unknown methyltransferase enzyme carries out methylation of the genome, similar to *D. melanogaster*.

Another target of aae-miR-2940-5p was identified as *Arginine methyltransferase 3 (ArgMet-3)*, which is also positively regulated by the miRNA (Zhang et al. 2014). Consequently, in *Wolbachia*-infected mosquitoes higher levels of ArgMet-3 were detected as compared to non-infected mosquitoes. Similar to MetP, silencing of *ArgMet-3* in *Wolbachia*-infected mosquito cells led to reduction in *Wolbachia* density in the cells. In addition to aae-miR-2940-5p, aae-miR-12 was also found in higher abundance in *Wolbachia*-infected mosquitoes (Hussain et al. 2011). Two targets of this miRNA, *DNA replication licensing factor (MCM6)* and *monocarboxylate transporter (MCT1)*, were found in lower abundance as a consequence of the upregulation of aae-miR-12 (Osei-Amo et al. 2012). Inhibition of aae-miR-12 by transfecting *Wolbachia*-infected mosquito cells with the synthetic inhibitor of the miRNA resulted in significant reductions in *Wolbachia* density in the host cells implying that it plays a role in *Wolbachia* persistence in the host cells.

Overall, the findings summarized above propose that alterations of the host miRNA profile in *Wolbachia*-infected cells and the genes they regulate are important in colonization and persistence of the endosymbiont in the host cells. The link between the different targets of the identified miRNAs and how they influence *Wolbachia* density remain to be explored. These effects on *Wolbachia* density could very well be indirect, for example, by affecting nutrient availability, cell proliferation, etc. rather than directly regulating colonization.

6.4 Effect of *Wolbachia* on Cellular Distribution of Agos and miRNAs

While the majority of miRNAs are found in the cytoplasm, it has been shown both in mammals and insects that a certain number of miRNAs could be translocated into the nucleus in association with Ago proteins (Hwang et al. 2007; Jeffries et al. 2011; Hussain et al. 2013b). These are most likely involved in transcriptional gene silencing (TGS) (e.g. Ameyar-Zazoua et al. 2012). The presence of Ago1 (mainly associated with the miRNA pathway) and Ago2 (associated with siRNA and miRNA pathways) in the nuclei of *Ae. aegypti* and *D. melanogater* was demonstrated (Hussain et al. 2013b). However, in *Wolbachia*-infected insects, less Ago1 was detected in the nucleus as compared to non-infected cells, while Ago2's presence in the nucleus was not affected (Hussain et al. 2013b). Further analysis revealed that importin β-4, that facilitates Ago1 distribution to the nucleus in mosquito cells, is downregulated by aae-miR-981, which is induced in *Wolbachia*-infected mosquitoes. Does this affect cellular distribution of miRNAs in *Wolbachia*-infected cells?

Deep sequencing analysis of small RNAs of nuclear and cytoplasmic fractions of mosquito Aag2 cells (derived from *Ae. aegypti*) showed changes in the abundance of a number of miRNAs when *Wolbachia*-infected and uninfected cells were compared (Mayoral et al. 2014a). Considering the 20 top abundant miRNAs in the nucleus of the mosquito cells, the ranking of 14 out of the 20 miRNAs changed in *Wolbachia*-infected cells as compared to uninfected cells. In *Wolbachia*-infected cells, aae-miR-989 and aae-miR-306-5p disappeared from the list of top 20 miRNAs in the nuclei, but on the other hand aae-miR-2765 and aae-bantam-3p appeared on the top 20 abundant list in *Wolbachia*-infected cells (Mayoral et al. 2014a). These changes coincided with alterations in the abundance of most of these miRNAs in the cytoplasmic fractions from infected and non-infected cells suggesting that distribution of the miRNAs between the two compartments might have been affected due to *Wolbachia* infection.

6.5 Effect of *Wolbachia*-Altered Host miRNAs on Mosquito-Virus Interactions

Conferring protection against virus infection by *Wolbachia* is a significant fitness advantage for infected insects, which has been utilized as an approach to reduce transmission of arboviruses. While a number of reports have shown contribution of competition for host resources, autophagy and immunity in the virus protection property of *Wolbachia* (reviewed in Rainey et al. 2014; Martinez et al. 2014), the exact mechanism(s) underlying this protection is largely unknown. Important known facts in this regard are that the protection is not systemic in insects (i.e. presence of *Wolbachia* in the cell is absolutely required for virus protection) (Moreira et al. 2009), and also *Wolbachia* density has a direct effect on the level of protection conferred (Osborne et al. 2012; Lu et al. 2012).

AaDnmt2 was found to enhance replication of dengue virus serotype 2 (DENV-2) in mosquito Aag2 cells without *Wolbachia* infection. This suggested that downregulation of AaDnmt2 in *Wolbachia*-infected *Ae. aegypti* mosquitoes, in which aae-miR-2940-5p is upregulated, may contribute to the virus protection effect observed in the mosquito. However, the expression of the homologous gene in a number of *Drosophila* species infected with various strains of *Wolbachia* was not significantly different to that of non-infected flies and did not show a correlation with the level of protection against Drosophila C virus or Flock house virus infections (Martinez et al. 2014). Therefore, the antiviral protection effect mediated by *Wolbachia* by altering Dnmt2 levels might be host/virus/*Wolbachia* strain specific and not a ubiquitous mechanism for *Wolbachia*-mediated protection. Also, it is worth mentioning that the *Wolbachia-Ae. aegypti* association is very recent as compared to very long association of *Wolbachia* with the flies, and the effect of AaDnmt2 on DENV was tested in a mosquito cell line, whereas the work carried out in *Drosophila* spp was done in whole flies.

MetP was found to facilitate West Nile virus (WNV) replication in *Aedes albopictus* C6/36 cells (Slonchak et al. 2014), and specific downregulation of aae-miR-2940-5p in the cells was assumed as an antiviral response of the host cells since the miRNA positively regulates *MetP* transcript levels (Hussain et al. 2011). Upregulation of aae-miR-2940-5p in *Wolbachia*-infected cells should therefore enhance WNV replication. This was in fact observed in *Wolbachia*-infected cells but only at the genomic RNA and protein levels (Hussain et al. 2013a). Titration of virions produced by the cells indicated suppression of virus production in the presence of *Wolbachia* both in cell line and mosquitoes, which could presumably be due to interference with viral assembly or secretion.

6.6 *Wolbachia* Small RNAs and Their Role in Host Regulation

In recent years, accumulating evidence suggests that exchanged small non-coding RNAs between microorganisms and their eukaryotic host may play a role in cross-kingdom communications. For example, miRNAs from sickle cell erythrocytes were found to be transferred to *Plasmodium falciparum* malaria parasite regulating the cAMP-dependent Protein Kinase (PKA-R) (LaMonte et al. 2012). As a consequence of this interaction, translation is inhibited in the parasite leading to the resistance of the sickle cells to malaria infection. Furthermore, other examples are covered in the other chapters in this volume. Analysis of small RNAs obtained by deep sequencing in *w*MelPop-infected Aag2 cells that mapped to the *Wolbachia* genome suggested that miRNA-like small RNAs could be produced by the endosymbiont (Mayoral et al. 2014b). Follow up experiments validated two *Wolbachia* small non-coding RNAs (*W*snRNAs 46 and -49) of about 30 nucleotides. While the majority of snRNAs found in bacteria are 50–200 nt (Caldelari et al. 2012), recent studies suggest that smaller ncRNAs with precursor stem-loops resembling pre-miRNAs could also be produced in bacteria (Shmaryahu et al. 2014; Furuse et al. 2014). However, the mechanism of the biogenesis of these miRNA-like snRNAs in bacteria is unknown. The small RNAs were also found in *D. melanogaster* and *D. simulans* infected with *w*MelPop and *w*Au strains, respectively (Mayoral et al. 2014b). This detection and further sequence homology searches indicated that the precursor and mature *W*snRNAs are conserved in most supergroup A *Wolbachia* strains. There are currently eight *Wolbachia* supergroups known (A-H), with most insect *Wolbachia* strains belonging to supergroups A and B (Werren et al. 1995).

Interestingly, *W*snRNAs were detected in the host cell cytoplasm suggesting that they could be exported from *Wolbachia* cells into the host (Mayoral et al. 2014b). In silico searches identified a number of potential host genes as targets of the *W*snRNAs. One of the targets that was experimentally validated is *Dynein heavy chain* (*Dhc*). In this instance, the target gene is positively regulated by *W*snRNA-46

(Mayoral et al. 2014b). Firstly, *Dhc* transcript levels were significantly higher in *Wolbachia*-infected mosquitoes as compared to non-infected ones. Secondly, injection of the mimic of *W*snRNA-46 into non-infected *Ae. aegypti* mosquitoes eventuated in increased transcript levels of *Dhc*. Thirdly, transfection of *W*snRNA-46 mimic into Aag2 cells led to increased transcript levels of the reporter gene *GFP* downstream of which *Dhc* target sequences were cloned. While the mechanism of positive regulation in this instance is not known, this type of interaction is not rare as several other examples have shown positive interactions of miRNAs with targets (e.g. Conrad et al. 2013; Orom et al. 2008; Bruno et al. 2011; Ma et al. 2010; Vasudevan 2012; Henke et al. 2008; Hussain et al. 2011).

In addition to regulation of *Dhc* host gene by a *W*snRNA, potential target genes of *W*snRNA-46 and -59 were also found in the bacterial genome (Mayoral et al. 2014b). One of these targets, UDP-N-acetylmuramoyalanine-D-glutamate ligase (*murD*), showed a negative correlation with *W*snRNA-46. Firstly, inhibition of the small RNA by transfection of *w*MelPop infected cells with the synthetic inhibitor of *W*snRNA-46 eventuated in higher transcript levels of *murD*. Secondly, when *murD* target sequences were cloned downstream of the *GFP* gene and co-transfected with the mimic of *W*snRNA-46 into mosquito Aag2 cells, significantly less *GFP* transcript levels were found.

In addition to *W*snRNAs that resemble miRNAs and are around 30 nucleotides, two putative snRNAs expressed from the intergenic regions in *w*Mel strain were reported (Woolfit et al. 2015). Of these two, one (ncrwmel02, non-coding RNA *Wolbachia w*Mel 02) is conserved in *w*MelCS, *w*MelPop, *w*Mel and *w*Au strains of *Wolbachia*. Expression of ncrwmel02 is regulated in a sex- and tissue-specific manner. The exact size of these non-coding RNAs is not known and whether they are functional remains to be shown.

6.7 Conclusions

Our understanding of the role of small non-coding RNAs in regulation of various biological aspects of living organisms is rapidly increasing. More recent findings allude to the role of sncRNAs in interactions of microorganisms with their hosts, whether these interactions are pathogenic or mutualistic. In the evolutionary arm race between hosts and pathogens, this may mean manipulation of host genes by the pathogen to reduce immune responses of the host or its recognition as foreign object. On the other hand, sncRNAs from the host may regulate pathogens genes to reduce their replication or survival. In mutualistic or endosymbiotic associations, sncRNAs could be utilized for establishment and maintenance of the association. In this area, we have only scratched the surface and follow up studies will determine the significance of sncRNAs in cross-kingdom communications, in particular between *Wolbachia* and insect hosts. Better understanding of the role of sncRNAs in the associations will help us to optimize utilization of microorganisms involved in the interactions.

Acknowledgments I would like to thank Dr Beth McGraw for helpful comments on the manuscript. I also acknowledge the contributions made by my lab members over the years that have been cited here. Work in my laboratory has been supported by the Australian Research Council (DP110102112) and National Health and Medical Research (APP1062983, APP1027110).

References

Alphey L, Benedict M, Bellini R, Clark G, Dame D, Service M, Dobson S (2010) Sterile-insect methods for control of mosquito-borne diseases: an analysis. Vector Borne Zoonotic Dis 10:295–311. doi:10.1089/vbz.2009.0014

Ameyar-Zazoua M, Rachez C, Souidi M, Robin P, Fritsch L, Young R, Morozova N, Fenouil R, Descostes N, Andrau J, Mathieu J, Hamiche A, Ait-Si-Ali S, Muchardt C, Batsché E, Harel-Bellan A (2012) Argonaute proteins couple chromatin silencing to alteRNAtive splicing. Nat Stru Mol Biol 19:998–1004. doi:10.1038/nsmb.2373

Asgari S (2013) MicroRNA functions in insects. Insect Biochem Mol Biol 43:388–397. doi:10.1016/j.ibmb.2012.10.005

Bian G, Xu Y, Lu P, Xie Y, Xi Z (2010) The endosymbiotic bacterium *Wolbachia* induces resistance to Dengue virus in *Aedes aegypti*. PLoS Pathog 6:e1000833. doi:10.1126/science.1236192

Bian G, Joshi D, Dong Y, Lu P, Zhou G, Pan X, Xu Y, Dimopoulos G, Xi Z (2013) *Wolbachia* invades *Anopheles stephensi* populations and induces refractoriness to *Plasmodium* infection. Science 340:748–751. doi:10.1371/journal.ppat.1000833

Bourtzis K, Braig HR, Karr TL (2003) Cytoplasmic incompatibility. In: Bourtzis K, Miller TA (eds) Insect Symbiosis. CRC Press, FL, USA, pp 217–246

Bruno IG, Karam R, Huang L, Bhardwaj A, Lou CH, Shum EY, Song HW, Corbett MA, Gifford WD, Gecz J, Pfaff SL, Wilkinson MF (2011) Identification of a microRNA that activates gene expression by repressing nonsense-mediated RNA decay. Mol Cell 42:500–510. doi:10.1016/j.molcel.2011.04.018

Caldelari I, Chao Y, Romby PJV (2012) RNA-mediated regulation in pathogenic bacteria. Cold Spring Harb Perspect Med 3:a010298. doi:10.1101/cshperspect.a010298

Carvalho D, McKemey A, Garziera L, Lacroix R, Donnelly C, Alphey L, Malavasi A, Capurro M (2015) Suppression of a field population of *Aedes aegypti* in Brazil by sustained release of transgenic male mosquitoes. PLoS Neg Trop Dis 9:e0003864. doi:10.1371/journal.pntd.0003864

Chambeyron S, Seitz H (2014) Insect small non-coding RNA involved in epigenetic regulations. Curr Opin Insect Sci 1:1–9. doi:10.15252/embr.201439092

Conrad K, Giering F, Erfurth C, Neumann A, Fehr C, Meister G, Niepmann M (2013) MicroRNA-122 dependent binding of Ago2 protein to hepatitis C virus RNA is associated with enhanced RNA stability and translation stimulation. PLoS ONE 8:e56272. doi:10.1371/journal.pone.0056272

Etebari K, Osei-Amo S, Blomberg S, Asgari S (2015) Dengue virus infection alters post-transcriptional modification of microRNAs in the mosquito vector *Aedes aegypti*. Sci Rep 5:15968. doi:10.1038/srep15968

Frentiu FD, Zakir T, Walker T, Popovici J, Pyke AT, van den Hurk A, McGraw EA, O'Neill SL (2014) Limited dengue virus replication in field-collected *Aedes aegypti* mosquitoes infected with *Wolbachia*. Plos Neglect Trop Dis 8:e2688. doi:10.1371/journal.pntd.0002688

Furuse Y, Finethy R, Saka H, Xet-Mull A, Sisk D, Smith K, Lee S, Coers J, Valdivia R, Tobin D, Cullen B (2014) Search for microRNAs expressed by intracellular bacterial pathogens in infected mammalian cells. PLoS ONE 9:e106434. doi:10.1371/journal.pone.0106434

Henke JI, Goergen D, Zheng J, Song Y, Schüttler CG, Fehr C, Jünemann C, Niepmann M (2008) microRNA-122 stimulates translation of hepatitis C virus RNA. EMBO J 27:3300–3310. doi:10.1038/emboj.2008.244

Hilgenboecker K, Hammerstein P, Schlattmann P, Telschow A, Werren JH (2008) How many species are infected with *Wolbachia*?—A statistical analysis of current data. FEMS Microbiol Lett 218:215–220. doi:10.1111/j.1574-6968.2008.01110.x

Hoffmann A, Iturbe-Ormaetxe I, Callahan A, Phillips B, Billington K, Axford J, Montgomery B, Turley A, O'Neill S (2014) Stability of the *w*Mel *Wolbachia* infection following invasion into *Aedes aegypti* populations. PLoS Neg Trop Dis 8:e3115. doi:10.1371/journal.pntd.0003115

Hoffman A, Ross P, Rašić G (2015) *Wolbachia* strains for disease control: ecological and evolutionary considerations. Evol Appl 8:751–768. doi:10.1111/j.1574-6968.2008.01110.x

Hoffman AA, Turelli M (1997) Cytoplasmic incompatibility in insects. In: O'Neill SL, Hoffman AA, Werren JH (eds) Influential passengers: inherited microorganisms and arthropod reproduction. Oxford University Press, Oxford, pp 42–80

Hoffmann AA, Montgomery BL, Popovici I, Iturbe-Ormaetxe I, Johnson PH, Muzzi F, Greenfield M, Durkan M, Leong YS, Dong Y, Cook H, Axford J, Callahan AG, Kenny N, Omodei C, McGraw EA, Ryan PA, Ritchie SA, Turelli M, O'Neill SL (2011) Successful establishment of *Wolbachia* in *Aedes* populations to suppress dengue transmission. Nature 476:454–457. doi:10.1038/nature10356

Hughes GL, Koga R, Xue P, Fukatsu T, Rasgon JL (2011) *Wolbachia* infections are virulent and inhibit the human malaria parasite *Plasmodium falciparum* in *Anopheles gambiae*. PLoS Pathog 7:e1002043. doi:10.1371/journal.ppat.1002043

Hussain M, Asgari S (2014) MicroRNAs as mediators of insect host–pathogen interactions and immunity. J Insect Physiol 70:151–158. doi:10.1016/j.jinsphys.2014.08.003

Hussain M, Frentiu FD, Moreira LA, O'Neill SL, Asgari S (2011) *Wolbachia* utilizes host microRNAs to manipulate host gene expression and facilitate colonization of the dengue vector *Aedes aegypti*. Proc Natl Acad Sci USA 108:9250–9255. doi:10.1073/pnas.1105469108

Hussain M, Lu G, Torres S, Edmonds JH, Kay BH, Khromykh AA, Asgari S (2013a) Effect of *Wolbachia* on replication of West Nile Virus in mosquito cell line and adult mosquitoes. J Virol 87:851–858. doi:10.1128/JVI.01837-12

Hussain M, O'Neill SL, Asgari S (2013b) *Wolbachia* interferes with the intracellular distribution of Argonaute 1 in the dengue vector *Aedes aegypti* by manipulating the host microRNAs. RNA Biol 10:1868–1875. doi:10.4161/rna.27392

Hwang H, Wentzel E, Mendell J (2007) A hexanucleotide element directs microRNA nuclear import. Science 315:97–100

Ishizu H, Siomi H, Siomi M (2012) Biology of PIWI-interacting RNAs: new insights into biogenesis and function inside and outside of germlines. Gene Dev 26:2361–2373. doi:10.1101/gad.203786.112

Ito K, Akiyama Y (2005) Cellular functions, mechanism of action, and regulation of FtsH protease. Annu Rev Microbiol 59:211–231

Jeffries CD, Fried HM, Perkins DO (2011) Nuclear and cytoplasmic localization of neural stem cell microRNAs. RNA 17:675–686. doi:10.1261/rna.2006511

Johnson KN (2015) Bacteria and antiviral immunity in insects. Curr Opin Insect Sci 8:97–103. doi:10.1016/j.cois.2015.01.008

LaMonte G, Philip N, Reardon J, Lacsina JR, Majoros W, Chapman L, Thornburg CD, Telen MJ, Ohler U, Nicchitta CV, Haystead T, Chi JT (2012) Translocation of sickle cell erythrocyte microRNAs into *Plasmodium falciparum* inhibits parasite translation and contributes to malaria resistance. Cell Host Microbe 12:187–199. doi:10.1016/j.chom.2012.06.007

Landmann F, Orsi GA, Loppin B, Sullivan W (2009) *Wolbachia*-mediated cytoplasmic incompatibility is associated with impaired histone deposition in the male pronucleus. PLoS Pathog 5:e1000343. doi:10.1371/journal.ppat.1000343

LePage D, Jernigan K, Bordenstein S (2014) The relative importance of DNA methylation and Dnmt2-mediated epigenetic regulation on *Wolbachia* densities and cytoplasmic incompatibility. PeerJ 2:e678. doi:10.7717/peerj.678

Liu Y, Zhou Y, Wu J, Zheng P, Li Y, Zheng X, Puthiyakunnon S, Tu Z, Chen X (2015) The expression profile of *Aedes albopictus* miRNAs is altered by dengue virus serotype-2 infection. Cell Biosci 5:16. doi:10.1186/s13578-015-0009-y

Lu P, Bian G, Pan X, Xi Z (2012) *Wolbachia* induces density-dependent inhibition to dengue virus in mosquito cells. Plos Neglect Trop Dis 6:e1754. doi:10.1371/journal.pntd.0001754

Ma F, Liu XG, Li D, Wang P, Li N, Lu L, Cao X (2010) MicroRNA-466 l upregulates IL-10 expression in TLR-triggered macrophages by antagonizing RNA-binding protein tristetraprolin-mediated IL-10 mRNA degradation. J Immunol 184:6053–6059. doi:10.4049/jimmunol.0902308

Martinez J, Longdon B, Bauer S, Chan Y, Miller W, Bourtzis K, Teixeira L, Jiggins F (2014) Symbionts commonly provide broad spectrum resistance to viruses in insects: a comparative analysis of *Wolbachia* strains. PLoS Pathog 10:e1004369. doi:10.1371/journal.ppat.1004369

Mayoral JG, Etebari K, Hussain M, Khromykh AA, Asgari S (2014a) *Wolbachia* infection modifies the profile, shuttling and structure of microRNAs in a mosquito cell line. PLoS ONE 9:e96107. doi:10.1371/journal.pone.0096107

Mayoral JG, Hussain M, Joubert DA, Iturbe-Ormaetxe I, O'Neill SL, Asgari S (2014b) *Wolbachia* small non-coding RNAs and their role in cross-kingdom communications. Proc Natl Acad Sci USA 111:18721–18726. doi:10.1073/pnas.1420131112

McGraw EA, O'Neill SL (2013) Beyond insecticides: new thinking on an ancient problem. Nat Rev Microbiol 11:181–193. doi:10.1038/nrmicro2968

McMeniman CJ, Lane RV, Cass BN, Fong AWC, Sidhu M, Wang Y-F, O'Neill SL (2009) Stable introduction of a life-shortening *Wolbachia* infection into the mosquito *Aedes aegypti*. Science 323:141–144. doi:10.1126/science.1165326

Miesen P, Girardi E, van Rij RP (2015) Distinct sets of PIWI proteins produce arbovirus and transposon-derived piRNAs in *Aedes aegypti* mosquito cells. Nucleic Acids Res 43:6545–6556. doi:10.1093/nar/gkv590

Morazzani E, Wiley M, Murreddu M, Adelman Z, Myles K (2012) Production of virus-derived ping-pong-dependent piRNA-like small RNAs in the mosquito soma. PLoS Pathog 8: e1002470. doi:10.1371/journal.ppat.1002470

Moreira LA, Iturbe-Ormaetxe I, Jeffery JA, Lu GJ, Pyke AT, Hedges LM, Rocha BC, Hall-Mendelin S, Day A, Riegler M, Hugo LE, Johnson KN, Kay BH, McGraw EA, van den Hurk AF, Ryan PA, O'Neill SL (2009) A *Wolbachia* symbiont in *Aedes aegypti* limits infection with Dengue, Chikungunya, and *Plasmodium*. Cell 139:1268–1278. doi:10.1016/j.cell.2009.11.042

O'Neill S, Hoffmann AA, Werren JH (1997) Influential passengers: inherited microorganisms and arthropod reproduction. Oxford University Press, Oxford

Orom UA, Nielsen FC, Lund AH (2008) MicroRNA-10a binds the 5′ UTR of ribosomal protein mRNAs and enhances their translation. Mol Cell 30:460–471. doi:10.1016/j.molcel.2008.05.001

Osborne SE, Iturbe-Ormaetxe I, Brownlie JC, O'Neill SL, Johnson KN (2012) Antiviral protection and the importance of *Wolbachia* density and tissue tropism in *Drosophila*. Appl Environ Microbiol 78:6922–6929. doi:10.1128/AEM.01727-12

Osei-Amo S, Hussain M, O'Neill SL, Asgari S (2012) *Wolbachia*-induced aae-miR-12 miRNA negatively regulates the expression of MCT1 and MCM6 genes in *Wolbachia*-infected mosquito cell line. PLoS ONE 7:e50049. doi:10.1371/journal.pone.0050049

Raddatz G, Guzzardo P, Olova N, Fantappié M, Rampp M, Schaefer M, Reik W, Hannon G, Lyko F (2013) Dnmt2-dependent methylomes lack defined DNA methylation patterns. Proc Natl Acad Sci USA 110:8627–8631. doi:10.1073/pnas.1306723110

Rainey S, Shah P, Kohl A, Dietrich I (2014) Understanding the *Wolbachia*-mediated inhibition of arboviruses in mosquitoes: progress and challenges. J Gen Virol 95:517–530. doi:10.1099/vir.0.057422-0

Schaefer M, Pollex T, Hanna K, Tuorto F, Meusburger M, Helm M, Lyko F (2010) RNA methylation by Dnmt2 protects transfer RNAs against stress induced cleavage. Genes Dev 24:1590–1595. doi:10.1101/gad.586710

Schnettler E, Donald CL, Human S, Watson M, Siu RW, McFarlane M, Fazakerley JK, Kohl A, Fragkoudis R (2013) Knockdown of piRNA pathway proteins results in enhanced Semliki Forest virus production in mosquito cells. J Gen Virol 94:1680–1689. doi:10.1099/vir.0.053850-0

Shmaryahu A, Carrasco M, Valenzuela P (2014) Prediction of bacterial microRNAs and possible targets in human cell transcriptome. J Microbiol 52:482–489. doi:10.1007/s12275-014-3658-3

Shrinet J, Jain S, Jain J, Bhatnagar R, Sunil S (2014) Next generation sequencing reveals regulation of distinct *Aedes* microRNAs during chikungunya virus development. Plos Neglect Trop Dis 8:e2616. doi:10.1371/journal.pntd.0002616

Siddle K, Tailleux L, Deschamps M, Loh Y, Deluen C, Gicquel B, Antoniewski C, Barreiro L, Farinelli L, Quintana-Murci L (2015) Bacterial infection drives the expression dynamics of microRNAs and their isomiRs. PLoS Genet 11:e1005064. doi:10.1371/journal.pgen.1005064

Slonchak A, Hussain M, Torres Morales S, Asgari S, Khromykh AA (2014) Expression of mosquito microRNA aae-miR-2940-5p is down-regulated in response to West Nile virus infection to restrict viral replication. J Virol 88:8457–8467. doi:10.1128/JVI.00317-14

Stiburek L, Cesnekova J, Kostkova O, Fornuskova D, Vinsova K, Wenchich L, Houstek J, Zeman J (2012) YME1L controls the accumulation of respiratory chain subunits and is required for apoptotic resistance, cristae morphogenesis, and cell proliferation. Mol Biol Cell 23:1010–1023. doi:10.1091/mbc.E11-08-0674

Sylvestre G, Gandini M, Maciel-de-Freitas R (2013) Age-dependent effects of oral infection with dengue virus on *Aedes aegypti* (Diptera: Culicidae) feeding behavior, survival, oviposition success and fecundity. PLoS ONE 8:e59933. doi:10.1371/journal.pone.0059933

Tuorto F, Liebers R, Musch T, Schaefer M, Hofmann S, Kellner S, Frye M, Helm M, Stoecklin G, Lyko F (2012) RNA cytosine methylation by Dnmt2 and NSun2 promotes tRNA stability and protein synthesis. Nat Stru Mol Biol 19:900–905. doi:10.1038/nsmb.2357

Vasudevan S (2012) Posttranscriptional upregulation by microRNAs. RNA 3(3):311–330. doi:10.1002/wrna.121

Vasudevan S, Tong Y, Steitz JA (2007) Switching from repression to activation: microRNAs can up-regulate translation. Science 318:1931–1934

Walker T, Johnson PH, Moreira LA, Iturbe-Ormaetxe I, Frentiu FD, McMeniman CJ, Leong YS, Dong Y, Axford J, Kriesner P, Lloyd AL, Ritchie SA, O'Neill SL, Hoffman AA (2011) The *w*Mel *Wolbachia* strain blocks dengue and invades caged *Aedes aegypti* populations. Nature 476:450–453. doi:10.1038/nature10355

Werren J, Zhang W, Guo L (1995) Evolution and phylogeny of *Wolbachia*: reproductive parasites of arthropods. Proc R Soc Lond B 261:55–63

Wilke A, Marrelli M (2015) Paratransgenesis: a promising new strategy for mosquito vector control. Parasit Vectors 8:342. doi:10.1186/s13071-015-0959-2

Woolfit M, Algama M, Keith J, McGraw E, Popovici J (2015) Discovery of putative small non-coding RNAs from the obligate intracellular bacterium *Wolbachia pipientis*. PLoS ONE 10:e0118595. doi:10.1371/journal.pone.0118595

Ye YH, Woolfit M, Huttley GA, Rancès E, Caragata EP, Popovici J, O'Neill SL, McGraw EA (2013) Infection with a virulent strain of *Wolbachia* disrupts genome wide-patterns of cytosine methylation in the mosquito *Aedes aegypti*. PLoS ONE 8:e66482. doi:10.1371/journal.pone.0066482

Zhang G, Hussain M, Asgari S (2014) Regulation of arginine methyltransferase 3 by a *Wolbachia*-induced microRNA in *Aedes aegypti* and its effect on *Wolbachia* and dengue virus replication. Insect Biochem Mol Biol 53:81–88. doi:10.1016/j.ibmb.2014.08.003

Zug R, Hammerstein P (2012) Still a host of hosts for *Wolbachia*: analysis of recent data suggests that 40 % of terrestrial arthropod species are infected. PLoS ONE 7:e38544. doi:10.1371/journal.pone.0038544

Chapter 7
Uptake and Reaction of *C. elegans* to Environmental RNAs

Ahmed Waqas and Ge Shan

Abstract Relation between animals and bacteria is just like marriage made by evolution. For *Caenorhabditis elegans*, bacteria are part of their food as well as the environments. In recent decades, *C. elegans* has been extensively used as a model organism for studying host–microbiota interactions. As dsRNAs expressed in *E. coli*, which is the laboratory food of *C. elegans*, can trigger RNAi in the worm, it is possible that bacteria could use their own RNAs to affect the physiology of *C. elegans*. The recent studies provide insights into these speculations. Our lab have demonstrated that some of the noncoding RNAs of *E. coli* modulate gene expression and regulate the life span and sensory response of *C. elegans*. These noncoding RNAs may be processed into tiny RNAs of ~17–18 nucleotides (nt) and utilize a distinct although partially overlapping pathway with the canonical feeding RNAi, which is triggered by exogenous double-stranded RNAs. There is a need for more studies and discussions about the interaction between *C. elegans* and RNAs in the environments, as a lot of details and mechanisms are still unknown.

7.1 Introduction

The discovery of RNA interference (RNAi) phenomenon dates back to late 1980s and 1990s, and researchers were trying to find the trigger of the RNAi (Meng et al. 2013). It was found that gene specific double-stranded RNA (dsRNA) is the most effective molecule to trigger RNAi in 1998 (Fire et al. 1991, 1998; Fischer 2010). In the 1998 research by Fire et al., potent RNAi was induced in a lot of cells by injecting the dsRNA into the body of *C. elegans* (Fire et al. 1998). This finding not only demonstrates the key roles of dsRNA in the RNAi but also points out that there is a pathway for uptaking and spreading exogenous RNAs in *C. elegans*. Alternative administration methods of RNAi in *C. elegans* by feeding or soaking

A. Waqas · G. Shan (✉)
School of Life Sciences and CAS Key Laboratory of Innate Immunity and Chronic Disease, University of Science and Technology of China, Hefei 230027, Anhui, China
e-mail: shange@ustc.edu.cn

A.L. Leitão and F.J. Enguita (eds.), *Non-coding RNAs and Inter-kingdom Communication*, DOI 10.1007/978-3-319-39496-1_7

were soon established (Liu et al. 2012; Meng et al. 2013). In feeding RNAi, engineered *E. coli* bacteria expressing gene specific dsRNA are given as food to the animal whereas in soaking RNAi worms are soaked in solution of dsRNA (Tabara et al. 1998; Timmons and Fire 1998). RNAi by injection, feeding, and soaking highlighted that environmental dsRNA (artificially made though) could be uptaken by intestine cells, and then the exogenous dsRNA or a processed RNAi molecule could be transported from cell to cell in *C. elegans* (Whangbo and Hunter 2008).

7.2 Environmental RNAi Pathway of *C. elegans*

The basic theme of RNAi is destabilizing RNA transcripts and/or inhibiting translation of mRNA by its interaction with complementary base-pairing dsRNA (Kong et al. 2007). This process is post transcriptional and induced by dsRNA. With more than 10 years of research, we have substantial knowledge about the environmental RNAi triggered by artificially expressed dsRNA. The very first step is the uptaking of exogenous RNAs from environment to the animal body or cells. In *C. elegans*, several genes have been identified by forward genetic approaches, which are crucial for this phenomenon. Out of multiple mentioned genes, *sid*-1 and *sid*-2 are essential for the uptake of exogenous RNAs to the intestinal cells (Saleh et al. 2006; Winston et al. 2002, 2007). The mutations in either gene abolished RNAi through feeding (Whangbo and Hunter 2008). Although *sid*-1 mutants showed no response to RNAi by injection, soaking or feeding, *sid*-2 mutants were sensitive to RNAi by injection yet were insensitive to feeding or soaking methods (Feinberg and Hunter 2003; Winston et al. 2002). The expression of *sid*-1 is mainly seen in non-neuronal cells while *sid*-2 is specifically expressed in intestinal cells (Feinberg and Hunter 2003; Jose et al. 2009).

Structure and function Studies of SID-1 and SID-2 demonstrate that SID-1 function as channel or component of the channel for uptake of dsRNAs (Feinberg and Hunter 2003; McEwan et al. 2012). On the other hand, SID-2 is assumed to be receptor or a component of the receptor, which helps SID-1 in uptaking of dsRNAs to the intestinal lumen (Feinberg and Hunter 2003; McEwan et al. 2012). These studies show that *sid*-2 is only needed for the uptake of dsRNA but not for spreading of RNAi from intestine to other cells, but *sid*-1 is playing dual role as it is involved in import of dsRNA as well as spreading of RNAi signal (McEwan et al. 2012). Genes involved in receptor-mediated endocytosis like *arl*-1, *vps*-34, and *vps*-41 are also linked with environmental RNAi both in *C. elegans* and *Drosophila* (Saleh et al. 2006). Similarly, researchers identified the roles of *sid*-3 and *sid*-5 in environmental RNAi (Hinas et al. 2012; Jose et al. 2012). A study shows that *sid*-3 plays a role in the import of silencing signal into cells but does not export it out of cells (Jose et al. 2012). A worm specific protein of 67 amino acids is encoded by *sid*-5 with a single transmembrane domain and is found to be colocalized with RAB-7 and LMP-1, which are late endosomal proteins. SID-5 function is predicted in transportation of silencing signal out of the intestine (Hinas et al. 2012; Meng

et al. 2013). Currently, it is not clear whether these proteins are involved in the same transporting pathway or alternative pathways, or there may be different cargos being transported by these different proteins.

Once inside the animal cells, dsRNA is processed by *dcr*-1, which encodes a bidentate ribonuclease to cut dsRNAs into small interference RNA (siRNA). RNAi defective gene 1 (*rde*-1), *alg*-1, and *alg*-2 are among the *C. elegans* Argonaute genes required for the feeding RNAi and the microRNA pathway. Another *rde* genes, *rde*-2, a gene specific for *Caenorhabditis*, is required for RNAi and transposon silencing. *Rde*-4, which encodes a *Caenorhabditis*-specific dsRNA-binding protein is essential for the initiation RNAi effect. The effect of feeding RNAi is also modified by *Caenorhabditis* specific RNA-dependent RNA polymerase genes such as *rrf*-3 (Fischer 2010).

7.3 Endogenous ncRNAs in *C. elegans* Bacterial Food

Essentially all the research above was performed with recombinant *E. coli*-expressing dsRNA. Some small regulatory RNAs are induced in bacteria such as *E. coli* under stress conditions. More than 100 small RNAs have been identified in bacteria so far (Meng et al. 2013). The major functional mechanism of bacterial ncRNAs is antisense, and by base pairing with target mRNAs, they may result in mRNA translation activation or inhibition, mRNA stabilization or degradation in harsh environments such as oxidative stress and low temperature (Gottesman and Storz 2011; Lalaouna et al. 2013).

Bacteria under stress may utilize all means for the survival of the species. We reasoned that bacterial RNAs, specifically in our previous research, the ncRNAs in stressed bacteria might possess a secondary function to interfere with the physiology of *C. elegans* and some other bacterial foragers, considering that artificially engineered dsRNAs can induce feeding RNAi in *C. elegans*.

7.4 Roles of DsrA and OxyS in *C. elegans*

In 2012, Liu et al. demonstrated the role of the two ncRNAs of *E. coli* in regulating gene expression? Or behavior? in *C. elegans* (Liu et al. 2012). These two ncRNAs are produced under stress conditions and affect physiology as well as life span of the worm, respectively. DsrA is an *E. coli* ncRNA highly induced at low temperature of 25 °C or below. DsrA decreased the life span of *C. elegans* by inhibiting the expression of a worm gene, diacylglycerol lipase gene *F42G9.6*, which is involved in longevity (Liu et al. 2012). OxyS is an *E. coli* ncRNA produced under oxidative stress. This ncRNA affected the behavior of the worms by down regulating the expression of a chemosensory gene, *che*-2 (Meng et al. 2013).

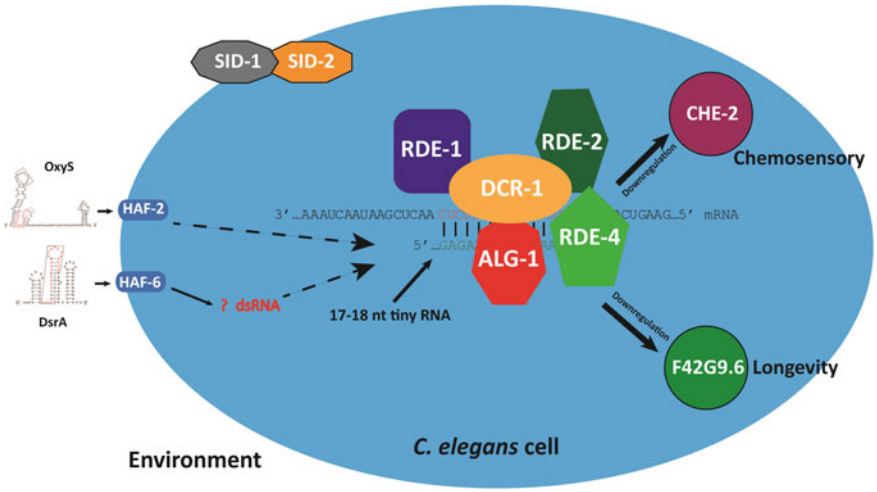


Fig. 7.1 The working model for the entry and roles of OxyS and DsrA in *C. elegans*. The entry of OxyS and DsrA are SID-1 and SID-2 independent, and they are transported into *C. elegans* cells by the ABC transporter HAF-2 and HAF-6, respectively. Presumably, DsrA is converted into dsRNAs to generate the antisense sequences. Tiny RNAs of ~17–18 nt originated from OxyS or DsrA are antisense to their target mRNA of *che*-2 or *F42G9.6* to inhibit the expression. CHE-2 and F42G9.6 are genes involved in the *C. elegans* chemosensory and longevity respectively. Sequences and proteins known to be essential for the effect of OxyS on *che*-2 expression are included, and for the effects of DsrA on *F42G9.6*, only RDE-4 is currently known to be essential

OxyS harbors a 17 nucleotide (nt) complementary sequence with *C. elegans che*-2 mRNA, and this sequence is distinct from the functional sequences of OxyS for base-pairing to target transcripts in *E. coli*. Down regulation of *che*-2 requires the 17 nt complementary sequence in OxyS, and is carried out with the involvement of *alg*-1 and *rde*-4. An ABC transporter gene, *haf*-2 was found to be another key player in the effect of OxyS (Liu et al. 2012). DsrA shares a 27 nt sequence (with 4 nt mismatch) with *F42G9.6* mRNA. Small RNAs antisense to this 27 nt sequence were identified in the RNA sequencing data available from databases, indicating that an unknown RNA-dependent RNA polymerase might convert DsrA into dsRNA (Liu et al. 2012; Meng et al. 2013). *rde*-4 and an ABC transporter gene *haf*-6 play vital roles for the targeting of F42G9.6 by DsrA. It was surprising to see that *sid*-1 and *sid*-2 were not necessary for the effects of both OxyS and DsrA (Fig. 7.1). This research further demonstrated that the effects of OxyS and DsrA in *C. elegans* might be mediated by small tiny RNAs of ~17–18 nt complementary to the mRNA target (Liu et al. 2012). When injected into *C. elegans*, these tiny RNAs could inhibit the expression of the target gene by lowing the levels of the corresponding mRNA.

It seems that the endogenous ncRNAs of *E. coli* regulate the gene expression of *C. elegans* in ways distinct from the canonical feeding RNAi pathway, which is

initiated by dsRNAs. First, their entry into the animal is not SID-1 or SID-2 dependent, but rather relies on ABC transporters. Secondly, the effects of OxyS and DsrA may be mediated by tiny RNAs of ~17–18 nt, not the typical siRNAs of ~20 nt. Thirdly, although genes involved in dsRNA feeding RNAi such as *rde*-4 are required, a lot of genes in the canonical feeding RNAi pathway are not essential for the effects of OxyS and DsrA.

Several studies showed that *sid*-1 and *sid*-2 may not be essential even for environmental RNAi triggered by dsRNA (Dalzell et al. 2011; Ghedin et al. 2007). The parasitic filarial nematode *Brugia malayi* is competent for environmental RNAi, although there is no *sid-1* and *sid*-2 homologs in its genome (Ghedin et al. 2007). Survey for orthologs of *C. elegans* RNAi pathway components in multiple nematode species revealed that most of the *C. elegans* genes for uptaking and spreading of the exogenous dsRNA are absent in these species, including those nematodes competent for environmental RNAi (Dalzell et al. 2011). It is possible that either these genes might be fast evolving or animals might have evolved alteRNAtive pathways for the uptaking of environmental dsRNA.

In 2015, Akay and colleagues published a research report examining effects of *E. coli* OxyS in *C. elegans* (Akay et al. 2015). Although not noticed by these authors, data from this research showed a switch for small RNAs in *C. elegans* with sequences corresponding to OxyS from shorter ($\leq$ 21 nt) to longer ($\geq$ 22 nt) RNAs between *wildtype* and *rde*-4 worms. This finding is in consistent with the prvious indications that ~17–18 nt tiny RNAs are the mediator of OxyS effect in *C. elegans*. Although they did not show reads of the 17 nt RNAs, levels of 18 nt small RNAs decreased significantly in *rde*-4 mutant worms. When examining *che*-2 mRNA, the authors showed a degree of variation but not statistically different levels in adult worms fed with OxyS expressing bacteria, compared to worms fed with bacteria without OxyS RNA (Akay et al. 2015). Besides the difference in the bacteria strains used, there is also difference in the stage of worms used in this research as contrast to research of Liu and colleagues (Liu et al. 2012). Younger L3 worms were examined in the 2012 paper, and it is possible that the suppression in *che*-2 mRNA levels by OxyS is not significant in older (L4 and adult) worms. These two researches actually agree in that OxyS does not trigger the canonical feeding RNAi pathway in *C. elegans*.

7.5 Natural Food and Environments of *C. elegans*

When reviewing the uptake and reaction of *C. elegans* to RNAs in their food and environments, we have to discuss what the natural food and environments are for this animal. It is surprising that we have very limited information about the real life of *C. elegans* in wild.

Although *C. elegans* is mostly raised on *E. coli* in laboratories, these bacteria are not a major *C. elegans* food source in nature (Kiontke and Sudhaus 2006). *Caenorhabditis* species including *C. elegans* may be associated with other

invertebrates or vertebrates for transportation, and may grow and reproduce on microorganism-rich organic material. In addition, in a confined laboratory space, mice could spread *C. elegans* and *E. coli* (e.g. *E. coli* inside mice) (Liu et al. 2012). Even *E. coli* may not be a major natural food of *C. elegans*, it remains possible that this bacterial species is on the menu of the nematode (Liu et al. 2012). Bacteria maybe found in the living environments of *C. elegans* such as *Bacillus mycoides* may also utilize their ncRNAs to interfere with gene expression of *C. elegans* (Liu et al. 2012).

One would expect that *Caenorhabditis* species and their bacterial food would endure cycles of "good" and "bad" time, and they have to "consider" each other in their life and evolutionary history. We speculate that ncRNAs in the food and environments of *C. elegans* could play roles possibly through a noncanonical feeding RNAi pathway in the co-evolution of worms and bacteria. It has been shown that OxyS and DsrA have protective effects on *E. coli* under the foraging pressure of *C. elegans* in a laboratory setup (Liu et al. 2012).

7.6 Conclusions and Future Perspectives

ncRNAs play vital roles in many life events in both eukaryotes and prokaryotes (Castel and Martienssen 2013; Hu et al. 2012; Lee 2012; Li et al. 2015; Ulitsky and Bartel 2013). The phenomenon of environmental RNAi demonstrates an interspecies role of ncRNA. Environmental RNAi pathway initiated by bacterial ncRNAs such as OxyS and DsrA is partially overlapping although distinct from the previously characterized feeding RNAi pathway triggered by exogenous dsRNA in *C. elegans*. The interspecies effects of OxyS and DsrA, secondary to its primary regulatory roles in bacterial stress responses, may have been formed as secondary adaptation during the course of co-evolution between bacteria (for example, *E. coli* and *B. mycoides*) and bacteria forager (for example, *C. elegans*). In *C. elegans,* small RNA pathways are very complex indicated by the expansion of Argonaute gene family in this animal. More target oriented studies will be helpful to understand the molecular pathway underneath the environmental RNAi triggered by bacterial ncRNAs.

A lot of unknowns are present for the effects of OxyS and DsrA in *C. elegans*. Whether and how ABC transporters play roles in environmental RNAi are still elusive. How OxyS and DsrA are processed in *C. elegans*, and how the resulting small RNAs modulate the gene expression are unclear. Presumably, *C. elegans* has to coordinate the effects of environmental RNAi and the effects initiated by their own small RNAs, and we know little about this. These intriguing questions demand vigorous research. DsRNAs, sRNAs such as OxyS and DsrA, or some other species of RNAs may go through different pathways.

Nevertheless, two *E. coli* endogenous ncRNAs, OxyS and DsrA, altered the physiology of *C. elegans* through regulating gene expression of the animal. Both ncRNAs are expressed in *E. coli* under stress conditions, and may protect *E. coli*

from overforaging by *C. elegans*. It remains to be seen how many other ncRNAs in the environments of *C. elegans* elicit this kind of effects in the life or even the evolutionary history of *C. elegans* in nature.

Acknowledgments We thank Bin Yu and other members of the Ge Shan lab for discussions. This work was supported by the National Basic Research Program of China (2015CB943000), National Natural Science Foundation of China (31471225 and 91519333), and the Fundamental Research Funds for the Central Universities (WK2070000034).

References

Akay A, Sarkies P, Miska EA (2015) E. coli OxyS non-coding RNA does not trigger RNAi in *C. elegans*. Sci Rep 5:9597. doi:10.1038/srep09597

Castel SE, Martienssen RA (2013) RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. Nat Rev Genet 14(2):100–112. doi:10.1038/nrg3355

Dalzell JJ, McVeigh P, Warnock ND, Mitreva M, Bird DM, Abad P, Fleming CC, Day TA, Mousley A, Marks NJ, Maule AG (2011) RNAi effector diversity in nematodes. PLoS Negl Trop Dis 5(6):e1176. doi:10.1371/journal.pntd.0001176

Feinberg EH, Hunter CP (2003) Transport of dsRNA into cells by the transmembrane protein SID-1. Science 301(5639):1545–1547. doi:10.1126/science.1087117

Fire A, Albertson D, Harrison SW, Moerman DG (1991) Production of antisense RNA leads to effective and specific inhibition of gene expression in *C. elegans* muscle. Development 113 (2):503–514

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature 391 (6669):806–811. doi:10.1038/35888

Fischer SE (2010) Small RNA-mediated gene silencing pathways in C. elegans. Int J Biochem Cell Biol 42(8):1306–1315. doi:10.1016/j.biocel.2010.03.006

Ghedin E, Wang S, Spiro D, Caler E, Zhao Q, Crabtree J, Allen JE, Delcher AL, Guiliano DB, Miranda-Saavedra D, Angiuoli SV, Creasy T, Amedeo P, Haas B, El-Sayed NM, Wortman JR, Feldblyum T, Tallon L, Schatz M, Shumway M, Koo H, Salzberg SL, Schobel S, Pertea M, Pop M, White O, Barton GJ, Carlow CK, Crawford MJ, Daub J, Dimmic MW, Estes CF, Foster JM, Ganatra M, Gregory WF, Johnson NM, Jin J, Komuniecki R, Korf I, Kumar S, Laney S, Li BW, Li W, Lindblom TH, Lustigman S, Ma D, Maina CV, Martin DM, McCarter JP, McReynolds L, Mitreva M, Nutman TB, Parkinson J, Peregrin-Alvarez JM, Poole C, Ren Q, Saunders L, Sluder AE, Smith K, Stanke M, Unnasch TR, Ware J, Wei AD, Weil G, Williams DJ, Zhang Y, Williams SA, Fraser-Liggett C, Slatko B, Blaxter ML, Scott AL (2007) Draft genome of the filarial nematode parasite *Brugia malayi*. Science 317 (5845):1756–1760. doi:10.1126/science.1145406

Gottesman S, Storz G (2011) Bacterial small RNA regulators: versatile roles and rapidly evolving variations. Cold Spring Harb Perspect Biol 3(12). doi:10.1101/cshperspect.a003798

Hinas A, Wright AJ, Hunter CP (2012) SID-5 is an endosome-associated protein required for efficient systemic RNAi in *C. elegans*. Curr Biol 22(20):1938–1943. doi:10.1016/j.cub.2012.08.020

Hu S, Wu J, Chen L, Shan G (2012) Signals from noncoding RNAs: unconventional roles for conventional pol III transcripts. Int J Biochem Cell Biol 44(11):1847–1851. doi:10.1016/j.biocel.2012.07.013

Jose AM, Smith JJ, Hunter CP (2009) Export of RNA silencing from *C. elegans* tissues does not require the RNA channel SID-1. Proc Natl Acad Sci USA 106(7):2283–2288. doi:10.1073/pnas.0809760106

Jose AM, Kim YA, Leal-Ekman S, Hunter CP (2012) Conserved tyrosine kinase promotes the import of silencing RNA into *Caenorhabditis elegans* cells. Proc Natl Acad Sci USA 109 (36):14520–14525. doi:10.1073/pnas.1201153109

Kiontke K, Sudhaus W (2006) Ecology of *Caenorhabditis* species. WormBook:1–14. doi:10.1895/wormbook.1.37.1

Kong Y, Ruan L, Ma L, Cui Y, Wang JM, Le Y (2007) RNA interference as a novel and powerful tool in immunopharmacological research. Int Immunopharmacol 7(4):417–426. doi:10.1016/j.intimp.2006.12.011

Lalaouna D, Simoneau-Roy M, Lafontaine D, Masse E (2013) Regulatory RNAs and target mRNA decay in prokaryotes. Biochim Biophys Acta 1829(6–7):742–747. doi:10.1016/j.bbagrm.2013.02.013

Lee JT (2012) Epigenetic regulation by long noncoding RNAs. Science 338(6113):1435–1439. doi:10.1126/science.1231776

Li Z, Huang C, Bao C, Chen L, Lin M, Wang X, Zhong G, Yu B, Hu W, Dai L, Zhu P, Chang Z, Wu Q, Zhao Y, Jia Y, Xu P, Liu H, Shan G (2015) Exon-intron circular RNAs regulate transcription in the nucleus. Nat Struct Mol Biol 22(3):256–264. doi:10.1038/nsmb.2959

Liu H, Wang X, Wang HD, Wu J, Ren J, Meng L, Wu Q, Dong H, Kao TY, Ge Q, Wu ZX, Yuh CH, Shan G (2012) Escherichia coli noncoding RNAs can affect gene expression and physiology of *Caenorhabditis elegans*. Nat Commun 3:1073. doi:10.1038/ncomms2071

McEwan DL, Weisman AS, Hunter CP (2012) Uptake of extracellular double-stranded RNA by SID-2. Mol Cell 47(5):746–754. doi:10.1016/j.molcel.2012.07.014

Meng LF, Chen L, Li ZY, Wu ZX, Shan G (2013) Environmental RNA interference in animals. Chin Sci Bull 35:4418–4425

Saleh MC, van Rij RP, Hekele A, Gillis A, Foley E, O'Farrell PH, Andino R (2006) The endocytic pathway mediates cell entry of dsRNA to induce RNAi silencing. Nat Cell Biol 8(8):793–802. doi:10.1038/ncb1439

Tabara H, Grishok A, Mello CC (1998) RNAi in C. elegans: soaking in the genome sequence. Science 282(5388):430–431

Timmons L, Fire A (1998) Specific interference by ingested dsRNA. Nature 395(6705):854. doi:10.1038/27579

Ulitsky I, Bartel DP (2013) lincRNAs: genomics, evolution, and mechanisms. Cell 154(1):26–46. doi:10.1016/j.cell.2013.06.020

Whangbo JS, Hunter CP (2008) Environmental RNA interference. Trends Genet 24(6):297–305. doi:10.1016/j.tig.2008.03.007

Winston WM, Molodowitch C, Hunter CP (2002) Systemic RNAi in *C. elegans* requires the putative transmembrane protein SID-1. Science 295(5564):2456–2459. doi:10.1126/science.1068836

Winston WM, Sutherlin M, Wright AJ, Feinberg EH, Hunter CP (2007) *Caenorhabditis elegans* SID-2 is required for environmental RNA interference. Proc Natl Acad Sci USA 104(25): 10565–10570. doi:10.1073/pnas.0611282104

Part II
Non-coding RNAs and Interactions Between Eukaryotic Organisms

Chapter 8
Rapid Evolution of Mosquito Anti-viral ncRNA Pathway Components

Corey L. Campbell, Laura B. Dickson and William C. Black IV

Abstract Vector-host non-coding RNA (ncRNA) interactions are important for the maintenance of arthropod-borne virus (arbovirus) infection cycles in nature. A major anti-viral pathway in mosquitoes is the RNA interference (RNAi) pathway. Using high throughput sequencing (HTS) data, population genetics analysis was performed on major RNAi components from natural populations of the dengue and zika virus vector, *Aedes aegypti*. Pairwise comparisons of four geographically separated wild-caught collections (Senegal, Thailand, and Mexico) were analyzed for this study. Mutation rate ratios of siRNA/miRNA orthologs indicated rapid evolution of anti-viral siRNA pathway components, as had been reported previously. Polymorphisms were present in key amino acid residues of RNAi pathway components, which could contribute to variability in arbovirus infection rates, and thereby influence transmission cycles. Haplotype analysis of Argonaute-2 (Ago2), the RNAi slicer enzyme, revealed a hypervariable region in the predicted N-terminus that varied widely among the populations, as well as polymorphisms in the PAZ and PIWI domains. Dicer-2 showed selected coding changes near the DEAD/DEAH helicase, dimerization, and RNase III domains, which also varied among the collections. Senegal (PK10) showed the highest number of coding changes in Ago2 and Dicer-2. Rapid evolution of RNAi components may influence emergence of arbovirus genotypes and contribute to the inter-kingdom arms race of arbovirus pathogens and their hosts.

8.1 Introduction

Arthropod-borne viruses (arboviruses) require alternating replication cycles in vector arthropods and vertebrate hosts to maintain their transmission cycles in nature. Vector competence is the extent to which arboviruses infect and are

C.L. Campbell (✉) · L.B. Dickson · W.C. Black IV
Department of Microbiology, Immunology and Pathology,
Colorado State University, Fort Collins, CO 80523, USA
e-mail: corey.campbell@colostate.edu

A.L. Leitão and F.J. Enguita (eds.), *Non-coding RNAs and Inter-kingdom Communication*, DOI 10.1007/978-3-319-39496-1_8

successfully transmitted from the vector insect to a new human host. Arboviruses evade the vector's immune response to establish a productive infection in the insect midgut, disseminate to peripheral tissues and finally release into saliva for subsequent transmission to new hosts. The required incubation period usually spans 10–14 days, but can be longer or shorter, depending on a variety of factors, including the vector's immune status (Lambrechts et al. 2011). Arbovirus infection of vectors is in some ways similar to that of animal reservoirs. As with reservoirs, vectors often do not show pathological signs of infection (reviewed in Bean et al. 2013). Importantly, once infected, both animal reservoirs and vectors remain infected for life (Schountz et al. 2012). The major anti-viral immune pathway in vector insects is an ncRNA regulatory pathway (Hess et al. 2011; Sanchez-Vargas et al. 2009). Because most arboviruses have RNA genomes (reviewed in Rosenberg et al. 2013), there is a triad relationship, and subsequent inter-kingdom communication, that occurs between RNA viruses and their vector and vertebrate hosts. Although inter-kingdom ncRNA communication can be initiated from viruses as well, this chapter focuses on vector-host ncRNA interactions.

Dengue, or breakbone fever, the most common human arbovirus disease, is caused by one of four dengue virus (DENV, family *Flaviviridae*, genus *Flavivirus*) serotypes (Bhatt et al. 2013; Gubler 2002; Nguyet et al. 2013). Mosquitoes' (*Ae. aegypti*) vector competence for DENV serotype 2 (DENV2) has evolved over time and, in hyper-endemic regions, no longer requires an animal reservoir (Weaver and Reisen 2010). Nevertheless, mosquitoes mount a multi-pronged response to virus infection, including the ncRNA regulatory pathway, regardless of whether they are resistant or susceptible to a given virus strain. In addition, there is growing evidence for gene-for-gene adaptive evolution of ncRNA anti-viral defense components and DENV genotypes in nature (Lambrechts et al. 2009, 2013; Moncayo et al. 2004).

A key requirement of the mosquito-virus relationship is the balance between successful virus replication and simultaneous avoidance of pathological effects that would diminish vector fitness levels. If such control were not attained, mosquito virus infections would be more similar in pathology to infections that occur in model organisms, such as drosophilids (van Rij et al. 2006). Pathogenic virus infection, rather than arbovirus infection, would be expected to drive anti-viral pathway evolution in mosquitoes. However, although arbovirus infection would not necessarily be expected to be a genetic driver of anti-viral defense evolution in mosquitoes, evidence is building for strong associations between these features (Lambrechts et al. 2009, 2013). Indeed, evolutionary selection of anti-viral defense genes may concomitantly affect arbovirus-mosquito interactions in a vector population-specific manner. Any effects on arbovirus evolution may be an incidental result of these evolutionary factors in mosquitoes.

Variability in the competence of vector mosquitoes for arbovirus infections is further complicated by a high degree of genetic variability within a given mosquito species. Although *Ae. aegypti* sensu lato is described as a single polytypic species, there is a high degree of genetic variability worldwide (Brown et al. 2011; Moore 1979). Recent support for a complex genetic structure was found in experiments showing reproductive incompatibilities between *Ae. aegypti aegypti*, the type form,

and west African populations (Dickson et al. 2016a), which provide support for cryptic subspecies in nature. This is important, because it is consistent with the hypothesis that vector competence determinants are also be quite variable in wild mosquito populations, which would facilitate emergence of new arboviruses, as has occurred with zika and chikungunya viruses (Roth et al. 2014).

There are three major small ncRNA regulatory pathways in multi-cellular organisms. Vector mosquitoes have all major pathway components that have been identified in model organisms (Campbell et al. 2008b). Each of these three pathways uses small ncRNAs of fewer than 50 nucleotides in length to control gene expression in a transcriptional or post-transcriptional manner. In this chapter, we discuss our current understanding of the anti-viral ncRNA pathway in the dengue vector, *Ae. aegypti*, and tie in new information about the diversity of major pathway components in natural mosquito populations. We also summarize genetic association studies of natural variability in vector competence levels in natural populations (Dickson et al. 2014; Sylla et al. 2009). HTS data was evaluated to confirm previously reported polymorphisms in small RNA regulatory pathway (SRRP) components, which contribute to rapid evolution of the anti-viral ncRNA pathway (Bernhardt et al. 2012), and determine whether these trends held true in wild-caught populations. Population-specific genetic associations in other anti-viral pathway components were also explored. We compared those differences among populations expected to have variable DENV2 transmission rates, in order to identify evolutionary trends in mosquitoes that may affect arbovirus transmission.

8.2 RNA Interference (RNAi): The Mosquito's Major Anti-viral Pathway

The mosquito's anti-viral ncRNA response exploits an SRRP specific for small interfering RNAs (siRNAs, ~19–23 nts) and stimulates cleavage of viral RNA genomes using an Ago2/Dicer-2 dependent RNAi (siRNAi) mechanism (Sanchez-Vargas et al. 2009). Another anti-viral avenue, which will not be addressed here, relies on the general innate immune response (Xi et al. 2008). Neither of these pathways, however, is sufficient to clear infection in susceptible mosquitoes. The RNA-induced silencing complex (RISC) contains components necessary and sufficient for cleavage of viral genomes (Lee et al. 2004; Pham and Sontheimer 2005; van Rij et al. 2006). *Ae. aegypti* has high molecular weight complexes, consistent with those of the RISC, which co-migrate with Ago2, the slicer catalytic enzyme (Hess et al. 2011). Within the RISC, Ago2 directly interacts with Dicer-2, which has a dsRNA binding domain that brings siRNAs in direct complementarity with target viral sequences.

Scientists have speculated that arboviruses evade or sequester their genomes away from siRNAi defense mechanisms in competent, or susceptible, vectors, while in resistant mosquitoes, all arbovirus genomes are successfully targeted by the

siRNA machinery and cleared, possibly prior to escape from the primary infection site, which is the midgut (Bennett et al. 2002; Bosio et al. 2000). In mosquitoes, injection of non-specific dsRNA as a negative control in gene-knockdown experiments results in mild stimulation of the RNAi pathway, which slightly diminishes viral replication rates (Sanchez-Vargas et al. 2009 and Campbell, unpublished). This is consistent with the hypothesis that stimulation of the immune response enhances anti-viral siRNAi activity to promote viral clearance. In contrast, sequestration of viral genomes to evade siRNAi could be due to membrane rearrangements commonly found in cell types supporting viral replication (Franco et al. 2010). This phenomenon has also been described for other flaviviruses (Ambrose and Mackenzie 2015; Hirano et al. 2014). Other possible viral immune evasion mechanisms include RNAi suppressor activity, which have been identified for a DENV nonstructural protein, ns4B, and for subgenomic flavivirus RNAs (Kakumani et al. 2013; Schnettler et al. 2012).

8.3 The PIWI RNA Pathway: An Alternate Anti-viral Response Pathway?

In dipterans, PIWI pathway ncRNAs (piRNAs, 24–30 nts) are best known for their control of transposon activation during germline development (Saito et al. 2006; Vagin et al. 2006). The first hint of a possible role for piRNA activity in the mosquito anti-viral response was observed in the malaria vector, *Anopheles gambiae*. The study showed that, in addition to *ago2* and *dicer-2,* the PIWI pathway component transcript *ago3* was enriched during arbovirus infection (Keene et al. 2004), consistent with a role in the host-virus response. Evidence of a possible anti-viral role in *Ae. aegypti* was first observed in a small RNA (sRNA) profiling study of mosquitoes exposed to DENV2 (Hess et al. 2011). We observed sRNAs (24–30 nts) of viral origin, which were of a size consistent with cleavage by piRNA pathway components. Virus-derived piRNAs have unique structural features in some virus-insect systems; these include a U at the 5′ end of the antisense strand and an A at position 10 of the sense piRNA strand. Although seemingly typical of flavivirus piRNAs in mosquito cell culture, these signature features are not always present in virus-derived piRNAs isolated from vector mosquitoes (Hess et al. 2011; Leger et al. 2013; Schirtzinger et al. 2015), however the reasons for this are unclear. In addition to the presumed involvement of *ago3* in the mosquito virus infection response, other studies have implicated *piwi4* and *piwi5* in viral piRNA biogenesis in alphavirus-infected *Ae. aegypti* cell culture (Miesen et al. 2015; Schnettler et al. 2013). However, little is understood about the precise mechanisms and implications of these associations with vector competence. Until more mechanistic insight is attained, we speculate that the PIWI SRRP acts in a compensatory manner during periods of viral replication in which canonical siRNAi anti-viral RISC binding sites are saturated.

8.4 Mosquito miRNA Pathway Host Responses

In dipterans, a third RNAi pathway uses miRNAs (~19–23 nts) (miRNAi) for post-transcriptional RISC-mediated regulation of housekeeping and developmental gene expression in an Argonaute-1/Dicer-1 dependent manner. miRNAs control gene expression of at least 15 % of the *Drosophila* genome (Grun et al. 2005). Because this pathway has a strong role in housekeeping regulation and therefore would be expected to be evolutionarily conserved, it was used as a negative control in analyses to assess evolutionary rates for anti-viral defense pathways (Bernhardt et al. 2012; Obbard et al. 2006). Using Sanger sequencing, we found evidence for diversifying selection of certain codons in *dicer-1* genes (Bernhardt et al. 2012). Moreover, a subsequent study showed that miRNA levels were significantly modulated in mosquitoes exposed to DENV2 (Campbell et al. 2014). However, it is not yet known whether subsequent miRNA target gene expression changes would benefit, mitigate or merely serve as a secondary response to virus replication effects on host cell metabolism.

8.5 Anti-viral Pathway Evolution Occurs in a Population-Specific Manner

Using traditional genetic mapping methods, we found that about 40 % of *Ae. aegypti* vector competence for DENV2 was associated with genomic quantitative trait loci (Bosio et al. 2000). These traditional methods may have missed focal amino acid replacements that conferred phenotypic changes. After components in the siRNAi pathway were identified as major players in anti-viral immunity (Campbell et al. 2008a; Keene et al. 2004; Sanchez-Vargas et al. 2009), we were able to refine our approach to the study of specific gene-for-gene interactions that influence vector competence. Sanger sequencing of all the major siRNA and miRNA pathway genes from colony mosquitoes, originally collected from locations similar to those used in the current study, indicated that replacement substitution rates, indicative of diversifying selection and evolutionary pressure, were significantly higher in siRNA pathway genes than in miRNA pathway genes, which was consistent with findings from a similar study of drosophilids (Bernhardt et al. 2012; Obbard et al. 2006). This finding provided an important precedent for study of variability in vector competence and evolution of antiviral ncRNA pathway genes.

8.6 Population Genomics Analysis Shows Evolution of piRNA and siRNA Pathway Components

DENV hyper-endemicity infection cycles often are initiated in urban areas due to high human population densities and coincident habitation of humans and competent mosquitoes (Cummings et al. 2004) and are indicative of simultaneous circulation of multiple DENV serotypes. In general, populations represented by the type form, *Ae. aegypti*, which prefer peridomestic container breeding sites, and preferentially feed on humans, have higher vector competence rates than sylvatic forms that lay their eggs in tree holes and prefer to feed on animals (Mattingly 1958; McBride et al. 2014). Lab studies of vector competence of Senegalese *Ae. aegypti* provided additional support for diversity in the association between mosquito collection site and DENV2 vector competence levels, and the given virus strain also influenced specific infection outcomes (Dickson et al. 2014; Sylla et al. 2009).

In our effort to understand the evolutionary basis for the relationship between anti-viral ncRNA pathways and DENV vector competence, our group has focused on population genetics comparisons of *Ae. aegypti* from Mexico, western Africa (Senegal) and Thailand. Human dengue cases are endemic in Mexico and hyper-endemic in Thailand, but less widely reported in western Africa (Cummings et al. 2004; Dantes et al. 2014); however, the lack of evidence for DENV transmission in western Africa is more likely due to under-reporting rather than a lack of cases, per se (Franco et al. 2010). Nevertheless, the high level of diversity among these reproductively isolated mosquito populations provided the basis for comparisons to lend insight into population-specific evolutionary selective pressures. The Mexico and Thai collections were chosen, because they represented populations of the type form, *Ae. aegypti* (Bennett et al. 2002) that are highly competent for DENV2 transmission. The Sengalese aedine populations were chosen, because they represent diverse ancient sylvatic forms with mixed vector competence levels (Diallo et al. 2008; Dickson et al. 2014; Sylla et al. 2009).

For the purposes of the current study, HTS exome-captured genomic DNA (gDNA) data from two high DENV2-competent collections (Mexico (Juan Pablo, Merida, Yucatan) and Thailand) of the *Ae. aegypti* type form were compared with sylvatic populations of mixed or reduced competence (Kaolack and PK10, Senegal, respectively) (Fig. 8.1). This strategy provided a useful measure of variable evolutionary pressures on mosquito anti-viral pathway components. Importantly, the advent of HTS and exome-capture technologies allowed us to address these paradigms in natural mosquito populations rather than having to rely on colonization prior to sampling. Individual adult mosquito gDNAs were pooled in stoichiometric amounts for each replicate (Fig. 8.1 legend). Dual replicate libraries for each collection were subjected to exome-capture, which preferentially enriched for whole-transcriptome sequences, using probes developed by Juneja and coworkers (Juneja et al. 2015). After sequencing, the reads were aligned to the aedine Liverpool strain transcriptome reference (genomebuild version Aaeg L1.3) and analyzed to identify

Fig. 8.1 Locations of mosquito collections used in this study. PK10, with F-type scale pattern, was collected from tree holes near Kedougou, Senegal; Kaolack, with mixed scale patterns, was collected from the village of Kaolack, Senegal. The Thai population was collected from Pai Lom, Thailand. The Mexico population was collected from Juan Pablo community of Merida, Mexico. Each biological replicate pool size was prepared from equal numbers of males and females, as follows: PK10 (24), Kaolack (14), Mexico (22), Thai (24), and two replicates were used for each pair-wise population comparison. Approximate library read counts per replicate were as follows: PK10, 38–39 M (million), Kaolack, 30–32 M, Mexico, 16–17 M, Thailand, ~24 M. See also (Dickson et al. 2016a)

single nucleotide polymorphisms (SNPs) between populations (Dickson et al. 2016b; Lawson et al. 2007), using a GSNAP/Varscan workflow (Koboldt et al. 2012; Wu and Nacu 2010). Calculations of −log10(pvalue) of chi-square contingency tables (LODs) were determined from Varscan outputs for each nucleotide in pair-wise population comparisons. From there, weighted average LOD values (adjLOD) were calculated per gene (Dickson et al. 2016a). We expected to observe increased diversifying selection of anti-viral pathway genes in comparisons of highly virus-competent populations (Thailand (Thai), and Mexico (Mex)) to the ancient genotypes represented by West African populations (PK10 and Kaolack (Kao), Senegal). Specific attributes of each SRRP are discussed in detail below, and all gene names and accession numbers for each pathway are identical to those indicated in Campbell and coworkers (2008a).

8.7 Antiviral siRNAi Pathway Evolution

Compared to the range of adjLOD values across the entire genome, siRNA pathway genes were not above the 95 % cI (Fig. 8.2a). This is likely because the important immune and regulatory functions of these pathway components results in some amount of purifying selection relative to all other genes in the genome. To test this idea, we examined the ratio of replacement to synonymous substitutions (Ka/Ks mutation rates). At the gene-wise level for all pair-wise population comparisons,

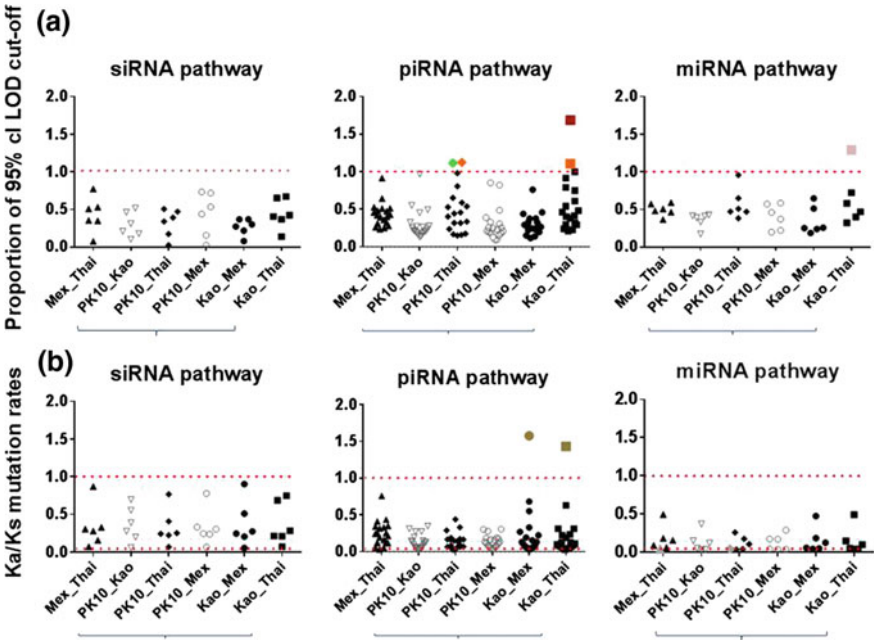


Fig. 8.2 Key ncRNA components are undergoing selection in wild *Ae. aegypti* populations. Gene-wise adjLOD values and mutation rates (Ka/Ks) are shown for all major pathway components for each pair-wise comparison indicated. Specific genes in each pathway were as previously reported (Campbell et al. 2008a). Only genes with significant differences are highlighted. **a** Plot values represent the relative percentage of the 95 % cI threshold cut-off for each gene in a given pathway. *Left panel*, canonical anti-viral defense siRNA pathway. *Center panel*, the germline and putative compensatory anti-viral regulatory piRNA pathway, PIWI3 (*red* symbol), PIWI5 (*bright green*) and two RNA helicase orthologs (*orange* symbols) show increased genetic association. *Right panel*, miRNA pathway component shows background association levels above the 95 % cI (Drosha, mauve label). Pairwise comparison genetic association values were calculated using contingency table Chi-square analysis of SNPs from HTS sequencing data, calculated using the probability with 3(n − 1) degrees of freedom where n was the number of nucleotides segregating at that site. This probability was converted to an LOD score as –logbase10 (prob). The 95 % confidence level (95 % cI) was calculated from transcriptome-wide LOD scores of all genes in a given comparison. Genes above the cut-off are considered significantly different. Sequencing details can be found in Dickson et al. (2016a). **b** Amino acid substitution rates (ω, Ka/Ks). *Left panel*, anti-viral defense siRNA pathway. *Center panel*, the piRNA pathway has several genes showing Ka/Ks rates >1, *armitage* (*olive* symbol) shows evidence of diversifying selection. All 6 comparisons showed *piwi1* with Ka/Ks rates <lower 5 % cI, indicating significant purifying selection. Similarly, *ago1a* was also <0.05 % cI for the PK10_Thai, Mex_Thai contrasts. *Right panel*, miRNA pathway components. *Blue brackets* indicate among-subspecies comparisons. Ka, replacement substitutions, Ks, synonymous substitutions; Ka/Ks rates indicate the mutation rates for each pair-wise comparison indicated. All raw fastq sequences are publicly available at NCBI sequence read archive under accession numbers PRJNA258086 and SRP061709

siRNA pathway genes showed amino acid substitution rates (Ka/Ks) < 1 (Fig. 8.2b), which were consistent with a lack of significant diversifying selection. Therefore we chose to assess the fold-change of siRNA pathway ortholog mutation rates relative to those of orthologous miRNA pathway components. In previous studies, this method provided insight into evolution of anti-viral pathway components in both drosophilids and colonized PK10 and Thai mosquitoes (Bernhardt et al. 2012; Obbard et al. 2006). Importantly, the current results were consistent with rapid evolution of anti-viral defense genes compared to orthologous non-immune genes for most pair-wise comparisons. At the per-gene level, *ago2*, the RISC catalytic enzyme and slicer of viral RNA genomes, showed the strongest evidence of rapid evolution, with *ago2/ago1* Ka/Ks ratio changes of 7.4 to 20-fold (Fig. 8.3). Ratios of *r2d2*, another key RISC component, and *r3d1* showed intermediate mutation rate ratios (range 4.7–8.2). Lastly, *dicer-2/dicer-1* ratios were lower (range 1.2–4.3), consistent with modest trends in rapid evolution, which was similar to our previous report (Bernhardt et al. 2012). The apparent lower rate change for *dicer-2* was due to higher *dicer-1* Ka/Ks values (range 0.12–0.18) compared to the miRNA orthologs *ago1* (range 0.03–0.06) or *r3d1* (range 0.04–0.06). In all comparisons, the within-subspecies comparison of Mexico and Thai collections yielded the lowest ratios, consistent with less evolutionary distance, even though the collections were widely separated geographically.

Selective pressure may have resulted in focal amino acid changes that were not reflected at the per-gene level shown above. Indeed, changes in protein-protein interactions of regulatory genes likely alter the overall complexion of immune responses. Therefore, we analyzed gene haplotypes produced from de novo assemblies of all library fastq files (Trinity package, version 2013-02-15) (Grabherr et al. 2011; Haas et al. 2013). Although the Trinity de novo assembly package was designed for RNA-Seq, it is suitable for assembling discrete coding regions to

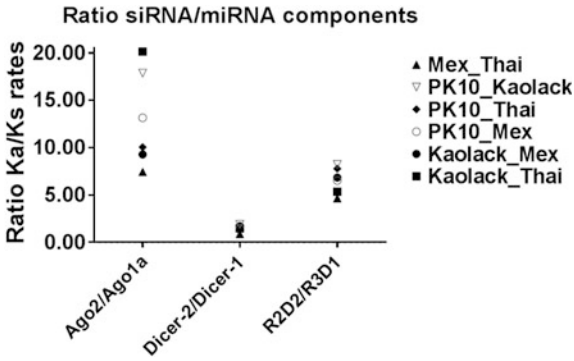


Fig. 8.3 Rapid evolution of siRNAi genes. Amino acid substitution (Ka/Ks) rates from Fig. 8.2 were graphed to display rate ratios for orthologous antiviral siRNAi pathway components compared to those of the miRNA pathway; *ago2/ago1*, *dicer-2/dicer-1*, *r2d2/r3d1*. The Mex_Thai comparison was a within *Ae. aegypti* subspecies comparison; the remaining comparisons are between putative subspecies

identify abundant polymorphisms in genomic DNA (B. Haas, personal communication). SNPs that were present in fewer than 5 % of all associated contigs would be excluded from the final output. We examined *ago2* and *dicer-2* haplotype fragments from raw fastq reads and found support for positively selected amino acids that were previously reported for PK10 colony mosquitoes (Fig. 8.4) (Bernhardt et al. 2012). It's important to note that the Thai mosquitoes used in the Bernhardt study were also colony reared, whereas the HTS libraries in the current study were from field-collected adults (Dickson et al. 2016a). The major advantages of the current approach over that used in the Bernhardt study were the use of gDNA and natural populations, which prevented the introduction of PCR artifacts and colony bottlenecks, respectively. Nevertheless, the identification of previously published positively selected replacement substitutions in PK10 provided evidence that Trinity de novo assembled haplotypes produce informative SNPs for population genetics-based gene model analyses.

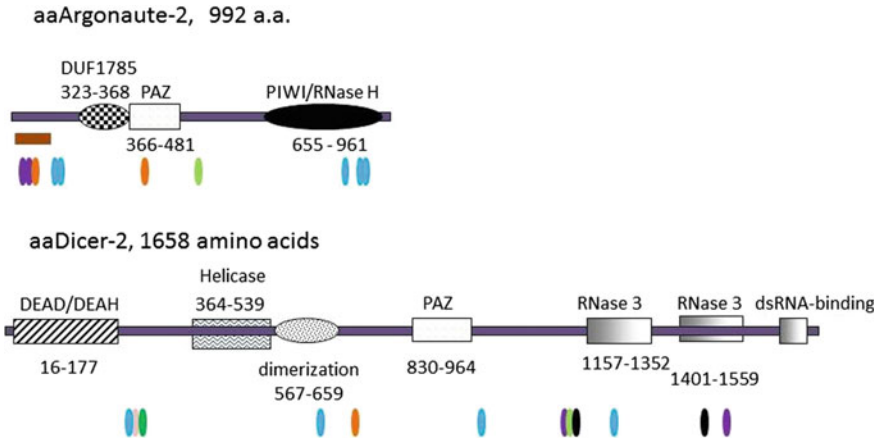


Fig. 8.4 Major anti-viral siRNAi proteins, aaArgonaute-2 and aaDicer-2, show predicted population-specific replacement substitutions. HTS exome-capture gDNA fastq libraries were subjected to limited de novo assembly in a genewise manner using default parameters in the Trinity package to generate haplotype contigs (Grabherr et al. 2011; Haas et al. 2013, personal communication, B. Haas). *Top panel*, aedine Ago2 (Vectorbase accession AAEL017251), 992 amino acids, with major protein domains marked. *Bottom panel*, Dicer-2 (Vectorbase accession AAEL006794), 1658 amino acids. Prosite profiles were used to annotate domains (Sigrist et al. 2002). *Colored ticks* below each schematic indicate collection-specific amino acid polymorphisms relative to the Liverpool reference gene (Vectorbase.org, version L3.3, Lawson et al. 2007). *Brown bar*, highly polymorphic among all populations. *Orange tab*, population-specific replacement substitution in Kaolack. *Light green tab*, replacement sites in PK10 and Kaolack. *Bright blue tab*, replacement substitution in PK10 and/or in previously published positively selected replacement substitutions (Bernhardt et al. 2012) but not in other collections. *Purple tab*, replacement substitution in PK10, Kaolack and Mexico. *Black tab*, replacement substitution site in Mexico. *Pink tab*, substitution in PK10, Thai and Mexico. *Dark green*, replacement was present in PK10, Kaolack and Thai

In general, the two Senegalese populations showed a higher number of coding substitutions, which is consistent with their status as the more ancient populations of those examined (Fig. 8.4). Among all populations, predicted Ago2 N-terminus showed a highly variable region, which likely contributed significantly to high Ka/Ks ratios. Little is known about the function of the Ago2 N-terminal domain in dipterans; however, the gene has a high degree of variability among mosquitoes and drosophilids (Campbell et al. 2008a) and similar polymorphisms in PCR data from colony mosquitoes (Bernhardt et al. 2012). Distal to the hypervariable region of *ago2*, up to 7 replacement substitutions were present in the haplotype data, depending on the collection. For example, in the Senegalese populations, there were two replacement substitutions in or near the PAZ domain at R375K in Kaolack pools and S533G in both Kaolack and PK10. The PAZ domain of Ago2 contains 3′-end sRNA binding sites; therefore, alteration of key amino acids could alter Ago2 slicing of viral genomes. Only the PK10 collection showed several replacement substitutions in the PIWI/RNase H domain. Amino acid substitutions Q838E, N916K, and N918S were found in this region. Two of these, N916K and N918S, are immediately adjacent to one of the RNase H catalytic residues and therefore are consistent with altered slicer activity.

Dicer-2 showed a total of 12 variable regions among the collections tested. Interestingly, there was a set of contiguous replacement substitutions in the intervening region between the DEAD and helicase C domains. All the collections had one or more substitutions in this region (Fig. 8.4). All collections except Kaolack had either N246D or were heterozygous at that site (N246D/N). Finally, all collections except Mexico showed a homozygous or heterozygous substitution at Q253K/Q. Although the function of the amino acids between the DEAD and C-terminal helicases is not completely understood, the conserved representation among all populations is consistent with a role in altered helicase specificity or function. In addition, PK10 showed a replacement substitution in the dimerization domain of *dicer-2*.

8.8 Evolution of PIWI Pathway Components

The mechanisms of PIWI pathway function and possible roles in vector competence, are much more poorly understood in mosquitoes, such as *Ae. aegypti*, than they are in model organisms. Indeed, evidence for mosquito anti-viral PIWI pathway function has primarily relied on descriptive studies showing viral sRNAs in the predicted size range of 24–30 nts (Hess et al. 2011; Morazzani et al. 2012), and limited cell culture studies showing possible roles for *piwi4* and *piwi5* in alphavirus infection (Miesen et al. 2015; Schnettler et al. 2013). It's important to note that cell culture studies are greatly limited in comparison to in vivo experiments, because natural tissue infection barriers are not present. With this in mind, and given the important role for regulation of germline development, high genetic association to population differences cannot necessarily be attributed to anti-viral traits (Fig. 8.2b).

Nevertheless, genetic association analysis showed that both *piwi3* and *piwi5* had LOD values above the 95 % cI in separate population comparisons between the *Ae. aegypti* subspecies and the Senegal collections (Fig. 8.2a, middle panel). However, neither of these genes showed evidence of strong diversifying selection at the gene-wise level (Fig. 8.2b, middle panel). These trends could be due in part to the proposed multi-functional roles played by these components, and perhaps counter-balancing evolutionary pressures of the mosquito life cycle. In contrast, *piwi1* showed Ka/Ks rates < lower 5 % cI for all pair-wise comparisons, indicative of strong purifying selection. Unfortunately, not enough is known about functional interactions of piRNA pathway orthologs to perform additional analysis of rapid evolution, as was done for the siRNAi/miRNAi pathway genes. Clearly additional work must be done to clarify the mechanisms of anti-viral function presumably played by piRNA pathway components.

8.9 Conclusion

Arbovirus pathogens and their associated vectors have co-evolved over time, which has likely resulted in reduced pathogenesis in vector hosts and altered efficiency of transmission dynamics. Analyses of HTS data, provided herein, were consistent with entomological evidence of rapid evolution of the peridomestic *Ae. aegypti* type form (Brown et al. 2011). In addition, we provided increased support for the very rapid evolution of *ago2* among type form and sylvatic *Ae. aegypti* populations, which could, in turn, drive arbovirus evolution. Focal replacement substitutions in large genes, such as *dicer-2*, which may also drive virus evolution, may not be revealed in gene-wise calculations of gene-wise Ka/Ks rates. Together these data have implications for evolution of the arbovirus-mosquito host arms race and will likely affect risk for emerging vector-borne pathogens.

Acknowledgments This work was funded under RO1 NIH AI-83368 and R21 AI109463 to WCB4.

References

Ambrose RL, Mackenzie JM (2015) Conserved amino acids within the N-terminus of the West Nile virus NS4A protein contribute to virus replication, protein stability and membrane proliferation. Virology 481:95–106. doi:10.1016/j.virol.2015.02.045

Bean AG, Baker ML, Stewart CR, Cowled C, Deffrasnes C, Wang LF, Lowenthal JW (2013) Studying immunity to zoonotic diseases in the natural host—keeping it real. Nat Rev Immunol 13(12):851–861. doi:10.1038/nri3551

Bennett KE, Olson KE, Munoz Mde L, Fernandez-Salas I, Farfan-Ale JA, Higgs S, Black WCt, Beaty BJ (2002) Variation in vector competence for dengue 2 virus among 24 collections of Aedes aegypti from Mexico and the United States. Am J Trop Med Hyg 67(1):85–92

Bernhardt SA, Simmons MP, Olson KE, Beaty BJ, Blair CD, Black WC (2012) Rapid intraspecific evolution of miRNA and siRNA genes in the mosquito Aedes aegypti. PLoS ONE 7(9): e44198. doi:10.1371/journal.pone.0044198

Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, Drake JM, Brownstein JS, Hoen AG, Sankoh O, Myers MF, George DB, Jaenisch T, Wint GR, Simmons CP, Scott TW, Farrar JJ, Hay SI (2013) The global distribution and burden of dengue. Nature 496(7446):504–507. doi:10.1038/nature12060

Bosio CF, Fulton RE, Salasek ML, Beaty BJ, Black WCT (2000) Quantitative trait loci that control vector competence for dengue-2 virus in the mosquito Aedes aegypti. Genetics 156 (2):687–698

Brown JE, McBride CS, Johnson P, Ritchie S, Paupy C, Bossin H, Lutomiah J, Fernandez-Salas I, Ponlawat A, Cornel AJ, Black WCt, Gorrochotegui-Escalante N, Urdaneta-Marquez L, Sylla M, Slotman M, Murray KO, Walker C, Powell JR (2011) Worldwide patterns of genetic differentiation imply multiple 'domestications' of Aedes aegypti, a major vector of human diseases. Proc Biol Sci 278(1717):2446–2454. doi:10.1098/rspb.2010.2469

Campbell CL, Black WCt, Hess AM, Foy BD (2008a) Comparative genomics of small RNA regulatory pathway components in vector mosquitoes. BMC Genomics 9:425. doi:10.1186/1471-2164-9-425

Campbell CL, Keene KM, Brackney DE, Olson KE, Blair CD, Wilusz J, Foy BD (2008b) Aedes aegypti uses RNA interference in defense against Sindbis virus infection. BMC Microbiol 8:47. doi:10.1186/1471-2180-8-47

Campbell CL, Harrison T, Hess AM, Ebel GD (2014) MicroRNA levels are modulated in Aedes aegypti after exposure to Dengue-2. Insect Mol Biol 23(1):132–139. doi:10.1111/imb.12070

Cummings DA, Irizarry RA, Huang NE, Endy TP, Nisalak A, Ungchusak K, Burke DS (2004) Travelling waves in the occurrence of dengue haemorrhagic fever in Thailand. Nature 427 (6972):344–347. doi:10.1038/nature02225

Dantes HG, Farfan-Ale JA, Sarti E (2014) Epidemiological trends of dengue disease in Mexico (2000–2011): a systematic literature search and analysis. PLoS Negl Trop Dis 8(11):e3158. doi:10.1371/journal.pntd.0003158

Diallo M, Ba Y, Faye O, Soumare ML, Dia I, Sall AA (2008) Vector competence of Aedes aegypti populations from Senegal for sylvatic and epidemic dengue 2 virus isolated in West Africa. Trans R Soc Trop Med Hyg 102(5):493–498. doi:10.1016/j.trstmh.2008.02.010

Dickson LB, Sanchez-Vargas I, Sylla M, Fleming K, Black WCI (2014) Vector competence in West African Aedes aegypti is flavivirus species and genotype dependent. PLoS Negl Trop Dis 8(10):e3153. doi:10.1371/journal.pntd.0003153

Dickson LB, Campbell CL, Juneja P, Jiggins FM, Sylla M, Black WCI (2016a) Exon enriched libraries reveal large genetic differences on the sex autosome within and among Aedes aegypti from Senegal, West Africa. (Personal communication)

Dickson LB, Sharakhova V, Timoshevskiy A, Fleming KL, Caspary A, Sylla M, Black WCI (2016b) Reproductive incompatibility involving Senegal Aedes aegypti (L) is associated with chromosome rearrangements. PLoS Negl Trop Dis 10(4):e0004626. doi: 10.1371/journal.pntd.0004626

Franco L, Di Caro A, Carletti F, Vapalahti O, Renaudat C, Zeller H, Tenorio A (2010) Recent expansion of dengue virus serotype 3 in West Africa. Euro Surveill 15(7):578–583

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29 (7):644–652. doi:10.1038/nbt.1883

Grun D, Wang YL, Langenberger D, Gunsalus KC, Rajewsky N (2005) microRNA target predictions across seven Drosophila species and comparison to mammalian targets. PLoS Comput Biol 1(1):e13. doi:10.1371/journal.pcbi.0010013

Gubler DJ (2002) The global emergence/resurgence of arboviral diseases as public health problems. Arch Med Res 33(4):330–342

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, Macmanes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, Leduc RD, Friedman N, Regev A (2013) De novo transcript sequence reconstruction from RNA-seq using the trinity platform for reference generation and analysis. Nat Protoc 8(8):1494–1512. doi:10.1038/nprot.2013.084

Hess AM, Prasad AN, Ptitsyn A, Ebel GD, Olson KE, Barbacioru C, Monighetti C, Campbell CL (2011) Small RNA profiling of Dengue virus-mosquito interactions implicates the PIWI RNA pathway in anti-viral defense. BMC Microbiol 11:45. doi:10.1186/1471-2180-11-45

Hirano M, Yoshii K, Sakai M, Hasebe R, Ichii O, Kariwa H (2014) Tick-borne flaviviruses alter membrane structure and replicate in dendrites of primary mouse neuronal cultures. J Gen Virol 95(Pt 4):849–861. doi:10.1099/vir.0.061432-0

Juneja P, Ariani CV, Ho YS, Akorli J, Palmer WJ, Pain A, Jiggins FM (2015) Exome and transcriptome sequencing of Aedes aegypti identifies a locus that confers resistance to Brugia malayi and alters the immune response. PLoS Pathog 11(3):e1004765. doi:10.1371/journal.ppat.1004765

Kakumani PK, Ponia SS, S RK, Sood V, Chinnappan M, Banerjea AC, Medigeshi GR, Malhotra P, Mukherjee SK, Bhatnagar RK (2013) Role of RNA interference (RNAi) in dengue virus replication and identification of NS4B as an RNAi suppressor. J Virol 87 (16):8870–8883. doi:10.1128/JVI.02774-12

Keene KM, Foy BD, Sanchez-Vargas I, Beaty BJ, Blair CD, Olson KE (2004) RNA interference acts as a natural antiviral response to O'nyong-nyong virus (Alphavirus; Togaviridae) infection of Anopheles gambiae. Proc Natl Acad Sci USA 101(49):17240–17245. doi:10.1073/pnas.0406983101

Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, Miller CA, Mardis ER, Ding L, Wilson RK (2012) VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 22(3):568–576. doi:10.1101/gr.129684.111

Lambrechts L, Chevillon C, Albright RG, Thaisomboonsuk B, Richardson JH, Jarman RG, Scott TW (2009) Genetic specificity and potential for local adaptation between dengue viruses and mosquito vectors. BMC Evol Biol 9:160. doi:10.1186/1471-2148-9-160

Lambrechts L, Paaijmans KP, Fansiri T, Carrington LB, Kramer LD, Thomas MB, Scott TW (2011) Impact of daily temperature fluctuations on dengue virus transmission by Aedes aegypti. Proc Natl Acad Sci USA 108(18):7460–7465. doi:10.1073/pnas.1101377108

Lambrechts L, Quillery E, Noel V, Richardson JH, Jarman RG, Scott TW, Chevillon C (2013) Specificity of resistance to dengue virus isolates is associated with genotypes of the mosquito antiviral gene Dicer-2. Proc Biol Sci 280(1751):20122437. doi:10.1098/rspb.2012.2437

Lawson D, Arensburger P, Atkinson P, Besansky NJ, Bruggner RV, Butler R, Campbell KS, Christophides GK, Christley S, Dialynas E, Emmert D, Hammond M, Hill CA, Kennedy RC, Lobo NF, MacCallum MR, Madey G, Megy K, Redmond S, Russo S, Severson DW, Stinson EO, Topalis P, Zdobnov EM, Birney E, Gelbart WM, Kafatos FC, Louis C, Collins FH (2007) VectorBase: a home for invertebrate vectors of human pathogens. Nucleic Acids Res 35 (Database issue):D503–505. doi:10.1093/nar/gkl960

Lee YS, Nakahara K, Pham JW, Kim K, He Z, Sontheimer EJ, Carthew RW (2004) Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell 117(1):69–81

Leger P, Lara E, Jagla B, Sismeiro O, Mansuroglu Z, Coppee JY, Bonnefoy E, Bouloy M (2013) Dicer-2- and Piwi-mediated RNA interference in Rift Valley fever virus-infected mosquito cells. J Virol 87(3):1631–1648. doi:10.1128/JVI.02795-12

Mattingly PF (1958) Genetical aspects of the Aedes aegypti problem. II. Disease relationships, genetics and control. Ann Trop Med Parasitol 52(1):5–17

McBride CS, Baier F, Omondi AB, Spitzer SA, Lutomiah J, Sang R, Ignell R, Vosshall LB (2014) Evolution of mosquito preference for humans linked to an odorant receptor. Nature 515 (7526):222–227. doi:10.1038/nature13964

Miesen P, Girardi E, van Rij RP (2015) Distinct sets of PIWI proteins produce arbovirus and transposon-derived piRNAs in Aedes aegypti mosquito cells. Nucleic Acids Res 43(13):6545–6556. doi:10.1093/nar/gkv590

Moncayo AC, Fernandez Z, Ortiz D, Diallo M, Sall A, Hartman S, Davis CT, Coffey L, Mathiot CC, Tesh RB, Weaver SC (2004) Dengue emergence and adaptation to peridomestic mosquitoes. Emerg Infect Dis 10(10):1790–1796. doi:10.3201/eid1010.030846

Moore DF (1979) Hybridization and mating behavior in Aedes aegypti (Diptera: Culicidae). J Med Entomol 16(3):223–226

Morazzani EM, Wiley MR, Murreddu MG, Adelman ZN, Myles KM (2012) Production of virus-derived ping-pong-dependent piRNA-like small RNAs in the mosquito soma. PLoS Pathog 8(1):e1002470. doi:10.1371/journal.ppat.1002470

Nguyet MN, Duong TH, Trung VT, Nguyen TH, Tran CN, Long VT, le Dui T, Nguyen HL, Farrar JJ, Holmes EC, Rabaa MA, Bryant JE, Nguyen TT, Nguyen HT, Nguyen LT, Pham MP, Luong TT, Wills B, Nguyen CV, Wolbers M, Simmons CP (2013) Host and viral features of human dengue cases shape the population of infected and infectious Aedes aegypti mosquitoes. Proc Natl Acad Sci USA 110(22):9072–9077. doi:10.1073/pnas.1303395110

Obbard DJ, Jiggins FM, Halligan DL, Little TJ (2006) Natural selection drives extremely rapid evolution in antiviral RNAi genes. Curr Biol 16(6):580–585. doi:10.1016/j.cub.2006.01.065

Pham JW, Sontheimer EJ (2005) Molecular requirements for RNA-induced silencing complex assembly in the Drosophila RNA interference pathway. J Biol Chem 280(47):39278–39283. doi:10.1074/jbc.M509202200

Rosenberg R, Johansson MA, Powers AM, Miller BR (2013) Search strategy has influenced the discovery rate of human viruses. Proc Natl Acad Sci USA 110(34):13961–13964. doi:10.1073/pnas.1307243110

Roth A, Mercier A, Lepers C, Hoy D, Duituturaga S, Benyon E, Guillaumot L, Souares Y (2014) Concurrent outbreaks of dengue, chikungunya and Zika virus infections—an unprecedented epidemic wave of mosquito-borne viruses in the Pacific 2012–2014. Euro Surveill 19(41)

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(16):2214–2222. doi:10.1101/gad.1454806

Sanchez-Vargas I, Scott JC, Poole-Smith BK, Franz AW, Barbosa-Solomieu V, Wilusz J, Olson KE, Blair CD (2009) Dengue virus type 2 infections of Aedes aegypti are modulated by the mosquito's RNA interference pathway. PLoS Pathog 5(2):e1000299. doi:10.1371/journal.ppat.1000299

Schirtzinger EE, Andrade CC, Devitt N, Ramaraj T, Jacobi JL, Schilkey F, Hanley KA (2015) Repertoire of virus-derived small RNAs produced by mosquito and mammalian cells in response to dengue virus infection. Virology 476:54–60. doi:10.1016/j.virol.2014.11.019

Schnettler E, Sterken MG, Leung JY, Metz SW, Geertsema C, Goldbach RW, Vlak JM, Kohl A, Khromykh AA, Pijlman GP (2012) Noncoding flavivirus RNA displays RNA interference suppressor activity in insect and Mammalian cells. J Virol 86(24):13486–13500. doi:10.1128/JVI.01104-12

Schnettler E, Donald CL, Human S, Watson M, Siu RW, McFarlane M, Fazakerley JK, Kohl A, Fragkoudis R (2013) Knockdown of piRNA pathway proteins results in enhanced Semliki Forest virus production in mosquito cells. J Gen Virol 94(Pt 7):1680–1689. doi:10.1099/vir.0.053850-0

Schountz T, Acuna-Retamar M, Feinstein S, Prescott J, Torres-Perez F, Podell B, Peters S, Ye C, Black WCt, Hjelle B (2012) Kinetics of immune responses in deer mice experimentally infected with Sin Nombre virus. J Virol 86(18):10015–10027. doi:10.1128/JVI.06875-11

Sigrist CJ, Cerutti L, Hulo N, Gattiker A, Falquet L, Pagni M, Bairoch A, Bucher P (2002) PROSITE: a documented database using patterns and profiles as motif descriptors. Brief Bioinform 3(3):265–274

Sylla M, Bosio C, Urdaneta-Marquez L, Ndiaye M, Black WCt (2009) Gene flow, subspecies composition, and dengue virus-2 susceptibility among Aedes aegypti collections in Senegal. PLoS Negl Trop Dis 3(4):e408. doi:10.1371/journal.pntd.0000408

Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD (2006) A distinct small RNA pathway silences selfish genetic elements in the germline. Science 313(5785):320–324. doi:10.1126/science.1129333

van Rij RP, Saleh MC, Berry B, Foo C, Houk A, Antoniewski C, Andino R (2006) The RNA silencing endonuclease Argonaute 2 mediates specific antiviral immunity in Drosophila melanogaster. Genes Dev 20(21):2985–2995. doi:10.1101/gad.1482006

Weaver SC, Reisen WK (2010) Present and future arboviral threats. Antiviral Res 85(2):328–345. doi:10.1016/j.antiviral.2009.10.008

Wu TD, Nacu S (2010) Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics 26(7):873–881. doi:10.1093/bioinformatics/btq057

Xi Z, Ramirez JL, Dimopoulos G (2008) The Aedes aegypti toll pathway controls dengue virus infection. PLoS Pathog 4(7):e1000098. doi:10.1371/journal.ppat.1000098

Chapter 9
Differential Expression of *Toxoplasma gondii* MicroRNAs in Murine and Human Hosts

Müşerref Duygu Saçar Demirci, Caner Bağcı and Jens Allmer

Abstract MicroRNAs are short RNA sequences involved in post-transcriptional gene regulation. MicroRNAs are known for a wide variety of species ranging from bacteria to plants. It has become clear that some cross-kingdom regulation is possible especially between viruses and their hosts. We hypothesized that intracellular parasites, like *Toxoplasma gondii*, similar to viruses would be able to modulate their host's gene expression. We were able to show that *T. gondii* produces many putative pre-miRNAs which are actually transcribed. Furthermore, some of these expressed pre-miRNAs have a striking resemblance to host mature miRNAs. Previous studies indicated that *T. gondii* infection coincides with increased abundance of some miRNAs. Here we were able to show that many of these miRNAs have close relatives in *T. gondii* which may not be distinguishable using PCR. Taken together, the similarity to host miRNAs, their confirmed expression, and their upregulation during infection, it suggests that *T. gondii* actively transfers miRNAs to regulate its host. We conclude, that this type of cross-kingdom regulation may be possible, but that targeted analysis is necessary to consolidate our computational findings.

9.1 Introduction

Mature microRNAs (miRNAs) are short (18–24 nt) nucleotide sequences which act as a recognition key via base-pairing with their target mRNAs involved in post-transcriptional gene regulation (Erson-Bensan 2014). Mature miRNAs are cut from larger stem-loop structures (Dicer for animals and DCL1 in the nucleus for

M.D. Saçar Demirci · J. Allmer (✉)
Molecular Biology and Genetics, Izmir Institute of Technology, 35430 Urla, Izmir, Turkey
e-mail: jens@allmer.de

C. Bağcı
Biotechnology, Izmir Institute of Technology, 35430 Urla, Izmir, Turkey

J. Allmer
Bionia Incorporated, IZTEKGEB A8, 35430 Urla, Izmir, Turkey

A.L. Leitão and F.J. Enguita (eds.), *Non-coding RNAs and Inter-kingdom Communication*, DOI 10.1007/978-3-319-39496-1_9

plants) and incorporated into protein complexes in the cytoplasm which then perform the regulatory action. The stem-loop structures are cut from pre-miRNAs (hairpins) within the nucleus (Drosha for animals and DCL1 for plants) and then exported into the cytosol. The pre-miRNAs are excised from pri-miRNA during transcription and they may derive from any part of a genome (Kim et al. 2009; Rodriguez et al. 2004) transcribed by either RNA-Polymerase II or III. The existence of miRNAs has been shown for a wide range of organisms ranging from protists (Gottesman 2005) and sponges (Kim et al. 2009) to plants (Xie et al. 2015) and animals (Okamura 2012). Even viruses contain miRNAs (Grey 2015); likely to also influence the host's gene expression (Li et al. 2014; Skalsky and Cullen 2010).

For *Toxoplasma gondii*, it has been shown that miRNAs and the necessary machinery for genesis and targeting exist (Braun et al. 2010) and that it is important during development (Hakimi and Menard 2010). Interestingly, the miRNA genesis pathway of *T. gondii* seems related to plant and fungal ones (Braun et al. 2010) which is striking since the plant miRNA genesis pathway may have evolved independently from the metazoan one (Chapman and Carrington 2007). It is even more curious that the targeting machinery in *T. gondii* seems related to the metazoan one (Braun et al. 2010). Around the same time with these findings, we predicted potential miRNA regulatory networks in *T. gondii* under the assumption of similarity to metazoan miRNAs (Cakir and Allmer 2010). This analysis may have to be redone using a mixture model although it led to the interesting finding that there are perhaps as many hub miRNAs in the *T. gondii* genome as there are major hosts. In summary, it has been established that *T. gondii* uses miRNA for its post transcriptional gene regulation.

Like viruses, obligatory intracellular parasites may benefit from influencing their hosts' gene expression. In this regard, it has been shown that *T. gondii* increases the amount of key miRNAs (miR-17 ~ 92 and miR-106b ~ 25), implicated in numerous hyperproliferative diseases, during infection (Zeiner et al. 2010). Zeiner et al. further demonstrated that this process is tied to *T. gondii* and that at least one other apicomplexan (*Neospora caninum*) does not impact the level of the selected miRNAs. Fast miRNA evolution has been reported before (Liang and Li 2009; Lu et al. 2008) and this seems to be further supported with the study by Zeiner et al. supporting their implicit idea that such regulation may be parasite-host specific. Other studies support this finding and show *T. gondii* dependent dysregulation of miRNAs (Cai et al. 2014; Cannella et al. 2014; Thirugnanam et al. 2013; Xiao et al. 2014).

We pondered how such dysregulation may be caused and hypothesized that a possible path would be for *T. gondii* to export miRNAs into its host. We then analyzed the miRNAs of *T. gondii* and concluded that communication between parasite and host via miRNAs is theoretically possible (Saçar et al. 2014) which is further supported by studies which show that secretion is important for parasite host communication (Boothroyd and Dubremetz 2008; Luder et al. 2009); for a recent

review see (English et al. 2015). Since *T. gondii* is present in a large number of hosts, we wondered whether some putative miRNAs are expressed host-specifically or whether some are highly expressed in multiple hosts. To investigate this, we used all human and mouse miRNAs available on miRBase (Griffiths-Jones et al. 2008) to establish machine learned models and applied them to all hairpins of the *T. gondii* ME49 genome. We acquired all available next generation sequencing datasets where any toxoplasma strain was sequenced within a host cell or singly from SRA (Leinonen et al. 2011) and evaluated the expression profiles of the predicted miRNAs. Clearly, the results can only guide further research since it is necessary to have evidence on the protein level for confident miRNA assignment as described in (Saçar and Allmer 2013c) something which can be achieved using mass spectrometry with quantitation (Allmer 2010). Mass spectrometric experiments can be guided by providing expected proteotypic peptide masses of proteins expected to be miRNA targets using multiple reaction monitoring type experiments (Kondrat et al. 1978).

T. gondii is a unicellular protist and as such is not expected to have many miRNAs with clear human or mouse like features. Yet, we found many such examples (2761066 and 2299419 hairpins passed human and mouse models at a prediction cutoff of 0.99) and they are conserved (5278 and 8190 hairpins showed high similarity to human and mouse mature microRNAs with a matchScore above or equal to 35) or differentially expressed (57789 and 64598 hairpins are differentially expressed in human and murine hosts with an absolute log2 fold change equal to or larger than 1 and a Benjamini Hochberg (BH) adjusted p-value equal to or smaller than 0.01) in mouse and human hosts. Since *T. gondii*'s targeting machinery is metazoan like, there are two modes of directly influencing the host's post transcriptional gene regulation. We suggest that either *T. gondii* transports naked stem-loop structures into the host which are then processed by the host machinery, or it can export loaded TgAgo (Al Riyahi et al. 2006) which may be directly functional in the host.

Host like miRNAs which are highly expressed by *T. gondii* in both hosts may have indispensable function for *T. gondii*; and gene enrichment using Reactome (Croft et al. 2014) indicates that this is very likely true for the miRNAs identified in this study. MicroRNAs that are differentially expressed between hosts may modulate *T. gondii*'s interference with host gene expression to specific hosts. Since *T. gondii* replicates in many hosts, and since it seems to be able to distinguish between primary and secondary hosts, there must be a mechanism allowing this distinction and miRNAs may play a role in communicating it. Both types of identified miRNAs can aid in developing new diagnostic biomarkers and may serve as drug targets (Hoy et al. 2014; Manzano-Roman and Siles-Lucas 2012). Although drugs against *T. gondii* are currently available, the need for novel ones is evident (Blader and Saeij 2009).

9.2 Materials and Methods

9.2.1 *Hairpin Extraction from the* T. gondii *Genome*

The genome assembly of *T. gondii* ME49 was downloaded from ToxoDB (Gajria et al. 2008) Release 25 (http://toxodb.org/common/downloads/release-25/TgondiiME49/fasta/data/). With an in-house script we divided the genome into 500 nt long sequences with 250 nt overlap for both strands. These sequence lengths were chosen since there is no human or murine hairpin in miRBase longer than 250 nt; so all potential pre-miRNAs in the *T. gondii* genome can be completely captured within one 500 nt fragment. The resulting 263964 fragments were then folded using RNAFold (Hofacker 2003). The structure provided by RNAFold was parsed using an in-house script which extracted all hairpin-like structures only requiring a stem with at least three consecutive bonds somewhere in the stem and a terminal loop with at least 3 nucleotides. All extracted hairpins ($\sim$4.8 million) were then examined with human and murine models established via machine learning.

9.2.2 *MicroRNA Feature Calculation*

There are no miRNA examples available on miRBase for *T. gondii* or other Apicomplexa and therefore targeted feature selection is not possible. Since the interest of the present study is on how *T. gondii* may regulate its host gene expression using miRNAs, this is not a problem and the study can draw from a larger body of knowledge describing human miRNA prediction. About a dozen studies performing ab initio pre-miRNA prediction have been published (Batuwita and Palade 2009; Bentwich 2008; Bentwich et al. 2005; Ding et al. 2010; Gao et al. 2012; Gudys et al. 2013; Jiang et al. 2007; Lopes Ide et al. 2014; Ng and Mishra 2007; Ritchie et al. 2012; van der Burgt et al. 2009; Xu et al. 2008; Xue et al. 2005). All features and their natural extensions and normalizations were implemented in JAVA™ and since calculations for the analysis of all putative hairpins extracted from the *T. gondii* genome were time consuming, HTCondor (Thain et al. 2005) was used for parallel computation.

9.2.3 *Machine Learning*

MicroRNA detection is quite involved experimentally and, therefore, computational methods are indispensable (Ng and Mishra 2007). Numerous methods for computational detection of pre-miRNAs have been developed (Allmer 2014; Allmer and Yousef 2012; Saçar et al. 2014). Most approaches are based in machine learning, specifically two-class classification although some use one-class

classification (Yousef et al. 2008, 2015). For successful machine learning, the pre-miRNAs need to be parameterized and many features to describe hairpins have been proposed (Saçar and Allmer 2013b, 2013c). The most important factor for effective machine learning is, however, data with positive and negative examples correctly assigned (Saçar and Allmer 2013b). We took into account problems with available data (Saçar et al. 2013) and class-imbalance (Saçar and Allmer 2013b) while training models using KNIME (Berthold et al. 2009). We previously showed that as little as 50 features may be enough for training a classifier and, therefore, we performed feature selection prior to training of the classifier (Fig. 9.1). We first clustered the about 700 features using *k*-Means clustering into 100 clusters. Then 50 clusters were selected based on calculation speed and information gain and from each cluster, the feature with highest information gain was included into our final feature set (Supplementary Table 2). In previous studies we have not seen a great impact of the classifier on classification efficiency (Saçar and Allmer 2013a) and, therefore, chose to use decision tree as our model since it is very fast and allows serialization of the model in KNIME.

For positive data we used all pre-miRNAs from miRBase for *Homo sapiens* (hsa) and *Mus musculus* (mmu) and as negative data we used the pseudo dataset (Ng and Mishra 2007). For the training of the classifier thousand fold Monte Carlo (MC) cross validation (Xu and Liang 2001) was performed and for each fold randomly sampled 70 % of the positive data for training and 30 % of it for testing. Negative data was randomly sampled from the pseudo dataset such the number of examples matched the number of positive examples. The models trained for hsa and mmu were then applied to all hairpins that could be extracted from the *T. gondii*

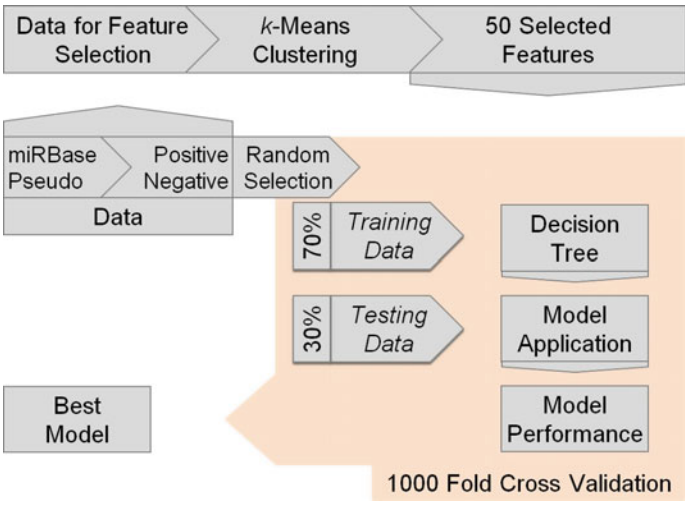


Fig. 9.1 Machine learning regime for establishing models that describe pre-miRNAs for different organisms. Feature selection was performed on a per species basis

ME49 genome assembly. Commonly used performance measures like specificity, sensitivity and accuracy were calculated.

9.2.4 MicroRNA Expression Analysis

All *T. gondii* sequencing datasets (74) available on the small sequencing read archive (Leinonen et al. 2011) were downloaded and analyzed using FastQC (Andrews 2010). 14 datasets were finally accepted for further analysis (Supplementary Table 1). Supplementary Table 1 also details why some of the datasets did not fit the present study and why they were excluded with the major reasons being low base-calling quality and overrepresentation of a dataset. The datasets were from parasite cultures (3), human foreskin fibroblast cells (3), murine bone-marrow macrophages (3) and mouse peritoneal exudate (5).

All accepted datasets were trimmed from adapter contaminations using an in-house script to determine adapter sequences and cut adapt (Martin 2010) in default settings to remove them. Remaining sequences were further trimmed from low-quality and ambiguous base callings with sickle (http://omictools.com/sickle-s714.html) using a quality threshold of 30 (implying a 0.001 base-calling error probability). Reads shorter than 20 nt after the two step trimming processes were discarded (Supplementary Table 1).

The ToxoDB Release 25 genome was indexed using Bowtie 2 (Langmead and Salzberg 2012) and trimmed reads were aligned to the genome using Tophat 2 (Kim et al. 2013) with default settings which include a maximum edit distance of 2 and take splicing into account. Reads that were ambiguously mapped to multiple locations on the genome and reads that did not map in respect to proper pairs for the paired-end sequencing datasets (mate pairs mapped on different chromosomes or wrong orientation) were excluded from further analysis.

Read counts were established using featureCounts (Liao et al. 2014) from the Subread package (http://subread.sourceforge.net) for both annotated *Toxoplasma gondii* genes from ToxoDB Release 25 and for the predicted hairpins that were mapped back to the genome by BLAST (Altschul and Gish 1996). All blast hits were recorded that pass two criteria: (1) predicted hairpin being longer than 40 nucleotides and (2) the hairpin matching to the genome perfectly. Reads spanning two or more genes/hairpins were assigned to the gene/hairpin with the largest overlap. For paired-end sequencing, fragments (read pairs) were counted instead of reads.

We previously established a similarity measure to measure the similarity of mature sequences among selected microRNAs and predicted hairpins (Saçar et al. 2014). Aligning them using BLAST with ‘blastn-short‘ option and disallowing any mismatches from negative strand of both the mature sequences and the hairpins. The resulting matchScore was assigned to each hit using the following formula:

$$matchScore + (SM \cdot 5) + M - (SMM \cdot 5) - MM$$

where SM is the number of matches in the seed region (first 8 nucleotides), M is the number of matches in the non-seed region, SMM is the number of mismatches and gaps in the seed region, MM is the number of mismatches and gaps in the non-seed region. Some factors were tested for weighting of the seed region and 5 was arbitrarily chosen from a number of promising candidates.

9.3 Results and Discussion

9.3.1 MicroRNA Detection Model Training

For human and murine miRNA detection models were trained. 50 selected features (Supplementary Table 2) were calculated for positive and negative data and 1000-fold MC cross validation was performed. Accuracy for both models was always above 75 % and never reached 90 % (Fig. 9.2). The median accuracy for human was slightly above 83 % and thus about 1 % better than the murine one which was about 82 %. Both range and interquartile range were slightly larger for the generated murine model.

The models selected for miRNA detection for human and murine achieved an accuracy of 0.875 and 0.865, respectively. This accuracy is slightly better than the best accuracy achieved in our previous study (Saçar et al. 2014), but expectedly lower when compared to accuracies achieved earlier using feature sets from

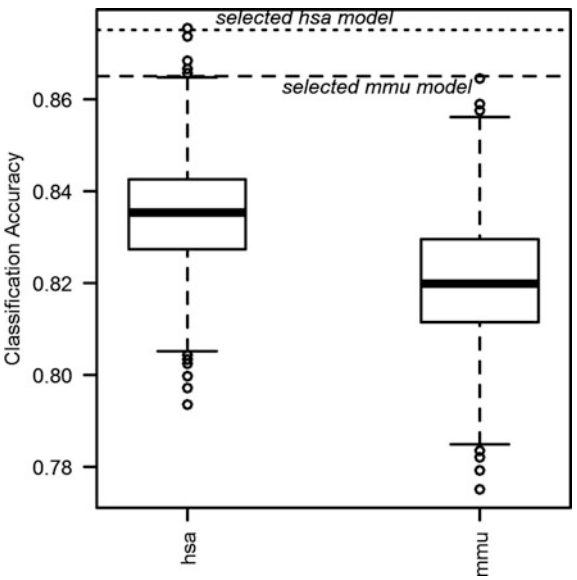


Fig. 9.2 Model accuracy for 1000-fold Monte Carlo cross validation during training of the human and murine classifiers. The selected human model (0.875) and the selected murine model (0.865) are displayed as *dotted* and *dashed lines*, respectively

ab initio studies trained with randomly generated negative data (Saçar and Allmer 2013b). Thus these models seemed applicable for the analysis of putative hairpins extracted from the *T. gondii* genome.

9.3.2 Hairpin Extraction and Analysis

Both strands of the genome were fragmented into 500 nt pieces with 250 nt overlap. The secondary structures for all fragments were predicted using RNAFold. About 2.4 million hairpin-like structures were found in both strands and two examples are given in Fig. 9.3.

Only hairpins with a minimum of three consecutive bonds in the stem (Fig. 9.3; A: 1-7) were extracted from the secondary structures of the folded 500-mers. The murine and human machine learned models were applied to all ~4.8 million hairpins extracted from the genome (Fig. 9.4). While forward and reverse strands show similar prediction score distributions for hsa and mmu, hsa and murine display distinct prediction score distributions. For all further analyses, we determined a prediction score threshold at 0.99 prediction score (Fig. 9.4).

Of the ~4.8 million extracted hairpins approximately 2.3 and 2.8 million passed the models at 99 % prediction score cutoff for murine and human, respectively. In our previous study (Saçar et al. 2014), far less miRNAs passed the threshold which may occur because we didn't use hairpin length or stem length cutoffs. Human and murine predictions for the forward strand share about half of the predicted hairpins at a cutoff of 99 % model score (Fig. 9.5) and the same is true for the reverse strand (data not shown).

Clearly, not all of these hairpins can represent true pre-miRNAs. The number may also be inflated by a factor of up to two due to the creation of overlapping

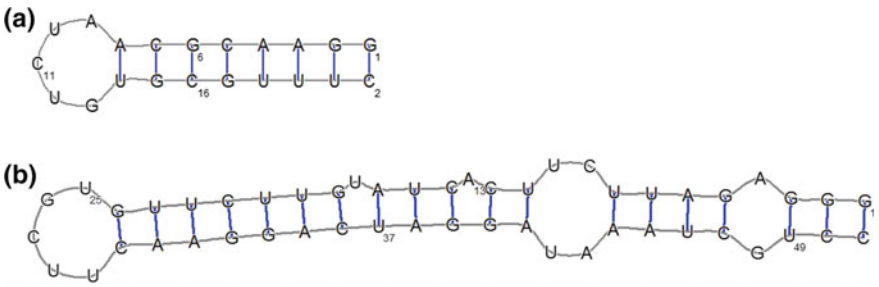


Fig. 9.3 The first two examples for stemloop structures from the forward strand of the *T. gondii* genome. The structures were drawn with RNAShapes (Steffen et al. 2006). Structure A is too short to be considered for further analysis whereas B fulfills expectations for further analysis

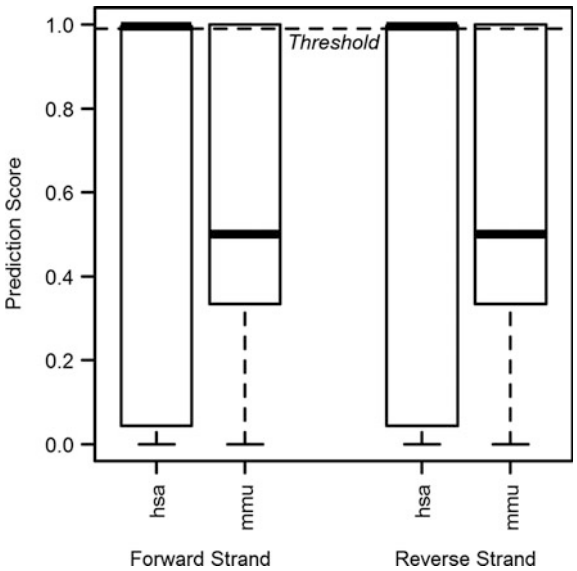


Fig. 9.4 Prediction score distribution for application of human (hsa) and mouse (mmu) models to putative hairpins extracted from the *T. gondii* genome. The 4.8 million hairpins were randomly sampled down to 100.000 to enable plotting. The *dashed line* indicates the 99 % score cutoff

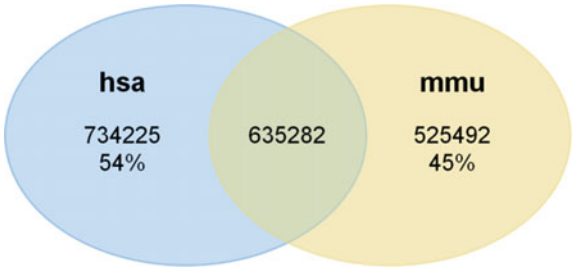


Fig. 9.5 Number of hairpins (coding strand only) passing the human (hsa) and mouse (mmu) models at a score cutoff of 0.99. The percentages are in respect to the organism i.e.: 46 % of predicted *T. gondii* hairpins that pass the hsa model also pass the mmu model whereas, when viewed from the mmu perspective it amounts to 55 %

fragments which was not accounted for, since folds are quite different and hairpins that are found at the same locus may still be quite different in consecutive overlapping fragments. Another convoluting factor is that the models achieved an accuracy of around 87 % which means that at least 13 % of hairpins that pass these models at a cutoff of 99 % are still wrongly predicted. A good way to filter such a large number of results is to check whether they are actually expressed.

9.3.3 Hairpin Expression Analysis

Many parts of a genome can be folded into structures strongly resembling pre-miRNAs. However, only if the hairpins are expressed can they perform any function and therefore expression was analyzed using RPKM (Mortazavi et al. 2008) as the measure (Fig. 9.6). We further analyzed whether there is a relationship between prediction score and hairpin expression abundance but were not able to find any (Fig. 9.1 in Supplementary File 1).

Most of the predicted hairpins were not expressed in at least one sample and, therefore, were not further considered. As can be expected, with increasing RPKM less hairpins fall into the ranges (Fig. 9.6). It is of note, that some of the expression values of the predicted hairpins for *T. gondii* are above the median expression of hairpins (hsa) provided in miRBase (~1000 RPKM).

Even more instructive than the pure expression of a predicted hairpin is whether it is shared among multiple organisms (Fig. 9.7). With the exception of Kunming mice which share only few highly expressed predicted pre-miRNAs with other organisms (perhaps due to the sequencing strategy), the larger part of highly expressed pre-miRNAs (~60 %) are shared among species (Fig. 9.7). 4071 predicted *T. gondii* pre-miRNAs are shared among the murine, human, and parasite

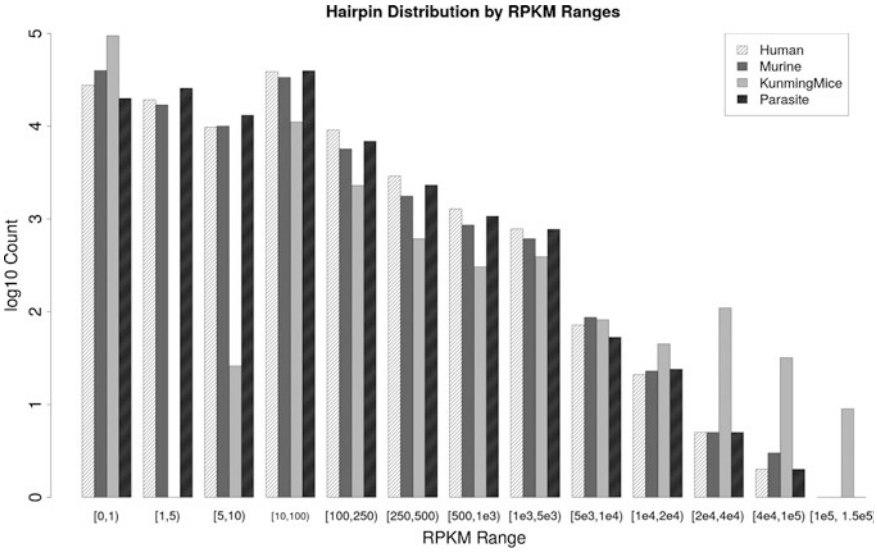


Fig. 9.6 Only hairpins that were expressed in at least one sample were considered for expression analysis. The predicted *T. gondii* hairpins were grouped by their RPKM and the counts are provided for the samples (pooled by species). The groups were chosen arbitrarily such that a small number of groups can capture the complete range of RPKM values

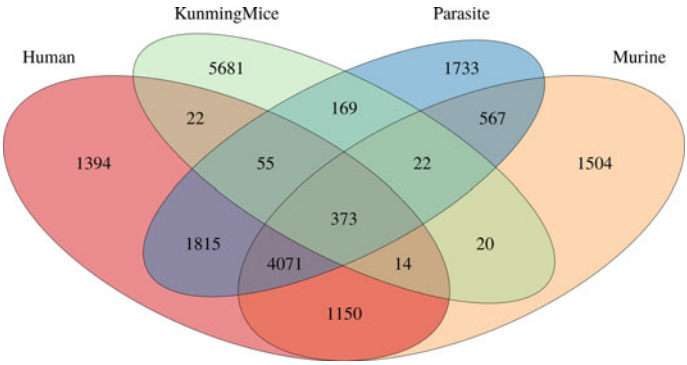


Fig. 9.7 Predicted *T. gondii* hairpins that are expressed above 1000 RPKM in the samples and how they are shared among species that were investigated

samples. Murine and parasite share an additional 5033, human and parasite 6314, and human and murine 5608 of the highly expressed predicted *T. gondii* pre-miRNAs. The number of predicted hairpins which are uniquely expressed in one of the species is, therefore, much lower than the miRNAs whose expression is shared among samples from different species (human: 1394, murine: 1504, and parasite: 1733). MicroRNAs modulate gene expression and this modulation can be systemic or specific to the environment. It seems likely that the predicted hairpins perform a fundamental function whereas the ones that are highly expressed in only one or two samples perform a more specific function.

The shared predicted hairpins discussed above are expressed at similar rates as hairpins typically presented on miRBase (http://www.mirbase.org/cgi-bin/miRNA_summary.pl?org=hsa) which for human range between 0.1 and 135000 with a median of 97 and an average of 1100 RPM. Unfortunately, these values cannot be translated into absolute concentrations which would be essential for drawing conclusions about the abundance of miRNAs in the different samples.

It is of interest whether some of the miRNAs previously found to be dysregulated in human or mouse hosts are expressed in *T. gondii*. For this we matched the predicted hairpins to human and mouse mature miRNAs using the matchScore with a threshold of 35. The hairpins were further required to be expressed in at least one sample.

It has been shown that miR-17 ~ 92, miR-106b ~ 25, and miR-106a ~ 363 increase in abundance within the host upon infection with *T. gondii* (Cai et al. 2014; Zeiner et al. 2010). Interestingly, predicted *T. gondii* hairpins similar to miR-18a, miR-20b, miR-92, and miR-363 were expressed in the human and murine samples (Supplementary Table 3). Another study showed that mmu-miR-712-3p, mmu-miR-511-5p and mmu-miR-217-5p are indicative for *T. gondii* infection in mice. Very similar hairpins to the former two miRNAs are also predicted for *T. gondii* in this study and they are expressed at relatively high levels (Supplementary Table 3). Xiao et al. reported the upregulation of miR-30c-1, miR-125b-2,

miR-132, mir-23b, and miR-17 ~ 92 (Xiao et al. 2014). Except for miR125-b-2 and miR-132 we also found similar *T. gondii* hairpins expressed in human and murine samples (Supplementary Table 3). Wang et al. analyzed *T. gondii* in respect to its miRNAs and found 17 conserved miRNAs 2 of which were related to metazoan miRNA families (Wang et al. 2012). Here, similar *T. gondii* pre-miRNAs to mmu-miR-466i-5p, mmu-miR-574-5p, and has-miR-574-5p are also found. Toxoplasma persistence in human brain seems to be associated with miR-146a, miR-155, and miR-1246 (Cannella et al. 2014). While we did not find miR-146a, we found similar *T. gondii* hairpins to miR-155 and miR-1246 to be expressed in human and/or murine samples (Supplementary Table 3). In summary, for many of the upregulated miRNAs indicative of toxoplasmosis there exist highly similar hairpins in *T. gondii*. Often PCR is performed to validate the existence of miRNAs in the referenced studies, but it is not clear whether the host miRNAs can be distinguished from the very similar *T. gondii* hairpins.

The targets of the miRNAs which have previously been reported to be upregulated in the host were submitted to Reactome analysis (D'Eustachio 2011). For all predicted *T. gondii* miRNAs similar to the previously reported miRNAs (both human and murine) that were expressed in both human and murine samples the gene targets are always signifying increased metabolism and infection (Supplementary Table 4).

9.3.4 Known Toxoplasma Gondii *Pre-microRNAs*

A recent study identified 339 novel miRNAs in Toxoplasma and compared expression between RH and ME49 strains (Wang et al. 2012). In order to check whether the miRNAs were also among the hairpins identified in this study, we acquired their data. Only perfect complete matches of their mature miRNA sequences to our putative pre-miRNAs were accepted using blastn-short (Supplementary Table 5). 48 of the mature sequence were not part of our predicted hairpins. 43 out of these were found in the sequence assembly we used, but were not part of hairpins that passed the human or mouse model. Finally, the remaining 4 mature sequences were not found in the *T. gondii* genome assembly we used in this study. Interestingly, the average expression of matches in their data is ~2080 and ~2270 while the average expression for the reported mature sequences without counterpart in our data is 1.4 and 4.3. All matching sequences also pass the human model at a prediction score cutoff of 0.99 which indicates that the identified miRNAs by Wang and colleagues (Wang et al. 2012) are very similar to human type pre-miRNAs and thus confirms the claim that the miRNAs are metazoan like (Braun et al. 2010).

9.4 Conclusion

MicroRNAs in *T. gondii* have metazoan like features in respect to targeting (Braun et al. 2010). Therefore, predictive models based on human and murine miRNAs seem to be applicable and the models trained here achieved around 87 % accuracy (Fig. 9.2). Due to the similarity and the fact that transfer from *T. gondii* to the host cell is possible, it can be reasoned that *T. gondii* hairpins may perform regulatory function in the host. Some of the predicted hairpins are expressed in various hosts and of these a large number is shared among them. Additionally, some predicted hairpins are highly similar to host miRNAs and are relatively abundant. Many of the miRNAs that have previously been shown to be upregulated in diseases have highly expressed close *T. gondii* homologs. Together, these clues are suggestive, but not fully conclusive, for the regulatory mechanism where *T. gondii* actively transfers miRNAs into the host to modulate gene expression.

T. gondii miRNAs, different from host miRNAs, which are highly expressed in all hosts would be good disease marker candidates. Differentially expressed miRNAs that are similar to host miRNAs may provide therapeutic leads for treatment of Toxoplasmosis. Unfortunately, sequencing data was only available for human and murine hosts. We hope that in the future we will be able to acquire data from more hosts; especially from cats.

Acknowledgments This work was supported by the Scientific and Technological Research Council of Turkey (Grant No. 113E326) awarded to JA.

Supplementary Materials Supplementary material is available at the following URL: http://bioinformatics.iyte.edu.tr/supplements/ncRNA2016

References

Al Riyahi A, Al-Anouti F, Al-Rayes M, Ananvoranich S (2006) Single Argonaute protein from *Toxoplasma gondii* is involved in the double-stranded RNA induced gene silencing. Int J Parasitol 36(9):1003–1014. doi:10.1016/j.ijpara.2006.04.014

Allmer J (2010) Label-free quantitation, an extension to 2DB. Amino Acids 38(4):1075–1087. doi:10.1007/s00726-009-0317-9

Allmer J (2014) Computational and bioinformatics methods for microRNA gene prediction. Methods Mol Biol 1107:157–175. doi:10.1007/978-1-62703-748-8_9

Allmer J, Yousef M (2012) Computational methods for ab initio detection of microRNAs. Front Genet 3:209. doi:10.3389/fgene.2012.00209

Altschul SF, Gish W (1996) Local alignment statistics. Methods Enzymol 266:460–480

Andrews S (2010) FASTQC: a quality control tool for high throughput sequence data. Babraham. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Batuwita R, Palade V (2009) microPred: effective classification of pre-miRNAs for human miRNA gene prediction. Bioinformatics 25(8):989–995. doi:10.1093/bioinformatics/btp107

Bentwich I (2008) Identifying human microRNAs. Curr Top Microbiol Immunol 320:257–269

Bentwich I, Avniel A, Karov Y, Aharonov R, Gilad S, Barad O, Barzilai A, Einat P, Einav U, Meiri E, Sharon E, Spector Y, Bentwich Z (2005) Identification of hundreds of conserved and nonconserved human microRNAs. Nat Genet 37(7):766–770. doi:10.1038/ng1590

Berthold MR, Cebron N, Dill F, Gabriel TR, Kötter T, Meinl T, Ohl P, Thiel K, Wiswedel B (2009) KNIME—The Konstanz Information Miner

Blader IJ, Saeij JP (2009) Communication between *Toxoplasma gondii* and its host: impact on parasite growth, development, immune evasion, and virulence. APMIS 117(5–6):458–476. doi:10.1111/j.1600-0463.2009.02453.x

Boothroyd JC, Dubremetz JF (2008) Kiss and spit: the dual roles of *Toxoplasma* rhoptries. Nat Rev Microbiol 6(1):79–88. doi:10.1038/nrmicro1800

Braun L, Cannella D, Ortet P, Barakat M, Sautel CF, Kieffer S, Garin J, Bastien O, Voinnet O, Hakimi MA (2010) A complex small RNA repertoire is generated by a plant/fungal-like machinery and effected by a metazoan-like Argonaute in the single-cell human parasite *Toxoplasma gondii*. PLoS Pathog 6(5):e1000920. doi:10.1371/journal.ppat.1000920

Cai Y, Chen H, Mo X, Tang Y, Xu X, Zhang A, Lun Z, Lu F, Wang Y, Shen J (2014) *Toxoplasma gondii* inhibits apoptosis via a novel STAT3-miR-17-92-Bim pathway in macrophages. Cell Signal 26(6):1204–1212. doi:10.1016/j.cellsig.2014.02.013

Cakir MV, Allmer J (2010) Systematic computational analysis of potential RNAi regulation in *Toxoplasma gondii*. In: 2010 5th international symposium on health informatics and bioinformatics. IEEE, Ankara, Turkey, pp 31–38

Cannella D, Brenier-Pinchart MP, Braun L, van Rooyen JM, Bougdour A, Bastien O, Behnke MS, Curt RL, Curt A, Saeij JP, Sibley LD, Pelloux H, Hakimi MA (2014) miR-146a and miR-155 delineate a MicroRNA fingerprint associated with *Toxoplasma* persistence in the host brain. Cell Rep 6(5):928–937. doi:10.1016/j.celrep.2014.02.002

Chapman EJ, Carrington JC (2007) Specialization and evolution of endogenous small RNA pathways. Nat Rev Genet 8(11):884–896. doi:10.1038/nrg2179

Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G, Caudy M, Garapati P, Gillespie M, Kamdar MR, Jassal B, Jupe S, Matthews L, May B, Palatnik S, Rothfels K, Shamovsky V, Song H, Williams M, Birney E, Hermjakob H, Stein L, D'Eustachio P (2014) The Reactome pathway knowledgebase. Nucleic Acids Res 42(Database issue):D472–477. doi:10.1093/nar/gkt1102

D'Eustachio P (2011) Reactome knowledgebase of human biological pathways and processes. Methods Mol Biol 694:49–61. doi:10.1007/978-1-60761-977-2_4

Ding J, Zhou S, Guan J (2010) MiRenSVM: towards better prediction of microRNA precursors using an ensemble SVM classifier with multi-loop features. BMC Bioinformatics 11(Suppl 11): S11. doi:10.1186/1471-2105-11-S11-S11

English ED, Adomako-Ankomah Y, Boyle JP (2015) Secreted effectors in *Toxoplasma gondii* and related species: determinants of host range and pathogenesis? Parasite Immunol 37 (3):127–140. doi:10.1111/pim.12166

Erson-Bensan AE (2014) Introduction to microRNAs in biological systems. Methods Mol Biol 1107:1–14. doi:10.1007/978-1-62703-748-8_1

Gajria B, Bahl A, Brestelli J, Dommer J, Fischer S, Gao X, Heiges M, Iodice J, Kissinger JC, Mackey AJ, Pinney DF, Roos DS, Stoeckert CJ, Jr., Wang H, Brunk BP (2008) ToxoDB: an integrated *Toxoplasma gondii* database resource. Nucleic Acids Res 36(Database issue):D553–556. doi:10.1093/nar/gkm981

Gao Z, Luo X, Shi T, Cai B, Zhang Z, Cheng Z, Zhuang W (2012) Identification and validation of potential conserved microRNAs and their targets in peach (Prunus persica). Mol Cells 34 (3):239–249. doi:10.1007/s10059-012-0004-7

Gottesman S (2005) Micros for microbes: non-coding regulatory RNAs in bacteria. Trends Genet 21:399–404. doi:10.1016/j.tig.2005.05.008

Grey F (2015) Role of microRNAs in herpesvirus latency and persistence. J Gen Virol 96(Pt 4):739–751. doi:10.1099/vir.0.070862-0

Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ (2008) miRBase: tools for microRNA genomics. Nucleic Acids Res 36(Database issue):D154–158. doi:10.1093/nar/gkm952

Gudys A, Szczesniak MW, Sikora M, Makalowska I (2013) HuntMi: an efficient and taxon-specific approach in pre-miRNA identification. BMC Bioinform 14:83. doi:10.1186/1471-2105-14-83

Hakimi MA, Menard R (2010) Do apicomplexan parasites hijack the host cell microRNA pathway for their intracellular development? F1000 Biol Rep 2. doi:10.3410/B2-42

Hofacker IL (2003) Vienna RNA secondary structure server. Nucleic Acids Res 31:3429–3431

Hoy AM, Lundie RJ, Ivens A, Quintana JF, Nausch N, Forster T, Jones F, Kabatereine NB, Dunne DW, Mutapi F, Macdonald AS, Buck AH (2014) Parasite-derived microRNAs in host serum as novel biomarkers of helminth infection. PLoS Negl Trop Dis 8 (2):e2701. doi:10.1371/journal.pntd.0002701

Jiang P, Wu H, Wang W, Ma W, Sun X, Lu Z (2007) MiPred: classification of real and pseudo microRNA precursors using random forest prediction model with combined features. Nucleic Acids Res 35(Web Server issue):W339–344. doi:10.1093/nar/gkm368

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol 14(4):R36. doi:10.1186/gb-2013-14-4-r36

Kim VN, Han J, Siomi MC (2009) Biogenesis of small RNAs in animals. Nat Rev Mol Cell Biol 10(2):126–139. doi:10.1038/nrm2632

Kondrat RW, McClusky GA, Cooks RG (1978) Multiple reaction monitoring in mass spectrometry/mass spectrometry for direct analysis of complex mixtures. Anal Chem 50:2017–2012. doi:10.1021/ac50036a020

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9 (4):357–359. doi:10.1038/nmeth.1923

Leinonen R, Sugawara H, Shumway M (2011) The sequence read archive. Nucleic Acids Res 39 (Database issue):D19–21. doi:10.1093/nar/gkq1019

Li C, Hu J, Hao J, Zhao B, Wu B, Sun L, Peng S, Gao GF, Meng S (2014) Competitive virus and host RNAs: the interplay of a hidden virus and host interaction. Protein Cell 5(5):348–356. doi:10.1007/s13238-014-0039-y

Liang H, Li WH (2009) Lowly expressed human microRNA genes evolve rapidly. Mol Biol Evol 26(6):1195–1198. doi:10.1093/molbev/msp053

Liao Y, Smyth GK, Shi W (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30(7):923–930. doi:10.1093/bioinformatics/btt656

Lopes Ide O, Schliep A, de Carvalho AC (2014) The discriminant power of RNA features for pre-miRNA recognition. BMC Bioinformatics 15:124. doi:10.1186/1471-2105-15-124

Lu J, Shen Y, Wu Q, Kumar S, He B, Shi S, Carthew RW, Wang SM, Wu CI (2008) The birth and death of microRNA genes in *Drosophila*. Nat Genet 40(3):351–355. doi:10.1038/ng.73

Luder CG, Stanway RR, Chaussepied M, Langsley G, Heussler VT (2009) Intracellular survival of apicomplexan parasites and host cell modification. Int J Parasitol 39(2):163–173. doi:10.1016/j.ijpara.2008.09.013

Manzano-Roman R, Siles-Lucas M (2012) MicroRNAs in parasitic diseases: potential for diagnosis and targeting. Mol Biochem Parasitol 186(2):81–86. doi:10.1016/j.molbiopara.2012.10.001

Martin M (2010) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10–12

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5(7):621–628. doi:10.1038/nmeth.1226

Ng KL, Mishra SK (2007) De novo SVM classification of precursor microRNAs from genomic pseudo hairpins using global and intrinsic folding measures. Bioinformatics 23(11):1321–1330. doi:10.1093/bioinformatics/btm026

Okamura K (2012) Diversity of animal small RNA pathways and their biological utility. Wiley Interdiscip Rev RNA 3(3):351–368. doi:10.1002/wrna.113

Ritchie W, Gao D, Rasko JE (2012) Defining and providing robust controls for microRNA prediction. Bioinformatics 28(8):1058–1061. doi:10.1093/bioinformatics/bts114

Rodriguez A, Griffiths-Jones S, Ashurst JL, Bradley A (2004) Identification of mammalian microRNA host genes and transcription units. Genome Res 14(10A):1902–1910. doi:10.1101/gr.2722704

Saçar MD, Allmer J (2013a) Comparison of four Ab initio MicroRNA prediction tools. In: Proceedings of the international conference on bioinformatics models, methods and algorithms. SciTePress—Science and and Technology Publications, Barcelona, Spain, pp 190–195

Saçar MD, Allmer J (2013b) Current limitations for computational analysis of miRNAs in cancer. Pakistan J Clin Biomed Res 1:3–5

Saçar MD, Allmer J (2013c) Data mining for microrna gene prediction: On the impact of class imbalance and feature number for microRNA gene prediction. In: 2013 8th international symposium on health informatics and bioinformatics. IEEE, Ankara, Turkey

Saçar MD, Hamzeiy H, Allmer J (2013) Can MiRBase provide positive data for machine learning for the detection of MiRNA hairpins? J Integr Bioinform 10(2):215. doi:10.2390/biecoll-jib-2013-215

Saçar MD, Bagci C, Allmer J (2014) Computational prediction of microRNAs from *Toxoplasma gondii* potentially regulating the hosts' gene expression. Genomics Proteomics Bioinformatics 12(5):228–238. doi:10.1016/j.gpb.2014.09.002

Skalsky RL, Cullen BR (2010) Viruses, microRNAs, and host interactions. Annu Rev Microbiol 64:123–141. doi:10.1146/annurev.micro.112408.134243

Steffen P, Voss B, Rehmsmeier M, Reeder J, Giegerich R (2006) RNAshapes: an integrated RNA analysis package based on abstract shapes. Bioinformatics 22:500–3. doi:10.1093/bioinformatics/btk010

Thain D, Tannenbaum T, Livny M (2005) Distributed computing in practice: the condor experience. Concurr Comput Pract Exp 17:2–4

Thirugnanam S, Rout N, Gnanasekar M (2013) Possible role of *Toxoplasma gondii* in brain cancer through modulation of host microRNAs. Infect Agent Cancer 8(1):8. doi:10.1186/1750-9378-8-8

van der Burgt A, Fiers MW, Nap JP, van Ham RC (2009) In silico miRNA prediction in metazoan genomes: balancing between sensitivity and specificity. BMC Genomics 10:204. doi:10.1186/1471-2164-10-204

Wang J, Liu X, Jia B, Lu H, Peng S, Piao X, Hou N, Cai P, Yin J, Jiang N, Chen Q (2012) A comparative study of small RNAs in *Toxoplasma gondii* of distinct genotypes. Parasit Vectors 5:186. doi:10.1186/1756-3305-5-186

Xiao J, Li Y, Prandovszky E, Karuppagounder SS, Talbot CC, Jr., Dawson VL, Dawson TM, Yolken RH (2014) MicroRNA-132 dysregulation in *Toxoplasma gondii* infection has implications for dopamine signaling pathway. Neuroscience 268:128–138. doi:10.1016/j.neuroscience.2014.03.015

Xie M, Zhang S, Yu B (2015) microRNA biogenesis, degradation and activity in plants. Cell Mol Life Sci 72 (1):87–99. doi:10.1007/s00018-014-1728-7

Xu Q-S, Liang Y-Z (2001) Monte Carlo cross validation. Chemom Intell Lab Syst 56:1–11. doi:10.1016/S0169-7439(00)00122-2

Xu Y, Zhou X, Zhang W (2008) MicroRNA prediction with a novel ranking algorithm based on random walks. Bioinformatics 24 (13):i50–58. doi:10.1093/bioinformatics/btn175

Xue C, Li F, He T, Liu GP, Li Y, Zhang X (2005) Classification of real and pseudo microRNA precursors using local structure-sequence features and support vector machine. BMC Bioinf 6:310. doi:10.1186/1471-2105-6-310

Yousef M, Jung S, Showe LC, Showe MK (2008) Learning from positive examples when the negative class is undetermined–microRNA gene identification. Algorithms Mol Biol 3:2. doi:10.1186/1748-7188-3-2

Yousef M, Allmer J, Khalifa W (2015) Sequence motif-based one-class classifiers can achieve comparable accuracy to two-class learners for plant microRNA detection. J Biomed Sci Eng 8:684–694. doi:10.4236/jbise.2015.810065

Zeiner GM, Norman KL, Thomson JM, Hammond SM, Boothroyd JC (2010) Toxoplasma gondii infection specifically increases the levels of key host microRNAs. PLoS One 5(1):e8742. doi:10.1371/journal.pone.0008742

Chapter 10
Hypothetical Plant-Mammal Small RNA Communication: Packaging and Stoichiometry

Kenneth W. Witwer

Abstract The dietary RNA hypothesis suggests that intact foreign RNAs from food can enter the ingesting organism and potentially function, for example, to silence endogenous transcripts. Studies of mammalian uptake have focused mostly on short RNA molecules known as microRNAs (miRNAs), or "xenomiRs" to denote their foreign origin. Enthusiasm about absorption and function of plant xenomiRs has been diminished by negative findings and evidence of contamination and experimental design flaws that account for apparently positive results. Nevertheless, some funding groups, regulatory agencies, and scientists remain interested and invested in the topic. Despite the relative lack of accepted evidence for the hypothesis, this interest is likely to continue into the foreseeable future. It may thus be helpful to identify questions that must be answered if, in the future, mammalian dietary RNA uptake and function are to be proven. Here, I examine packaging and stoichiometry considerations for potential plant-mammal RNA communication.

10.1 Introduction

There is current interest in the idea that our diet is a source, not just of nutrition, but of molecular information: food constituents capable of reprogramming our biological systems. What better component of the diet to act in this way than nucleic acids, the DNA and RNA that encode, communicate, and control information in our cells? Could these molecules be absorbed from the diet and distributed in the mammalian body in a functional form? Could this happen even for plant material, in a sort of cross-kingdom communication? Several groups have indeed reported detection in ingesting mammals of dietary RNA from plants (Chin et al. 2016;

K.W. Witwer (✉)
The John Hopkins University School of Medicine, Department of Molecular and Comparative Pathobiology, Department of Neurology, 733 North Broadway, Miller Research Building Rm 829, Baltimore, MD 21205, USA
e-mail: kwitwer1@jhmi.edu

A.L. Leitão and F.J. Enguita (eds.), *Non-coding RNAs and Inter-kingdom Communication*, DOI 10.1007/978-3-319-39496-1_10

Liang et al. 2015; Yang et al. 2015; Zhang et al. 2012a), and even transmission of complete genes (Spisak et al. 2013). However, these findings have been controversial, and some have already been attributed to contamination, experimental error, or artifact (Dickinson et al. 2013; Lusk 2014; Tosar et al. 2014; Witwer 2015; Witwer and Hirschi 2014).

To be sure, the nematode *Caenorhabditis elegans* provides a precedent as the undisputed champion of cross-kingdom RNA-mediated communication through the diet. (Other, non-dietary forms of cross-kingdom RNA communication have also been reported and are not discussed here). Amenable to culture in the lab and the source of much of our knowledge about RNA interference (RNAi) (Fire et al. 1998; Lee et al. 1993). *C. elegans* imports double stranded RNA from the digestive tract (Tabara et al. 1998; Timmons and Fire 1998; Winston et al. 2007). Once inside the cell, the RNA is enzymatically processed into smaller pieces: primary small interfering RNAs (siRNAs) that can function, e.g., as a type of immune system. Like microRNAs (miRNAs), which arise through processing of hairpin precursors, siRNAs are incorporated into Argonaute (AGO) proteins that exert effects as part of an RNA-induced silencing complex (RISC) (Carmell et al. 2002). RNA-dependent RNA polymerases use primary siRNAs to diversify and amplify the signal into an abundance of secondary siRNAs (Parrish and Fire 2001; Tabara et al. 1999). Absorbed RNA can also be efficiently passed from cell to cell throughout the nematode (Hunter et al. 2006; Winston et al. 2002).

In summary, *C. elegans* has evolved well characterized mechanisms that allow it to (1) take up, (2) amplify, and (3) distribute RNAi effectors throughout the body. Somewhat discouragingly for widespread application of dietary RNA cross-kingdom communication, though, similar mechanisms are unknown not only in distant organisms, but also in species within the same genus (Felix 2008). Some other organisms are susceptible to environmental RNAi to a lesser degree (Attasart et al. 2013; Bachman et al. 2013; Whyard et al. 2009; Zhou et al. 2008). Mammals, however, are not known to have any of the necessary mechanisms. Mammals do not use dsRNA as information, but in fact recognize it as a threat, stimulating inflammatory innate immune responses (Alexopoulou et al. 2001; Peisley and Hur 2013). In part for this reason, studies of hypothetical plant RNA uptake in mammals have focused mostly on short, single stranded RNAs. Short regulatory RNA molecules mediate posttranscriptional regulation by guiding repressive proteins to fully (siRNA) or partially (miRNA) complementary recognition elements in target transcripts. They are also perceived to be relatively stable, owing to the tight association of Argonaute (AGO) proteins and mature small RNAs (Elkayam et al. 2012; Schirle and MacRae 2012).

In the first study of dietary transfer with positive results, several rice miRNAs were reportedly found in blood of human donors, and one xenomiR, MIR168a, was said to increase circulating cholesterol in a mouse model (Zhang et al. 2012a). However, another group that was active in the field had found no evidence of RNA uptake in feeding studies of humans, knockout mice, and bees (Snow et al. 2013).

Other studies likewise found no appreciable uptake of plant small RNAs (Baier et al. 2014; Dickinson et al. 2013; Wang et al. 2012; Witwer et al. 2013). Two pivotal reports revealed that the difference between positive and negative studies went beyond methodologic differences. Tosar and colleagues found that plant xenomiR reads in a different dataset from the Zhang et al. group correlated with those in the human study. Since this second dataset used RNA from an aquatic organism not exposed to plants, it was apparent that both studies shared common plant contaminants (Tosar et al. 2014). Dickinson et al. repeated the original mouse feeding study, adding a crucial experimental condition to control for the nutritional insufficiency of raw rice (Dickinson et al. 2013). They found no uptake of MIR168a or downregulation of its purported mammalian target, and showed that the increase in cholesterol in the original study was a starvation response to a nutritionally insufficient diet.

As another example, a fruit feeding experiment with human subjects appeared to show a striking, time-dependent absorption of large quantities of specific plant miRNAs (Liang et al. 2015). However, the best-absorbed and most abundant xenomiR in this study was a sequence specific to monocots (Witwer 2015), in spite of the fact that only dicots (watermelon) were administered to the volunteers. That the monocot miRNA appeared even in qPCR of dicot plant material (Liang et al. 2015) indicates that a pervasive contamination problem, not dietary uptake, was responsible for the results. This recalls a previous finding that the dicot/monocot split reveals contaminant issues in public datasets (Zhang et al. 2012b).

Despite the relative lack of evidence to date for significant uptake or function of plant xenomiRs, the implications of plant-mammal RNA-mediated communication have kept interest in the hypothesis alive. Funding opportunities from the "Extracellular RNA Communication Consortium" of the United States National Institutes of Health have emphasized the possibility. Uses of RNA interference mechanisms in agricultural biotechnology (Sherman et al. 2015) have prompted attention from regulatory agencies (Ramon et al. 2014). For example, the US Environmental Protection Agency convened a special advisory panel in 2014 (http://www.epa.gov/sites/production/files/2015-06/documents/012814min.pdf) while the European Food Safety Authority held a workshop on the issue (http://www.efsa.europa.eu/sites/default/files/corporate_publications/files/705e.pdf). Since this interest is unlikely to abate in the near future, it might be useful to embark on a theoretical excursion to identify features of small RNA regulation that are relevant to postulated plant-mammal RNA communication through the diet. Specifically, I will examine how xenomiRs must be packaged if they are to survive inside or outside the cell. I will then investigate the stoichiometry of small RNA regulation to identify conditions that must be met if a dietary xenomiR is to enter and function in the recipient mammal.

10.2 Packaging Considerations for Plant-Mammal RNA Communication

10.2.1 First of All, Allow no Harm

Any RNA that is to function in dietary cross-kingdom communication must be protected from its environment: first in the plant, then in the alimentary canal, and all the way into the cytoplasm of the end recipient cell. In potential RNA plant-to-mammal cross-kingdom communication through the diet, the degradation process for native RNA begins in the food plant. Degradation is ongoing during tissue death in harvest, transportation, and storage. Processing further accelerates RNA degradation: in highly processed material, most RNA may be degraded before ingestion (Howard et al. 2015; Zhang et al. 2012a). For all food sources, mechanical, chemical, and enzymatic processes expose RNAs as part of digestion. The aqueous environment of the digestive tract and the acidic conditions of the stomach render RNA of any form susceptible to hydrolysis. Furthermore, all forms and structures of RNA are actively degraded by RNases (Silverman et al. 2014; Sorrentino 2010). Even relatively resistant structures including dsRNA are digested by RNases released into the gut, for example from the pancreas or by members of the gut microbiome. Oligonucleotides, like other macromolecules, are thus broken into molecular components that serve as nutrition for the ingesting mammal as well as for gut bacteria.

Do we know of any exceptions? Might certain RNA sequences be resistant to degradation? Two groups have reported that partially degraded ribosomal RNA sequences with mostly GC composition may be exceptionally hardy (Yang et al. 2015; Zhou et al. 2015). GC-mediated formation of complex multimeric structures might contribute to this resistance. Such degradation products are unlikely to serve as regulatory small RNAs, certainly not in multimeric form, and were not found to be associated with AGO (Yang et al. 2015). As a rule, however, RNA on its own is unlikely to survive passage through the digestive tract.

10.2.2 Protection Versus Accessibility

Of course, not all ingested material is digested completely. Some plant cells, for example, protected by thick cell walls, may pass through the ingesting animal without being broken down. This undigested material results in RNA that can be detected in gut contents or excreted fecal matter. However, RNA in undigested material including plant cells is in no position to enter or influence the cells of the ingesting mammal. When reporting that certain RNAs can be found at low abundance in stomach or intestinal contents, it is important to distinguish between cell-associated and cell-free RNA before considering whether RNAs that are present may consequently communicate with the host. Within an undigested plant cell

or in some other large protective complex, RNAs would not have any possibility of interaction with host cells.

10.2.3 AGO and Small RNA

Small RNAs such as miRNAs are often described as being exceptionally stable (e.g., in blood) because they are protected inside extracellular vesicles (EVs) or protein complexes. This statement is accurate but incomplete: it would better read that microRNAs are exceptionally stable as long as they are associated with long-lived proteins, such as AGO, no matter if the complex is found in the cell, in an EV, or free. Most miRNAs in mammalian blood are in free AGO complexes, with a small percentage within EVs (Arroyo et al. 2011; Turchinovich et al. 2011). Since EV-associated miRNA can be amplified despite its relative rarity (Chevillet et al. 2014), RNA is much easier to detect than the AGO that protects it, but this does not mean that free, stable miRNAs exist, even in the EV.

In a plant or a mammal, a small RNA without a protein protector such as AGO has no canonical regulatory function, a point that is obvious since the unprotected small RNA is also rapidly degraded. Usually, when microRNA precursors are processed in the cytoplasm, one strand is preferentially loaded into the AGO protein while the other is discarded. The unused small RNA, once known as the "star-strand" and denoted by '*,' is not associated with protein and is usually rapidly degraded. microRNAs have no special quality that renders them exempt from degradation. Indeed, for most of the non-preferentially incorporated strands, only the most sensitive assays are able to detect their short-lived existence. In contrast, the main mature products are loaded into AGO and thereby protected whether they are actively regulating or not. The association of AGO and a specific miRNA molecule is mutually stabilizing (Elkayam et al. 2012) and can persist through multiple rounds of regulation, or can be exceptionally long-lived in the absence of target regulation (Olejniczak et al. 2013).

10.2.4 No Transfer of miRNA Between AGOs

Once loaded, AGO does not appear to give up its small RNA (Martinez et al. 2002). AGOs are not known to transfer their small RNA cargo, within the same species or across kingdoms. This is an important point and perhaps the most significant consideration for plant-mammal small RNA communication. Take the case of the plant xenomir (xmiR), protected in its plant AGO (xAGO). This complex is stable in the plant and, freed during digestion, may even survive into or through the mammalian digestive tract. Yet it seems that the xmiR-xAGO could act in a mammalian cell only if the complex was to (1) enter the cytoplasm intact and (2) either function directly or incorporate into and function within a mammalian

RISC. Endogenous AGO is recognized by endogenous TNRC6 (GW182 family) proteins to incorporate into functional RISC (Schraivogel et al. 2015); this might not be possible for plant AGOs due to sequence and structural differences. Alternatively, an unknown mechanism of transfer would need to allow the xmiR to relocate from xAGO to endogenous AGO.

10.2.5 Cell Entry: Active Transport or EV Fusion?

For xmiR-xAGO to enter the cell, the complex must be actively transported by a currently unknown transporter or, if the xmiR-xAGO is first inside a lipid entity such as an EV, inserted into the cell cytoplasm after vesicle-membrane fusion. xmiR-xAGO transporters are not known, and very little is known about plant EVs. It seems reasonable to assume that they exist, since most organisms produce them. However, neither their contents nor their possible functions have been interrogated well. Their presence in foodstuffs and survival during digestion have not been widely investigated, either. In several studies, rodents were exposed to biological nanoparticles produced by high-speed blending of plant materials (Mu et al. 2014; Wang et al. 2013); however, such particles are not natural, and it is not known if they are found in ingested food. Also unknown is how plant EVs—native or produced during the digestive process—might fuse with the recipient cell membrane.

On a technical note, when investigating uptake mechanisms, it is important to remember that uptake of lipid dye molecules not necessarily is equivalent to uptake of EV contents. Simply detecting fluorescence in or near a host cell is insufficient to prove passage. Lipid dyes can move promiscuously from the EV membrane to the cell membrane without fusion and transfer of EV contents. Some lipid dyes also form EV-sized structures that transfer dye to cell membranes [see Brunger et al. 2015 and references therein]. Thus, transfer of signal within (host EV to cell) or across kingdoms (e.g., from bacteria to host gut epithelium), cannot be assumed to represent actual fusion events without additional evidence. Similarly, labels released from RNA molecules may engender erroneous interpretations of RNA uptake.

One might posit that the xmiR-xAGO, free or in an EV, does not necessarily have to enter cells of the gut. Perhaps transcytosis or a similar process could deliver xmiR-AGO across the barriers of the gut and into the bloodstream. Alternatively, immune cells could endocytose complexes and move them elsewhere in the body before releasing them again. However, these transport mechanisms, while theoretically possible, would only delay the dilemma, postponing eventual cellular uptake. Presence of xmiR-xAGO in the cytoplasm is absolutely essential for communication via miRNA-mediated regulation, as has been proposed, to occur.

In conclusion, in the digestive tract, xenomiRs in cells and other large undigested matter are relatively protected from degradation but irrelevant to possible

cross-kingdom communication. Free RNAs are rapidly degraded, and there are no established mechanisms for their uptake into mammalian cells. Even if there were, free xenomiRs are essentially "dead": so short-lived in or outside the cell as to be non-existent for canonical regulatory purposes, and unable to incorporate significantly into AGO. Some dietary plant small RNAs may be found in EV-like structures that may exist in the plant or may be generated during breakup of cells in food matter. The extent to which these vesicles survive digestion and whether and how they might be taken up by gut epithelium is not clear. The functional unit of miRNA-mediated regulation, the miRNA:AGO complex, would have to enter the mammalian cell through vesicle fusion or active transport of free complexes, neither of which has been reported. Transfer of miRNAs between AGOs and function of plant AGO in mammalian RISC have not been reported. To date, all studies of plant xenomiRs have focused on mature miRNAs only, and have not considered the necessary xmiR-xAGO complexes and the resulting dilemmas of trans-membrane delivery and cross-kingdom protein interactions.

10.3 Experimental Recommendations

If xenomiRs are to be shown to participate in plant-mammal communication, the following steps related to functional packaging must be taken:

(1) Demonstrate and quantitate plant xenomiR survival in the complex enzymatic environment of digestion in a potentially functional form (probably as xmiR-xAGO complexes);
(2) Show that surviving xenomiR complexes are cell-free (potentially accessible for active transport into mammalian cells);
(3) Discover and unravel molecular mechanisms of active xmiR-xAGO transport, plant vesicle fusion with mammalian gut cells, or other, currently unknown means of uptake;
(4) Provide evidence that imported xmiR-xAGO is present in the cytoplasm of mammalian cells (not xmiRs alone, not just associated with the cell surface, and not in the endosomal system).
(5) Demonstrate at least one of the following:
    a. that xmiR-xAGO can replace endogenous miR-AGO in mammalian RISC, or
    b. that xmiR-xAGO can tranfser its miRNA to an endogenous AGO
(6) Show that xmiR-xAGO-RISC or xmiR-AGO-RISC is functional.

10.4 Stoichiometric Considerations for Plant-Mammal RNA Communication

Having examined the challenges of xmiR-xAGO packaging for xenomiR communication, let us now turn to the problem of stoichiometry. Some of these considerations will be packaging-independent, so putting packaging aside for the moment, might expected exposure levels be consistent with xenomiR function in ingesting mammals?

10.4.1 Roughly: 100–1000 Copies per Cell Needed for Regulation

Roughly, the more copies of a small RNA in the cell, the more likely that specific RNA is engaged in canonical silencing. Although there is some danger in establishing absolute cutoffs, 100–1000 copies of a small RNA per cell may be sufficient to achieve measurable regulation of confirmed target mRNAs (Brown et al. 2007; Mullokandov et al. 2012). Similarly, in sequencing experiments, 100–1000 reads per million seems to be a reasonable lower cutoff for predicting potential canonical function for a detected small RNA. The necessary copy number may be much higher, depending on the number and accessibility of target sites and the ratio of target to small RNA. Adding or taking away copies of a small RNA may produce detectable changes in regulated targets only within a narrow concentration range, likely different for each miRNA and cellular environment (Mayya and Duchaine 2015). Because of these complexities, let us consider the simplified case of the xenomiR that is present at baseline at zero copies per cell and that is taken up after exposure. How many copies of the xenomiR are needed to achieve 1000 copies per cell?

A recent estimate of cell count in the human body suggested that a 70 kg human consists of approximately 40 trillion (4×10^{13}) cells (Bianconi et al. 2013). This estimate should probably be seen as a lower bound, since not all cells in all organs could be taken into account. If they could be instantly and evenly distributed throughout the body, at least 4×10^{16} xenomiRs would be required to achieve 1000 copies per cell. What if the miRNA could target endothelial cells only? Approximately 2.5×10^{15} miRNAs would be needed (2.5×10^{12} endothelial cells). In comparison, the liver has been estimated to contain about 400 billion cells (4×10^{11}). A hypothetical, 100 % liver-targeted miRNA would require 4×10^{14} copies to reach 1000/cell.

Cooked rice has been reported to retain about 0.95 fmol/g of one of the most abundant plant miRNAs, MIR168a, or just under 600 million copies per gram (Zhang et al. 2012a). In other plant foods, it was less abundant or even

undetectable. Following the above estimate of cell counts, molecules of MIR168a equivalent to 1000 copies per cell of an adult human could be found in 67 kg of cooked rice. For endothelial cells, 4 kg would be needed, and for liver, 0.7 kg. These estimates resemble those of Petrick et al. and Snow et al., who similarly calculated how much food would contain regulation-relevant levels of xenomiRs (Petrick et al. 2013; Snow et al. 2013). Although it would clearly be impossible to obtain enough xenomiRs through the diet to have systemic effects, could our adult human hope to obtain such levels of xenomiRs in specific organs (through unknown targeting), just by eating piles of rice? Most certainly not, since our oversimplified estimates rely on numerous false assumptions.

10.4.2 Degradation and Digestion

The first incorrect assumption is that all RNA survives the trip through the alimentary canal. As we have seen above in the packaging section, most dietary RNA is degraded, including essentially all RNAs that are exposed. Mlotshwa et al. were unable to detect two of three miRNAs in the mouse gut (the third was close to background), despite repeated delivery of approximately 200 trillion copies each, and even though part of the alimentary canal was bypassed by gavage (Mlotshwa et al. 2015). Another consideration, as mentioned, is that much of the plant RNA remaining in the gut after a plant meal will be found in undigested cells and is thus unavailable for uptake.

10.4.3 Crossing into the Circulation

A second faulty assumption is that all RNA in the gut reaches the circulation. Transport mechanisms for moving RNA from the gut lumen to the circulation have yet to be fully described but, as we have seen, must involve more than uncomplexed RNA. Transport avenues would have to accommodate xmiR-xAGO on its own or within extracellular vesicles. To date, there is some evidence for EV-mediated transfer of small RNAs across simplified barriers, but these in vitro experiments have used nonphysiologic conditions. For example, using a monolayer model (Caco-2 cells in transwell), dietary EV miRNA transport was estimated by adding approximately 200 μg of milk EVs to cells seeded at 9000 cells/well (Wolf et al. 2015). Using Sverdlov's estimate of 100 nm EVs per μg (Sverdlov 2012), these experiments involved approximately 400 million EVs per cell, or 400 billion EVs per 75 μl well, fully one-tenth the estimated number of EVs in the average human circulation.

10.4.4 Organ Targeting

Even so, let us suppose, generously, that 1 % of dietary small RNA is both protected from digestion and finds its way through unknown mechanisms into the circulation. Now, even for the 100 % liver-targeting scenario above, 70 kg of rice would contain enough xenomiR to target liver cells. This introduces our third obvious faulty assumption: xenomiRs that enter the circulation are unlikely to be targeted 100 % to one organ. They will encounter phagocytic cells, a wide variety of other blood cells that might wish to sample their environment, and endothelial cells. Some might leave the circulation to enter tissues. Free (or freed) RNAs will be rapidly cleared by the kidneys. Other particles are removed by the liver; extracellular vesicles have a half-life on the order of an hour or so. We cannot suppose that all circulating xenomiRs make their way immediately to a single organ, such as the liver.

10.4.5 Entry into the Cytoplasm

The next unrealistic assumption is that miRNAs in the extracellular environment can enter the target cell cytoplasm at 100 % efficiency and regulate endogenous target RNAs directly. As discussed above in the packaging section, there is currently no knowledge of how xenomiRs could be freed from the endosomal system; no compelling evidence that mature xenomiRs, contained as they must be in xAGO, could be incorporated into endogenous RISC and exert influence; and no reports that mature xenomiRs, alternatively, could transfer between xAGO and AGO. Nevertheless, we know from in vitro experiments that even if these substantial barriers are sidestepped by artificial delivery, vastly excess copies of small RNA are needed to achieve measurable canonical small RNA-mediated regulation.

Several key differences separate the “best case scenario” of artificial RNA delivery from hypothetical biological uptake. First, the RNA enters the cell via processes (electroporation) or reagents (such as lipids that disrupt/fuse with cell membranes) that were developed for efficient delivery directly into the cytoplasm as compared with conjectured uptake of EVs or xmiR-xAGO. Second, the introduced RNA is uncomplexed, “naked” immature RNA, handled under artificial RNase-free conditions and without a need for xAGO-to-AGO transfer. Third, the RNA is either in a pre-miRNA “mimic” form (i.e., dsRNA or annealed strands that are processed and loaded into native AGO) or transcribed from a plasmid and then processed and loaded. That is, unlike the mature xenomiRs proposed in the literature, these miRNAs are endogenous in every way except their sequence and any modifications introduced during synthesis.

Despite these favorable, manipulable conditions that hardly resemble the realities of xmiR-xAGO from the diet, in vitro experiments with small RNAs and optimized target reporters typically employ miRNA or siRNA transfections in the

10 to 100 nM range, equivalent in an average experiment to tens of millions of molecules per cell, in order to have a measurable effect, albeit often within the natural range of variation for endogenous genes (Seitz 2009). [For an example, see Zhang et al., where 120 million molecules per cell (20 pmol per 100,000 cells) were transfected in xenomiR experiments (Zhang et al. 2012a)] When delivering small RNA from outside the cell, excess molecules are needed, thousands- to millions-fold of the required intracellular copy number threshold. Adding to this the uncertainties of xmiR-xAGO uptake and function in exposure to plant material, not synthetic molecules in liposomes, a very large and currently incalculable excess would be needed to assure xenomiR function in the cell. Clearly, we have long since passed the possibility of biological relevance in most systems as well as the point of calculability. We can be confident only that plant xenomiR function in mammals is implausible in the context of known mechanisms, dietary intake, and stoichiometry: systemically or even if restricted to organs such as the liver.

10.4.6 Active Versus Inactive miR-AGO in Somatic Cells

There is, however, another consideration: is the recipient cell susceptible to small RNA-mediated regulation? Our collective knowledge of mammalian small RNA regulation has been gleaned mostly from experiments in transformed or proliferating, activated cells in culture, where an AGO-incorporated small RNA can be assumed to be active. However, recent findings suggest that this condition is not universal. In many, perhaps most, cells of the adult mammal, miR-AGO complexes are in low molecular weight form, not involved in canonical regulation (La Rocca et al. 2015). Only in tissues with high metabolic activity or active proliferation or differentiation were high molecular weight complexes detected, consisting of miR-AGO associated with other RISC components and target transcripts. We have found that even in isolated PBMCs, the majority of miR-AGO is low molecular weight (unpublished data). The implications are important: in many cell types, even if apparently regulation-relevant numbers of xenomiRs were to be found and incorporated into AGO, they would have no regulatory effect unless the recipient cell were simultaneously stimulated to shift miR-AGO into high molecular weight, active fractions.

In conclusion, hundreds to thousands of molecules of specific small RNAs are needed to achieve regulation in the cell. Combined with conservative estimates of degradation, this consideration means that mammals are not able to ingest enough dietary plant material to supply sufficient quantities of even high-abundance xenomiRs, even if it were possible to target these RNAs to specific organs. Adding considerably to the unlikelihood, in vitro experiments with optimized delivery systems and duplexed small RNA precursors show that even under ideal conditions, thousands to millions-fold RNA excess may be needed to achieve regulation. The challenges of packaging and delivery must also be considered, as well as the ratio of active to inactive miR-AGO.

10.5 Recommended Experiments

Allowing for the possibility of unknown biological mechanisms, the stoichiometry and functional aspects of suspected xenomiR-mammal communication could be explored further through the following steps:

(1) Quantitate xmiR-xAGO (or other currently unknown forms of protected xenomiR) in the food source, also accounting for any processing.
(2) Characterize digestion: how many *cell-free* copies of the protected xenomiR of interest remain at relevant locations in the alimentary tract (and beyond, if mechanisms of gut absorption and distribution have been found)?
(3) Identify the target population of cells, most likely a very small population in or with access to the alimentary canal, such that the number of cells and surviving, cell-free xenomiRs might be consistent with stoichiometric considerations.
(4) Show overwhelming concentration of the small RNA of interest within the cytoplasm of the identified cell population relative to others, with copy numbers consistent with regulatory effects. It is key to show that the measured RNA is in the cytoplasm, not just associated with the outside of the cell or the endosomal system.
(5) Confirm that these rare target cells contain high molecular weight RISC under native conditions. Since the method of La Rocca et al. has not been optimized for such small numbers of cells, alternative approaches such as imaging of co-localized molecules may be needed.
(6) Validate direct canonical small RNA-mediated regulation in the target cells.

References

Alexopoulou L, Holt AC, Medzhitov R, Flavell RA (2001) Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3. Nature 413(6857):732–738. doi:10.1038/35099560

Arroyo JD, Chevillet JR, Kroh EM, Ruf IK, Pritchard CC, Gibson DF, Mitchell PS, Bennett CF, Pogosova-Agadjanyan EL, Stirewalt DL, Tait JF, Tewari M (2011) Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. Proc Natl Acad Sci USA 108(12):5003–5008. doi:10.1073/pnas.1019055108

Attasart P, Namramoon O, Kongphom U, Chimwai C, Panyim S (2013) Ingestion of bacteria expressing dsRNA triggers specific RNA silencing in shrimp. Virus Res 171(1):252–256. doi:10.1016/j.virusres.2012.11.012

Bachman PM, Bolognesi R, Moar WJ, Mueller GM, Paradise MS, Ramaseshadri P, Tan J, Uffman JP, Warren J, Wiggins BE, Levine SL (2013) Characterization of the spectrum of insecticidal activity of a double-stranded RNA with targeted activity against Western Corn Rootworm (*Diabrotica virgifera* virgifera LeConte). Trans Res 22(6):1207–1222. doi:10.1007/s11248-013-9716-5

Baier SR, Nguyen C, Xie F, Wood JR, Zempleni J (2014) MicroRNAs are absorbed in biologically meaningful amounts from nutritionally relevant doses of cow milk and affect gene expression in peripheral blood mononuclear cells, HEK-293 kidney cell cultures, and mouse livers. J Nutr 144(10):1495–1500. doi:10.3945/jn.114.196436

Bianconi E, Piovesan A, Facchin F, Beraudi A, Casadei R, Frabetti F, Vitale L, Pelleri MC, Tassani S, Piva F, Perez-Amodio S, Strippoli P, Canaider S (2013) An estimation of the number of cells in the human body. Ann Hum Biol 40(6):463–471. doi:10.3109/03014460.2013.807878

Brown BD, Gentner B, Cantore A, Colleoni S, Amendola M, Zingale A, Baccarini A, Lazzari G, Galli C, Naldini L (2007) Endogenous microRNA can be broadly exploited to regulate transgene expression according to tissue, lineage and differentiation state. Nat Biotechnol 25(12):1457–1467. doi:10.1038/nbt1372

Brunger AT, Cipriano DJ, Diao J (2015) Towards reconstitution of membrane fusion mediated by SNAREs and other synaptic proteins. Crit Rev Biochem Mol Biol 50(3):231–241. doi:10.3109/10409238.2015.1023252

Carmell MA, Xuan Z, Zhang MQ, Hannon GJ (2002) The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. Genes Dev 16(21):2733–2742. doi:10.1101/gad.1026102

Chevillet JR, Kang Q, Ruf IK, Briggs HA, Vojtech LN, Hughes SM, Cheng HH, Arroyo JD, Meredith EK, Gallichotte EN, Pogosova-Agadjanyan EL, Morrissey C, Stirewalt DL, Hladik F, Yu EY, Higano CS, Tewari M (2014) Quantitative and stoichiometric analysis of the microRNA content of exosomes. Proc Natl Acad Sci USA 111(41):14888–14893. doi:10.1073/pnas.1408301111

Chin AR, Fong MY, Somlo G, Wu J, Swiderski P, Wu X, Wang SE (2016) Cross-kingdom inhibition of breast cancer growth by plant miR159. Cell Res 26(2):217–228. doi:10.1038/cr.2016.13

Dickinson B, Zhang Y, Petrick JS, Heck G, Ivashuta S, Marshall WS (2013) Lack of detectable oral bioavailability of plant microRNAs after feeding in mice. Nat Biotechnol 31(11):965–967. doi:10.1038/nbt.2737

Elkayam E, Kuhn CD, Tocilj A, Haase AD, Greene EM, Hannon GJ, Joshua-Tor L (2012) The structure of human argonaute-2 in complex with miR-20a. Cell 150(1):100–110. doi:10.1016/j.cell.2012.05.017

Felix MA (2008) RNA interference in nematodes and the chance that favored Sydney Brenner. J Biol 7(9):34. doi:10.1186/jbiol97

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature 391(6669):806–811. doi:10.1038/35888

Howard KM, Jati Kusuma R, Baier SR, Friemel T, Markham L, Vanamala J, Zempleni J (2015) Loss of miRNAs during processing and storage of cow's (Bos taurus) milk. J Agric Food Chem 63(2):588–592. doi:10.1021/jf505526w

Hunter CP, Winston WM, Molodowitch C, Feinberg EH, Shih J, Sutherlin M, Wright AJ, Fitzgerald MC (2006) Systemic RNAi in *Caenorhabditis elegans*. Cold Spring Harb Symp Quant Biol 71:95–100. doi:10.1101/sqb.2006.71.060

La Rocca G, Olejniczak SH, Gonzalez AJ, Briskin D, Vidigal JA, Spraggon L, DeMatteo RG, Radler MR, Lindsten T, Ventura A, Tuschl T, Leslie CS, Thompson CB (2015) In vivo, Argonaute-bound microRNAs exist predominantly in a reservoir of low molecular weight complexes not associated with mRNA. Proc Natl Acad Sci USA 112(3):767–772. doi:10.1073/pnas.1424217112

Lee RC, Feinbaum RL, Ambros V (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75(5):843–854

Liang H, Zhang S, Fu Z, Wang Y, Wang N, Liu Y, Zhao C, Wu J, Hu Y, Zhang J, Chen X, Zen K, Zhang CY (2015) Effective detection and quantification of dietetically absorbed plant microRNAs in human plasma. J Nutr Biochem 26(5):505–512. doi:10.1016/j.jnutbio.2014.12.002

Lusk RW (2014) Diverse and widespread contamination evident in the unmapped depths of high throughput sequencing data. PLoS ONE 9(10):e110808. doi:10.1371/journal.pone.0110808

Martinez J, Patkaniowska A, Urlaub H, Luhrmann R, Tuschl T (2002) Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. Cell 110(5):563–574

Mayya VK, Duchaine TF (2015) On the availability of microRNA-induced silencing complexes, saturation of microRNA-binding sites and stoichiometry. Nucleic Acids Res 43(15):7556–7565. doi:10.1093/nar/gkv720

Mlotshwa S, Pruss GJ, MacArthur JL, Endres MW, Davis C, Hofseth LJ, Pena MM, Vance V (2015) A novel chemopreventive strategy based on therapeutic microRNAs produced in plants. Cell Res 25(4):521–524. doi:10.1038/cr.2015.25

Mu J, Zhuang X, Wang Q, Jiang H, Deng ZB, Wang B, Zhang L, Kakar S, Jun Y, Miller D, Zhang HG (2014) Interspecies communication between plant and mouse gut host cells through edible plant derived exosome-like nanoparticles. Mol Nutr Food Res 58(7):1561–1573. doi:10.1002/mnfr.201300729

Mullokandov G, Baccarini A, Ruzo A, Jayaprakash AD, Tung N, Israelow B, Evans MJ, Sachidanandam R, Brown BD (2012) High-throughput assessment of microRNA activity and function using microRNA sensor and decoy libraries. Nat Methods 9(8):840–846. doi:10.1038/nmeth.2078

Olejniczak SH, La Rocca G, Gruber JJ, Thompson CB (2013) Long-lived microRNA-Argonaute complexes in quiescent cells can be activated to regulate mitogenic responses. Proc Natl Acad Sci USA 110(1):157–162. doi:10.1073/pnas.1219958110

Parrish S, Fire A (2001) Distinct roles for RDE-1 and RDE-4 during RNA interference in *Caenorhabditis elegans*. RNA 7(10):1397–1402

Peisley A, Hur S (2013) Multi-level regulation of cellular recognition of viral dsRNA. Cell Mol Life Sci 70(11):1949–1963. doi:10.1007/s00018-012-1149-4

Petrick JS, Brower-Toland B, Jackson AL, Kier LD (2013) Safety assessment of food and feed from biotechnology-derived crops employing RNA-mediated gene regulation to achieve desired traits: a scientific review. Regul Toxicol Pharmacol 66(2):167–176. doi:10.1016/j.yrtph.2013.03.008

Ramon M, Devos Y, Lanzoni A, Liu Y, Gomes A, Gennaro A, Waigmann E (2014) RNAi-based GM plants: food for thought for risk assessors. Plant Biotechnol J 12(9):1271–1273. doi:10.1111/pbi.12305

Schirle NT, MacRae IJ (2012) The crystal structure of human Argonaute2. Science 336 (6084):1037–1040. doi:10.1126/science.1221551

Schraivogel D, Schindler SG, Danner J, Kremmer E, Pfaff J, Hannus S, Depping R, Meister G (2015) Importin-beta facilitates nuclear import of human GW proteins and balances cytoplasmic gene silencing protein levels. Nucleic Acids Res 43(15):7447–7461. doi:10.1093/nar/gkv705

Seitz H (2009) Redefining microRNA targets. Curr Biol 19(10):870–873. doi:10.1016/j.cub.2009.03.059

Sherman JH, Munyikwa T, Chan SY, Petrick JS, Witwer KW, Choudhuri S (2015) RNAi technologies in agricultural biotechnology: the Toxicology forum 40th annual summer meeting. Regul Toxicol Pharmacol 73(2):671–680. doi:10.1016/j.yrtph.2015.09.001

Silverman IM, Li F, Alexander A, Goff L, Trapnell C, Rinn JL, Gregory BD (2014) RNase-mediated protein footprint sequencing reveals protein-binding sites throughout the human transcriptome. Genome Biol 15(1):R3. doi:10.1186/gb-2014-15-1-r3

Snow JW, Hale AE, Isaacs SK, Baggish AL, Chan SY (2013) Ineffective delivery of diet-derived microRNAs to recipient animal organisms. RNA Biol 10(7):1107–1116. doi:10.4161/rna.24909

Sorrentino S (2010) The eight human "canonical" ribonucleases: molecular diversity, catalytic properties, and special biological actions of the enzyme proteins. FEBS Lett 584(11):2194–2200. doi:10.1016/j.febslet.2010.04.018

Spisak S, Solymosi N, Ittzes P, Bodor A, Kondor D, Vattay G, Bartak BK, Sipos F, Galamb O, Tulassay Z, Szallasi Z, Rasmussen S, Sicheritz-Ponten T, Brunak S, Molnar B, Csabai I (2013) Complete genes may pass from food to human blood. PLoS ONE 8(7):e69805. doi:10.1371/journal.pone.0069805

Sverdlov ED (2012) Amedeo Avogadro's cry: what is 1 microg of exosomes? BioEssays 34 (10):873–875. doi:10.1002/bies.201200045

Tabara H, Grishok A, Mello CC (1998) RNAi in *C. elegans*: soaking in the genome sequence. Science 282(5388):430–431

Tabara H, Sarkissian M, Kelly WG, Fleenor J, Grishok A, Timmons L, Fire A, Mello CC (1999) The rde-1 gene, RNA interference, and transposon silencing in *C. elegans*. Cell 99(2):123–132

Timmons L, Fire A (1998) Specific interference by ingested dsRNA. Nature 395(6705):854. doi:10.1038/27579

Tosar JP, Rovira C, Naya H, Cayota A (2014) Mining of public sequencing databases supports a non-dietary origin for putative foreign miRNAs: underestimated effects of contamination in NGS. RNA 20(6):754–757. doi:10.1261/rna.044263.114

Turchinovich A, Weiz L, Langheinz A, Burwinkel B (2011) Characterization of extracellular circulating microRNA. Nucleic Acids Res 39(16):7223–7233. doi:10.1093/nar/gkr254

Wang K, Li H, Yuan Y, Etheridge A, Zhou Y, Huang D, Wilmes P, Galas D (2012) The complex exogenous RNA spectra in human plasma: an interface with human gut biota? PLoS ONE 7 (12):e51009. doi:10.1371/journal.pone.0051009

Wang Q, Zhuang X, Mu J, Deng ZB, Jiang H, Zhang L, Xiang X, Wang B, Yan J, Miller D, Zhang HG (2013) Delivery of therapeutic agents by nanoparticles made of grapefruit-derived lipids. Nat Commun 4:1867. doi:10.1038/ncomms2886

Whyard S, Singh AD, Wong S (2009) Ingested double-stranded RNAs can act as species-specific insecticides. Insect Biochem Mol Biol 39(11):824–832. doi:10.1016/j.ibmb.2009.09.007

Winston WM, Molodowitch C, Hunter CP (2002) Systemic RNAi in C. elegans requires the putative transmembrane protein SID-1. Science 295(5564):2456–2459. doi:10.1126/science.1068836

Winston WM, Sutherlin M, Wright AJ, Feinberg EH, Hunter CP (2007) Caenorhabditis elegans SID-2 is required for environmental RNA interference. Proc Natl Acad Sci USA 104 (25):10565–10570. doi:10.1073/pnas.0611282104

Witwer KW (2015) Contamination or artifacts may explain reports of plant miRNAs in humans. J Nutr Biochem 26(12):1685. doi:10.1016/j.jnutbio.2015.09.004

Witwer KW, Hirschi KD (2014) Transfer and functional consequences of dietary microRNAs in vertebrates: concepts in search of corroboration: negative results challenge the hypothesis that dietary xenomiRs cross the gut and regulate genes in ingesting vertebrates, but important questions persist. BioEssays 36(4):394–406. doi:10.1002/bies.201300150

Witwer KW, McAlexander MA, Queen SE, Adams RJ (2013) Real-time quantitative PCR and droplet digital PCR for plant miRNAs in mammalian blood provide little evidence for general uptake of dietary miRNAs: limited evidence for general uptake of dietary plant xenomiRs. RNA Biol 10(7):1080–1086. doi:10.4161/rna.25246

Wolf T, Baier SR, Zempleni J (2015) The intestinal transport of bovine milk exosomes is mediated by endocytosis in human colon carcinoma Caco-2 cells and rat small intestinal IEC-6 cells. J Nutr 145(10):2201–2206. doi:10.3945/jn.115.218586

Yang J, Farmer LM, Agyekum AA, Elbaz-Younes I, Hirschi KD (2015) Detection of an abundant plant-based small RNA in healthy consumers. PLoS ONE 10(9):e0137516. doi:10.1371/journal.pone.0137516

Zhang L, Hou D, Chen X, Li D, Zhu L, Zhang Y, Li J, Bian Z, Liang X, Cai X, Yin Y, Wang C, Zhang T, Zhu D, Zhang D, Xu J, Chen Q, Ba Y, Liu J, Wang Q, Chen J, Wang J, Wang M, Zhang Q, Zhang J, Zen K, Zhang CY (2012a) Exogenous plant MIR168a specifically targets mammalian LDLRAP1: evidence of cross-kingdom regulation by microRNA. Cell Res 22 (1):107–126. doi:10.1038/cr.2011.158

Zhang Y, Wiggins BE, Lawrence C, Petrick J, Ivashuta S, Heck G (2012b) Analysis of plant-derived miRNAs in animal small RNA datasets. BMC Genom 13:381. doi:10.1186/1471-2164-13-381

Zhou X, Wheeler MM, Oi FM, Scharf ME (2008) RNA interference in the termite *Reticulitermes flavipes* through ingestion of double-stranded RNA. Insect Biochem Mol Biol 38(8):805–815. doi:10.1016/j.ibmb.2008.05.005

Zhou Z, Li X, Liu J, Dong L, Chen Q, Kong H, Zhang Q, Qi X, Hou D, Zhang L, Zhang G, Liu Y, Zhang Y, Li J, Wang J, Chen X, Wang H, Zhang J, Chen H, Zen K, Zhang CY (2015) Honeysuckle-encoded atypical microRNA2911 directly targets influenza A viruses. Cell Res 25(1):39–49. doi:10.1038/cr.2014.130

Chapter 11
Different Types of Small RNAs in Protozoa

Ling-Ling Zheng, Jian-You Liao, Yan-Zi Wen, Geoff Hide, Liang-Hu Qu and Zhao-Rong Lun

Abstract The origin and evolution of small RNA molecules is a long-standing mystery. The protozoa represent the earliest eukaryotes and encompass rich diverse genetic resources. However, there are many contradictions and disputes prevalent in studies on protozoan microRNAs. In this review, we summarize the research on the analysis of small RNA transcriptome data on two representative protozoans, *Trypanosoma brucei* and *Giardia lamblia*, as generated by high throughput sequencing. The results show that these protozoans do not have canonical miRNAs. Unexpectedly, there are many types of endogenous small interfering RNAs (endo-siRNAs). In addition, stress induced tRNA derived small RNAs (sitRNAs) were observed in different life cycle stages of these parasites. In total, there are six kinds of sitRNAs in *G. lamblia*. These small RNAs are then found to be involved in the differentiation processes of primitive eukaryotes. These results showed that systematic analysis of small RNAs in the protozoa revealed a more complex picture than previously thought.

L.-L. Zheng · Y.-Z. Wen · L.-H. Qu (✉) · Z.-R. Lun (✉)
Key Laboratory of Gene Engineering of the Ministry of Education, State Key Laboratory of Biocontrol, School of Life Sciences, Sun Yat-Sen University, Guangzhou 510275, People's Republic of China
e-mail: lssqlh@mail.sysu.edu.cn

Z.-R. Lun
e-mail: lsslzr@mail.sysu.edu.cn

J.-Y. Liao
Guangdong Provincial Key Laboratory of Malignant Tumor Epigenetics and Gene Regulation, Sun Yat-Sen Memorial Hospital, Research Center of Medicine, Sun Yat-Sen University, Guangzhou 510120, People's Republic of China

Y.-Z. Wen · Z.-R. Lun
Key Laboratory of Tropical Diseases and Control of the Ministry of Education, Zhongshan School of Medicine, Sun Yat-Sen University, Guangzhou 510080, People's Republic of China

G. Hide · Z.-R. Lun
Ecosystems and Environment Research Centre/Biomedical Research Centre, School of Environment and Life Sciences, University of Salford, Salford M5 4WT, UK

A.L. Leitão and F.J. Enguita (eds.), *Non-coding RNAs and Inter-kingdom Communication*, DOI 10.1007/978-3-319-39496-1_11

11.1 Introduction to Small RNAs in Protozoa

Small RNAs are defined by their length (20–30 nucleotides) and play important roles in post-transcriptional regulation (Ha and Kim 2014). There are two main types of well-studied small RNAs: miRNAs (microRNAs) and endo-siRNAs (endogenous small interfering RNAs). Although, these two types of small RNAs are all involved in RNAi (RNA interference) pathways, their source, processing mechanisms and functions are very distinct.

MicroRNAs have been well studied in the past two decades since their first identification in *Caenorhabditis elegens*. They were later found in a diverse range of animals and these microRNAs appear to be widely expressed in animal somatic cells. They probably drive many of the epigenetic developmental processes that transform embryonic stem cells to terminally differentiated cells. Different tissues or cell lines contain their own specific profile of miRNAs, which regulate the expression of genes that are essential for cell-fate determination in cell differentiation (Sood et al. 2006; Li and Belmonte 2015) and hematopoietic development (Aravin and Hannon 2008). In the genome, microRNAs are self-governed genes which exist in inter-gene and intron regions. In animals, microRNA genes are typically transcribed by RNA polymerase II into long precursors pri-miRNAs (Li and Rana 2014). They are then processed by the Drosha enzyme into stem loop structure pre-miRNAs in the nucleus. The pre-miRNAs are then transported into cytoplasm, where pre-miRNAs are cleaved by the protein Dicer to generate a pair of 18–22 nt small RNAs—mature miRNA and miRNA star. The mature miRNA are loaded into the Argonaute complex and guide the complex to the 3′ UTR of mRNAs through base pair complementarity in the seed region (2–8 nt) of the microRNA. The expression of protein-coding genes are then suppressed at the post-transcriptional level. In general, the miRNA star is degraded in miRNA biogenesis, however, in some cases, these miRNA stars can also target mRNAs and have biological functions in development of diseases such as cancer (Yang et al. 2013; Shan et al. 2013; Wu and Arora 2014).

Endo-siRNAs are derived from long double-strand RNAs (dsRNAs), and then processed into 21–26 nt small RNA clusters. These genes were first found in *Caenorhabditis elegans* and can silence the expression of genes by means of perfect complementarity with the transcripts (Wightman et al. 1993). Endo-siRNAs are mainly derived from Transponsable Elements (TEs) and are believed to be the main mechanism for retrotransposon element silencing. There are many types of endo-siRNAs in plants (Nilsen 2008), however, it is difficult to ascertain their characteristics in animals (Okamura and Lai 2008). The pioneer work in the investigation of endo-siRNAs was performed by two groups and both of them identified a substantial number of endo-siRNAs in mouse oocytes (Tam et al. 2008; Watanabe et al. 2008). A range of studies, conducted at a similar time, used deep-sequencing technology to reveal that there were also an abundant repertoire of endo-siRNAs in *Drosophila melanogaster* (Ghildiyal et al. 2008; Okamura et al. 2008a, b; Kawamura et al. 2008; Czech et al. 2008; Chung et al. 2008).

MicroRNAs have been discovered in sponges and in *Chlamydomonas reinhardtii*, which are the lowest multicellular animals and plants, respectively. But homologous sequences and the characteristic microRNAs were not found between these two taxa (Wheeler et al. 2009). This suggested that microRNAs in plants and animals seem to have originated independently from different ancestors. A few studies have attempted to analyze the small RNAs in protozoans such as trypanosomes (Wen et al. 2011; Mallick et al. 2008; Tschudi et al. 2012a; Zheng et al. 2013), trichomonads (Chen et al. 2009; Lin et al. 2009; Huang et al. 2012), giardias (Chen et al. 2009; Liao et al. 2014; Saraiya and Wang 2008; Huang et al. 2012), amoeba (De et al. 2006) and toxoplasma (Braun et al. 2010). Some of them have reported that microRNAs demonstrate homology with those found in higher species. This implied that microRNAs might have originated from a common ancestor in the eukaryotes. Later other studies questioned the microRNAs identified in protists. Tarver et al. used critical criteria to re-evaluate these microRNAs and concluded that most of the identified microRNAs in protists were not bona fide microRNA (Tarver et al. 2012). These results bring more questions about the origin of microRNAs. Do microRNAs exist in ancient eukaryotes? Whether the features of these so called "microRNAs" are in common with/or specific from those that have been demonstrated in the plants and animals? If the molecules previously reported are not microRNAs, do they perform other specific functions? To answer these questions, a systematic analysis of small RNAs in protists is required. However, it is difficult work because of the following problems.

11.2 Current Problems Associated with the Identification of Small RNAs, and Particularly Authentic microRNAs, from Protozoan Organisms

There are at least four difficulties in small RNA identification in protozoan species:

1. The protozoans are a very diverse group of unicellular eukaryotic organisms that may well have had extensive polyphyletic origins (Cavalier-Smith 2003). The characteristics of these organisms are very broad, some are animal-like, some are plant-like and others also like fungi. Thus, there are a wide range of potential model protozoa organisms but may not be representative of all (Montagnes et al. 2012). Therefore, there are no common characteristics of small RNAs in protozoan organisms which makes it very difficult to investigate them systematically.
2. Most of the parasitic or symbiotic protozoan species have complicated life-cycles, due to the fact that they can infect or inhabit multiple types of hosts (including a wide range of plants and animals). To adapt to the very different environments within different hosts, protozoans have to change their morphology frequently. These changes are dependent on the expression of genes

(Lun et al. 2015). Gene expression and regulation in protozoan organisms is, therefore, potentially very complicated.

3. The current analysis software for microRNA identification has been specifically designed for use with higher animals and plants. However, the small RNAs in protozoan species are very different from those found in higher organisms. Thus, this software cannot be used directly for the investigation of protozoan small RNAs.
4. The next generation sequencing (deep-sequencing) technology has improved the convenience and accuracy of microRNA identification. However, compared with higher organisms, there is very limited small RNA transcriptome data for protozoan species lodged in databases. This is especially true for comparative data between different life cycle stages.

11.3 Computational Methods for Genome-Wide Analysis of Small RNAs in Animal-Like Protozoans

To specifically investigate small RNAs in protozoans, we developed two computational methods to analyze endo-siRNAs and microRNAs from deep-sequencing results which could be used for animal-like protozoans. The origins of siRNAs have been found to be from longer double stranded RNA which have been cleaved by the protein Dicer. Thus siRNAs can be distinguished from random degradation or small RNA production by pretreating extracted protozoan RNA with Dicer (Fig. 11.1). The families of fragments are then subjected to deep sequencing using standard approaches. Next the 3′ adaptors were removed from the illumina-generated sequence reads and then the remaining 18–30 nt reads were mapped to the desired protozoan genome database(s). For reads that mapped to multiple loci, they were excluded if there were greater than nine hits on the genome. Figure 11.1 shows the flowchart for siRNA cluster identification. We first counted the numbers (N) of small RNAs that mapped to both strands of the genome to generate two numbers: one for the sense and the other for the antisense strand. These small RNAs are linked into a cluster if they have at least one base that is overlapping.The length of the longest cluster is then calculated (L). These numbers were then divided by the length of the cluster (L) to obtain the densities (N/L) of small RNA enrichment. We chose the density of 0.04 as a cutoff in order to select candidate NATs with high potential to produce NAT derived siRNAs. This threshold is the same as used in the study by Guo et al. on trans-NATs in rice (Guo et al. 2009). Identification of microRNA was carried out as shown in Fig. 11.2. These are usually found in a single location and have no adjacent clusters. They are recognized by regions of reverse complementarity enabling the formation of hairpin secondary structure.

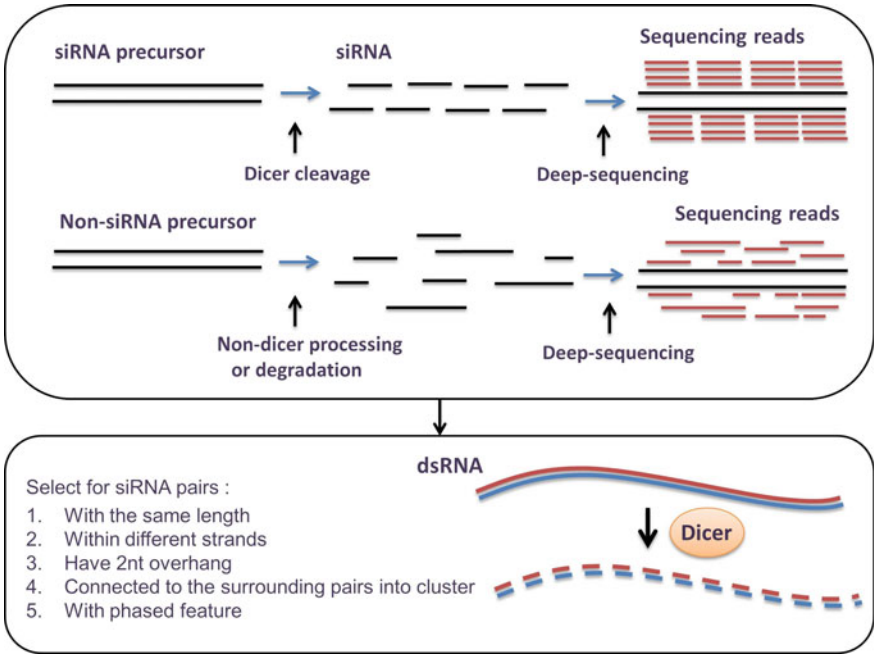


Fig. 11.1 siRNA cluster identification

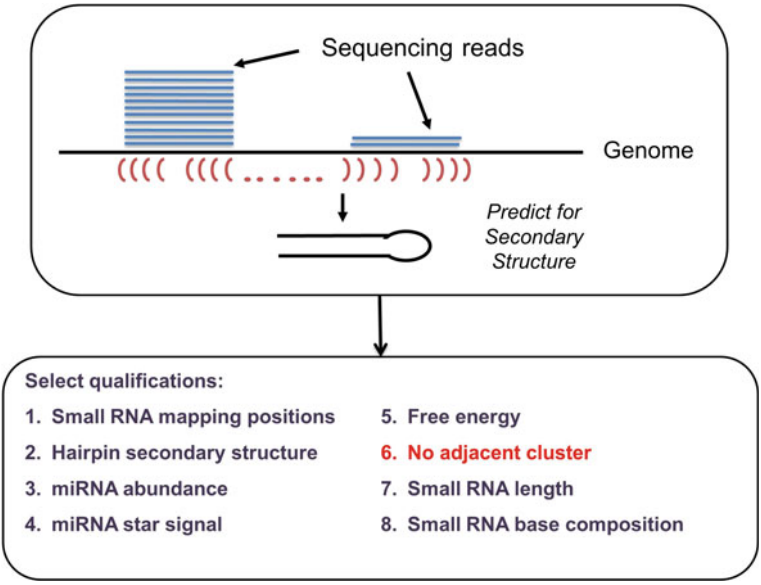


Fig. 11.2 microRNA identification

11.4 Small RNA Investigation in *Trypanosoma brucei*

We combined computational analysis and high-throughput sequencing to study siRNAs and microRNAs in *T. brucei*, produced during the two main life cycle stages, bloodstream form (BF) and procyclic form (PF). We focused specifically on Natural Antisense Transcript (NAT) pairs derived siRNAs in order to gain further insight into the extent and functional role of these RNAs in protozoan organisms. We identified hundreds of NATs on the annotated transcript-scale and whole genome-wide analyses (Tschudi et al. 2012a). We identified the majority of these small RNAs as siRNAs, clearly indicating the prevalence of this mechanism and its potential importance in protozoan parasites.

11.4.1 cis- *and* trans-*NATs Were Identified Within the* T. brucei *Genome*

Putative cis and trans-NATs identification was based on the annotation of *T. brucei* transcripts (Fig. 11.3). To identify the potential cis-NAT pairs, we compared the genomic loci of all annotated transcripts to search for gene pairs that in the same loci but from opposite strands. Eleven enclosed pairs of transcript elements were identified by overlapping reverse-complementary sequences and all of which were formed by protein coding genes and non-coding genes (snoRNAs and snRNA).

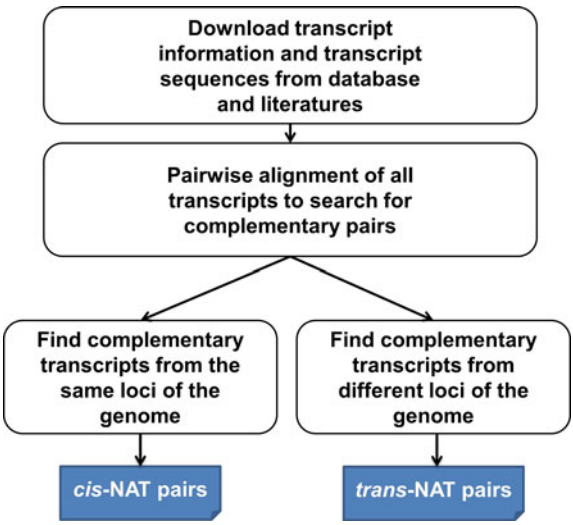


Fig. 11.3 Genome-wide identification of cis- and trans-NATs

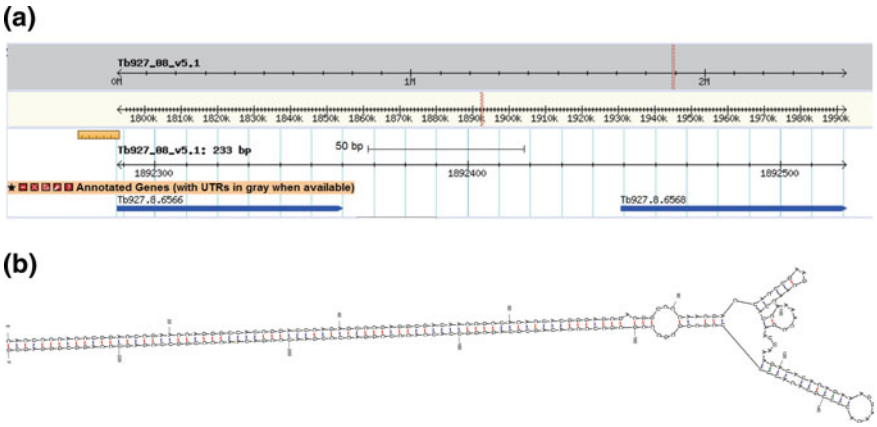


Fig. 11.4 Location and structure of two adjacent tRNAs which can form into NAT pairs. **a** Two Glutamine tRNAs (Tb927.8.6566 and Tb927.8.6568) at the adjacent locus of Tb927_08_v4 from 1892288 to 1892520. **b** Predicted structures that would be formed by base-pair interactions of the region. Secondary structures are folded using Mfold (Zuker 1989)

We carried out a pairwise alignment on all annotated transcript sequences to find *trans*-NAT pairs and identified 209 pairs of *trans*-NAT candidates. Specially, two tRNAs in adjacent locations could form into hairpin structures. Both of the two tRNAs were glutamine tRNA, locating in adjacent loci in plus strand of Tb927_08_v4 (Fig. 11.4a). The two inverted tRNAs were composed of perfect palindromes with an 89 nt internal sequence and were able to form highly stable hairpin loops (Fig. 11.4b).

We did not find any NATs in the mitochondrial genome. This was in accordance with the results in *P. falciparum* (Gunasekera et al. 2004) but inconsistant with the results in humans, mice and worms in which the abundance of antisense transcripts were low in the nuclear genome but high in abundance in the mitochondrial genome (Zhang et al. 2006). This might be an interesting subject for further study.

11.4.2 NAT-Derived Small RNAs Show Features of siRNAs

The NATs in both life stages were distributed across all of the 12 chromosomes of *T. brucei* TREU927 strain (Fig. 11.5). Chromosome sequences of *T. brucei* were downloaded from the TriTrypDB (release-2.1) component of EuPathDB (http://tritrypdb.org/tritrypdb/) (Aslett et al. 2010). Some NATs exhibit specific clustering across particular regions of the chromosomes (Fig. 11.5). Furthermore, we realized that these clusters were the regions where duplication events occurred frequently. For example, the longest NAT extended 8436 bp and spanned a cluster of

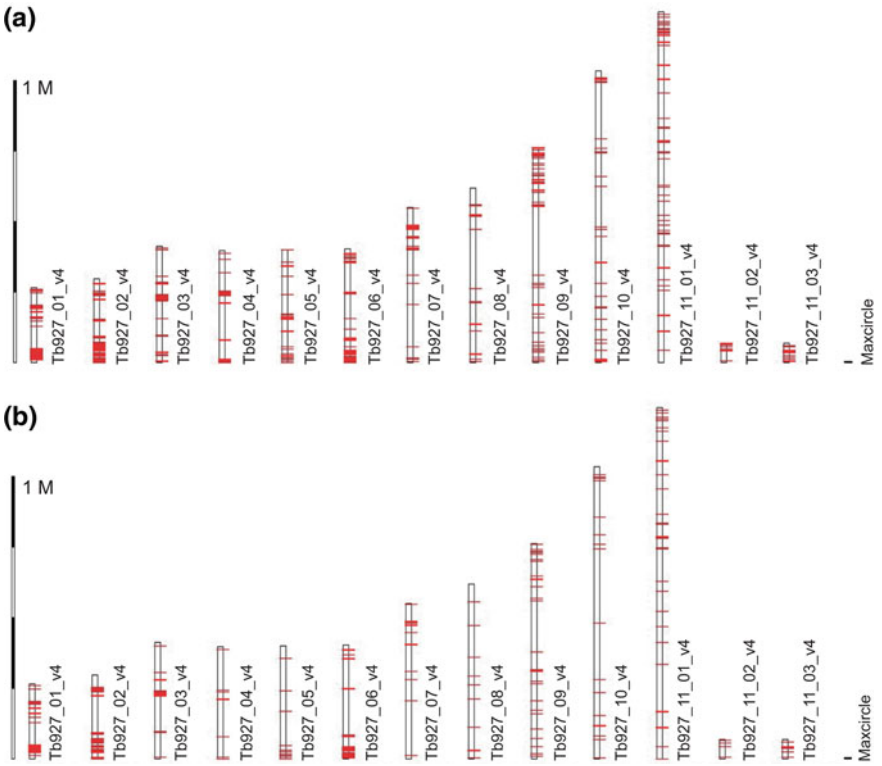


Fig. 11.5 Chromosome map of NATs. All of the NATs are mapped by there location on the chromosome in BF stage (**a**) and PF stage (**b**)

transcripts which were mainly transcribed from retrotransposon hot spot protein (RHS) pseudogenes. As shown in Fig. 11.6a, the NAT on Tb927_01_v4 from 126434 to 134869, where 6 transcripts were crossed, which included four RHS pseudogenes and two protein-coding genes with unknown function.

For each NAT type (i.e. cis-NATs or trans-NATs), we computed the densities of small RNA loci in the overlapping region. Small RNAs which derived from sense and antisense of the genome region are widely distributed in the bloodstream form and procyclic form (Fig. 11.6b). We then calculated the features of these small RNAs. In total, we got 259433 and 84339 NAT-derived small RNAs in bloodstream form and procyclic form, respectively. Most of these small RNAs were in the length region of 23–26 nt and had a characteristic "U" at the 5′-terminal (Fig. 11.7). These features correspond to the known siRNAs in *T. brucei* (Djikeng et al. 2001). Our further experiments showed that these small RNAs are indeed siRNAs and dependent on dicer processing (Zheng et al. 2013).

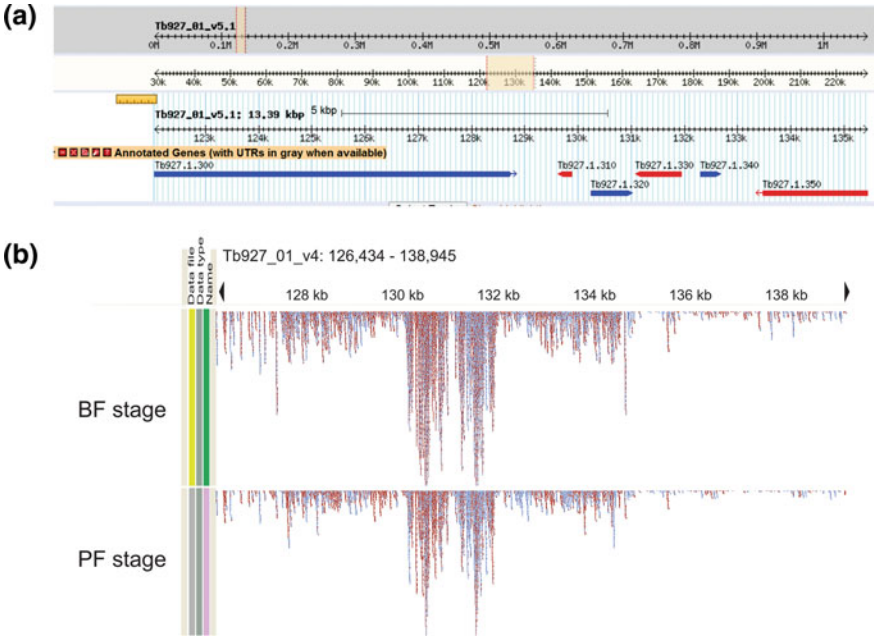


Fig. 11.6 Example of small RNAs derived from NAT. **a** is the screenshot of a region of T. brucei genome browser from TriTrypDB (Aslett et al. 2010) which show transcripts position information across the genome. *Arrow* indicates the transcription direction (5′ to 3′). *Color* also indicates the transcription direction: *blue* for sense transcripts and *red* for antisense transcripts. **b** Is the screenshot of a region of T. brucei genome browser from IGV (Thorvaldsdottir et al. 2013) The *small lines* with different *color* indicate the small RNAs originated position the *color* indicate their transcription direction (*blue* for sense *red* for antisense). *Top* is BF stage and *bottom* is PF stage

11.4.3 Endo-siRNAs Have Multiple Sources and Differential Expression in Two Life Cycle Stages

In addition to NATs, it has been shown that there are many regions in the genome that could generate siRNAs (Djikeng et al. 2001; Zheng et al. 2013; Tschudi et al. 2012a). These sources are listed in Table 11.1. Stage-specific siRNAs have been proved to be induced to respond to special environmental stress (Borsani et al. 2005; Hao et al. 2010). But what about the siRNAs in *T. brucei*? In our results, 78.6 % of siRNAs were only identified in the bloodsteam slender form, whereas only 9.2 % were specifically expressed in the insect procyclic form. Interestingly, the stage-specific siRNAs tended to be transcribed from clustered regions. For example, the region of Tb927_07_v4 (317344–621335) is a source of siRNAs. Surprisingly, although plenty of siRNAs were derived from this region in the slender form, little was found in the procyclic form (Fig. 11.8). Next, we analyzed the expression patterns of siRNAs in both stages and found 534 siRNAs exhibited

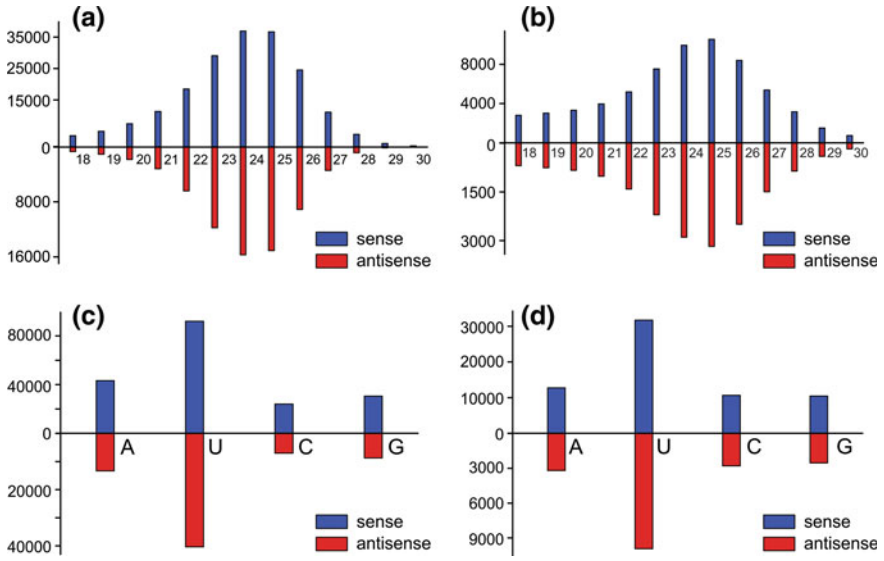


Fig. 11.7 Feature distribution of BTR-derived small RNAs. The length distribution of small RNAs derived from BTRs in BF stage (**a**) and PF stage (**b**). The 5′ end nucleotide distribution of small RNAs derived from BTRs in BF stage (**c**) and PF stage (**d**)

Table 11.1 The sources of siRNAs in *T. brucei*

Abb	Full Name	References
CIR147	Chromosome inteRNAl repeats	Obado et al. (2005)
SLACS	Spliced leader-associated conserved sequence	Djikeng et al. (2001)
Ingi	Ingi	Djikeng et al. (2001)
IR	Long inverted repeats	Tschudi et al. (2012a)
CTU	Convergent transcription unit	Tschudi et al. (2012a)
miscRNA	cMiscellaneous,	Tschudi et al. (2012a)
NAT	Nature antisense transcripts	Zheng et al. (2013)

significant expression differences ($p < 0.001$), 79 % are expressed significantly higher in the BF while only 21 % are expressed significantly higher in the PF stage.

11.4.4 Non-existence of Canonical microRNA in Selected Stages of T. brucei

In 2008, Mallick et al. reported that they used a bioinformatic approach to identify microRNA genes in *Trypanosoma brucei* and to predict their functions

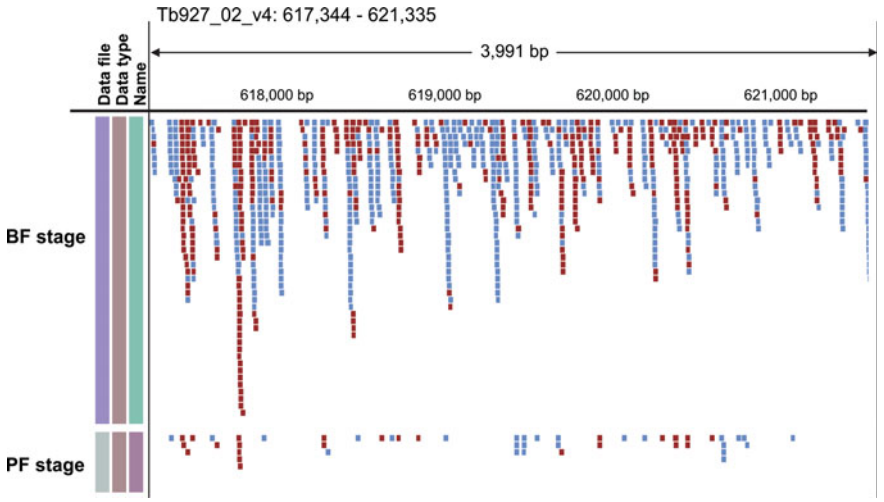


Fig. 11.8 Is the screenshot of a region of *T. brucei* genome browser from IGV(Thorvaldsdottir et al. 2013). The *small lines* with different *color* indicate the small RNAs original position and the *color* indicates their transcriptional direction (*blue* for sense *red* for antisense). *Top* is BF stage and *bottom* is PF stage

(Mallick et al. 2008). In total, they identified 1162 microRNA candidates and most of them targeted to the variant surface glycoprotein (VSG) in this parasite protozoan. They suggested that the microRNAs in *T. brucei* might play roles as genetic switches in modulating host-parasite interaction.

However, our results from the high-throughput sequencing on *T. brucei* did not identify any microRNA molecules from this parasite. Combined with the previous studies (Tschudi et al. 2012b; Kolev et al. 2011) we did not support the existence of microRNA in this parasite, at least in these two stages. We do not know the reasons behind this different result but it might be related to the microRNA genes which are expressed only in different stages of the life cycle of *T. brucei*. More studies need to be carried out to investigate this interesting topic.

No matter whether microRNAs exist in *T. brucei* or not, our results have shown that siRNAs as well as other small RNAs, like tRNA-derived small RNAs and rRNA derived small RNAs, are important regulators in the differentiation of *T. brucei* (Zheng et al. 2013). By comparison with the procyclic form, we found that some genes in the VSG family were specifically regulated by siRNAs in the slender form of *T. brucei*. VSG is an important antigen in the slender form of *T. brucei*. The extensive variety of this antigen allows the parasite to escape from the immunological defense of the mammalian hosts. Therefore, we propose that the siRNAs correspondingly evolve to eliminate the obsolete VSG transcripts as rapidly as possible, so that the novel VSG can be effective in a timely way. Taken together, our results provide genome-wide identification of small RNAs in *T. brucei*. These results will facilitate functional studies of siRNAs in this model

protozoan, as well as in other protozoan species, and help to unravel complex gene and non-coding RNA regulatory networks in eukaryotes.

11.5 Investigation into Small RNAs in *Giardia lamblia*

Giardia lamblia is one of the most primitive eukaryotes, and the study of its sRNA transcriptome could shed light on the origin and evolution of various sRNAs. We found that the *G. lamblia* sRNA transcriptome mainly consists of two kinds of sRNAs: endo-siRNA and tRNA derived sRNA (Liao et al. 2014). Both of these two kinds of sRNAs are also found in higher eukaryotic genomes, such as humans, suggesting they might be involved in the regulation of basic biological processes in eukaryotes.

11.5.1 Retrotransposon and eSGR Derived Endo-siRNAs

G. lamblia endo-siRNAs were primarily derived from two sources, retrotransposon and endo-siRNA generating regions (eSGR). The *G. lamblia* genome has three retrotransposons: two of these (GilT and GilM) are located in the telomeric regions, while the third retrotransposon (GilD) is dead (Arkhipova and Morrison 2001). Both functional retrotransposons, GilT and GilM, generate large number of endo-siRNAs. The dead GilD does not appear to generate any sRNAs. The endo-siRNA generating region (eSGR) is a region in the *G. lamblia* genome that seems to be specifically used to produce endo-siRNAs (Liao et al. 2014) and encodes three eSGRs (Fig. 11.9), named eSGRI, eSGRII and eSGRIII respectively. The eSGRI region is the longest and generates most of the eSGR derived endo-siRNAs (87.1 %). It spans

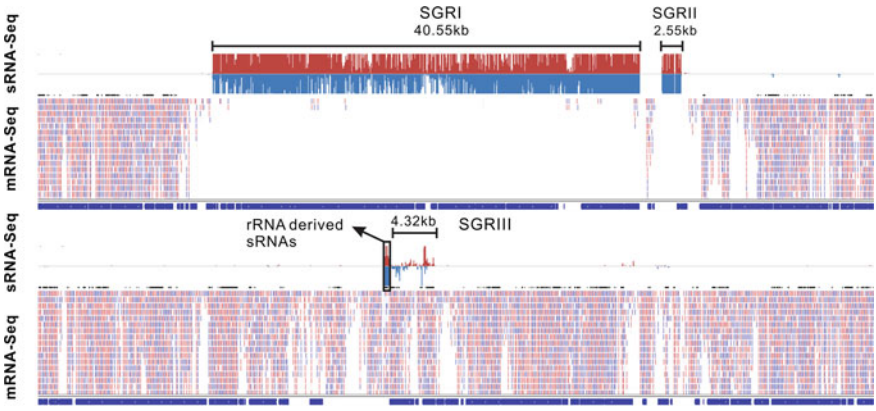


Fig. 11.9 The mRNA and sRNA expression statuses of three SGRs

40.55 kb in length and contains 34 protein-coding genes. Intriguingly, all these 34 protein-coding genes in eSGRI are not expressed. It is possible, therefore, that their expression is blocked by eSGRI derived endo-siRNA. In contrast, eSGRII and eSGRIII produce much less endo-siRNAs than eSGRI and they occupy shorter regions, 2.55 kb for eSGRII and 4.32 kb for eSGRIII (Fig. 11.9). These two regions also contain some protein coding genes, but these mRNAs are being expressed (Fig. 11.9). This might be due to the low expression levels of eSGRII and eSGRIII derived endo-siRNAs. The eSGR derived endo-siRNAs have the same characteristics as retrotransposon derived endo-siRNAs, which may be caused by them sharing an identical biogenesis pathway.

G. lamblia endo-siRNAs share very similar characteristics with endo-siRNAs of other eukaryotes including *T. brucei*. For instance, like endo-siRNAs in other eukaryotes, the *G. lamblia* endo-siRNAs have a length of between 20–30 nt, they are derived from both strands of the endo-siRNA generating regions and have one or more post-transcriptionally added 3′ untemplated nucleotides. The conservation of these characteristics suggests that they may be important for the function of endo-siRNAs. Unexpectedly, the 5′ end nucleotide of *G. lamblia* endo-siRNAs does not appear to be "U" like the endo-siRNAs of higher eukaryotes. It's unclear yet how and when endo-siRNA are processed to acquire their 5′ end "U". This is an interesting question worth pursuing in the future.

11.5.2 The Function of the G. lamblia *Endo-siRNA Pathway*

It's hard to clarify the detailed function of each endo-siRNA, because of the huge number of endo-siRNAs. We have identified more than a hundred thousand endo-siRNAs in the *G. lamblia* trophozoite, and the expression of each of these endo-siRNAs is low. A good way to investigate endo-siRNA function is to block the key protein in the endo-siRNA biogenesis pathway. Through knock-down of the *G. lamblia* Dicer gene (GLDICER), which is responsible for the slicing of the endo-siRNA precursor to generate mature endo-siRNAs (Macrae et al. 2006), we significantly reduced the expression level of all endo-siRNAs and observed that the ability of *G. lamblia* to differentiate was impaired. This result indicated that the endo-siRNA pathway might be involved in the regulation of *G. lamblia* differentiation. Another function of the *G. lamblia* endo-siRNA is in regulating antigenic variation (Prucca et al. 2008). This was found by Prucca et al. through knock-down of two proteins GLDICER and *G. lamblia* RNA dependent RNA Polymerase (GLRDRP). GLRDRP is a protein used to generate appropriate RNA for GLDICER slicing. Both the silencing of GLDICER or GLRDRP lead to the failure of mechanisms of *G. lamblia* antigenic variation. It's unclear yet how endo-siRNAs regulates differentiation and antigenic variation. The detailed study of the regulation mechanisms of endo-siRNAs is key to an understanding of the functional evolution of endo-siRNA in eukaryotes.

11.5.3 *The* G. lamblia *Genome Encodes Six Kinds of tRNA Derived sRNAs*

Recent studies have shown that tRNA derived sRNA (tsRNAs) is involved in the regulation of translation and various pathological processes like cancer (Goodarzi et al. 2015) and sperm tsRNAs from father may mediate intergenerational inheritance of diet-induced metabolic disorders (Chen et al. 2016; Sharma et al. 2016). We found that tRNA derived sRNAs are conserved in *G. lamblia* and exhibited changes in expression profiles between different life stages (Liao et al. 2014; Li et al. 2008), suggesting that the role of tRNA derived sRNA in cells might be more important than previous expected. In total, six kinds of tRNA derived sRNAs were found in *G. lamblia*, they were named as 5tasRNA, actasRNA, 3tasRNA, 5sitRNA, 3sitRNA and 5EsRNA respectively. The 5tasRNA, actasRNA, 3tasRNA members were derived from the 5′ end, middle region (most of the anticodon stem loop) and 3′ end of the mature tRNA, with lengths of around 20–22, 24–26, 28–30 nt respectively. Similarly 5tasRNA and 3tasRNA, 5sitRNA and 3sitRNA were also processed from the 5′ and 3′ ends of mature tRNA respectively. However these had much longer lengths, 50 nt for 5sitRNA and 46 nt for 3sitRNA. 5EsRNA was derived from the 5′ end of pseudotRNA-Gln (TTT), with a length of 36 nt.

11.5.4 *The Biogenesis and Potential Function of* G. lamblia *tRNA Derived sRNAs*

With the exception of 5EsRNA, all five remaining kinds of tRNA derived sRNAs were generated from various mature tRNAs. We identified two cleavage sites, Cleavage Site I and II (CSI and CSII), in the clover leaf structure of various tRNAs (Fig. 11.10). We found that 5tasRNA, actasRNA, 3tasRNA, and 5sitRNA, 3sitRNA might be all generated from cleavage at these two sites by an, as yet,

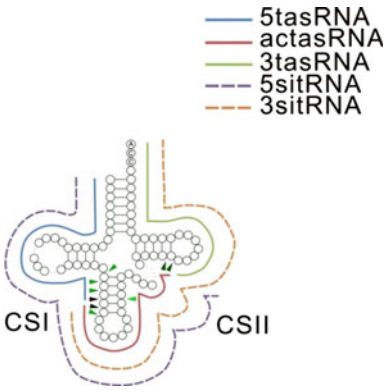


Fig. 11.10 Five kinds of sRNAs derived from mature tRNAs. CS refers to cleavage site

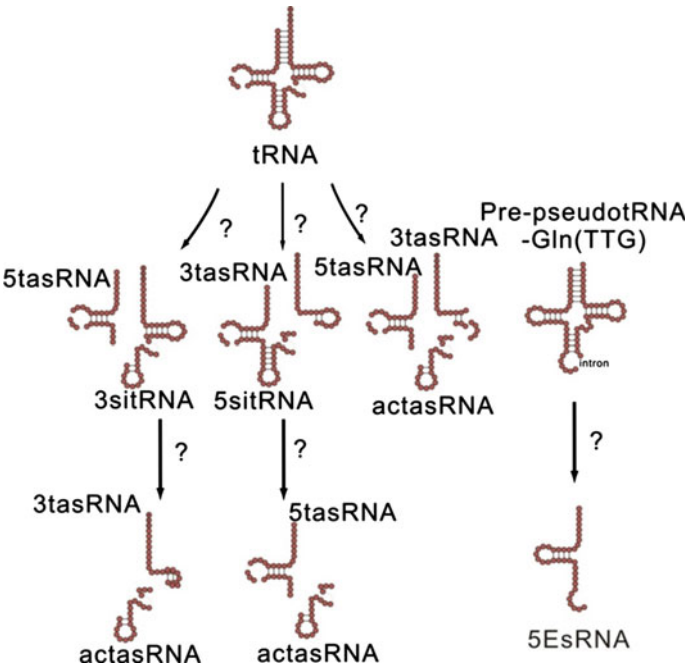


Fig. 11.11 The potential biogenesis pathway of six kinds of *G. lamblia* tRNA derived sRNAs

unknown RNA endonuclease. Considering that the 5tasRNA and 3sitRNA can be integrated into an intact tRNA, as does 3tasRNA and 5sitRNA, and also 5tasRNA, actasRNA and 3tasRNA, it's possible that the cleavage of mature tRNAs may generate one of these three groups of products (Fig. 11.11). The sequence of 5EsRNA covers the whole 5′-end exon of pseudotRNA-Gln(TTG), its biogenesis might borrow the pecursor process machinery from pseudotRNA-Gln(TTT).

The detailed functions of G. lamblia tRNA derived sRNAs are as yet unclear. 3sitRNA was the first identified *G. lamblia* tRNA derived sitRNA (Li et al. 2008). It is not expressed in trophozoites, but can be induced by several stress factors, thus connecting them with a potential stress resistant related function. Similarly, 3sitRNA, 5sitRNA, 5tasRNA, actasRNA and 3tasRNA are not expressed in the trophozoite stage either. However, all of these five kinds of sRNAs are expressed when the trophozoite starts to differentiate into cysts. They are also expressed at a high level in cysts, indicating that they might be involved in the differentiation process in *Giardia*. It's difficult to investigate the function of tRNA derived sRNAs as yet, because we currently lack appropriate methods to manipulate their gene expression. However, the identification of endonucleases responsible for the biogenesis of these sRNAs may overcome this obstacle.

11.5.5 *Does the* G. lamblia *Genome Possess microRNA-Encoding Genes?*

Because miRNAs share the same biogenesis machinery and effectors with endo-siRNA, it allows the possibility that the genome of *G. lamblia* may encode miRNA genes. Indeed, many studies have identified 166 potential miRNA in *G. lamblia* (Saraiya et al. 2011; Huang et al. 2012; Zhang et al. 2009; Chen et al. 2009; Saraiya and Wang 2008; Li et al. 2011, 2012). However, 76 % (126) of these miRNAs could not be detected by our very high depth sRNA sequencing. Among the 40 miRNAs that were revealed in our sequencing data, 12.5 % (5) were derived from ncRNAs, 50 % (20) were derived from an eSGR region, and the expression level of the remaining 37.5 % (15) miRNA is very low and only accounts for 0.0048 % of all reads in the trophozoite sRNA library. Therefore, there is no conclusive evidence, yet, that the *G. lamblia* genome encodes any miRNAs.

11.6 Concluding Remarks

In this review, we summarize the research on the analysis of small RNA transcriptome data on *Trypanosoma brucei* and *Giardia lamblia*. The results show that most of the protozoans do not have canonical miRNAs. However, there are many types of endogenous small interfering RNAs (endo-siRNAs) derived from pseudogenes, transcriptional elements, bidirectional transcription and NATs (natural antisense transcripts). In addition, widespread variation among the tRNA derived small RNAs were also observed in different stages of life cycle in these parasites. These results shed light on questions regarding the potential origins and evolution of mechanisms of small RNA-based control of gene expression.

Acknowledgments This work was supported by the grants from Ministry of Science and Technology of China; National Basic Research Program (No. 2011CB811300) from the National Basic Research program ("973" program) to L.-H.Q.; grants from the National Natural Science Foundation of China: 31401975 (to L.-L.Z.), 31401092 (to J.-Y.L.), 31301876 (to Y.-Z.W.), 31272305 and 31472058 (to Z.-R.L.) and 31471223 (to L.-H.Q.), and grants from Natural Science Foundation of Guangdong Province 2014A030313175 (to J.-Y.L.) and 2015A030311042 (to Z.-R. L). This was also supported in part by the Guangdong Province Key Laboratory of Computational Science and the Guangdong Province Computational Science Innovative Research Team. The authors declare no conflict of interest.

References

Aravin AA, Hannon GJ (2008) Small RNA silencing pathways in germ and stem cells. In Cold Spring Harbor symposia on quantitative biology 73:283–290. doi:10.1101/sqb.2008.73.058

Arkhipova IR, Morrison HG (2001) Three retrotransposon families in the genome of *Giardia lamblia*: two telomeric, one dead. Proc Natl Acad Sci USA 98(25):14497–14502. doi:10.1073/pnas.231494798

Aslett M, Aurrecoechea C, Berriman M, Brestelli J, Brunk BP, Carrington M, Depledge DP, Fischer S, Gajria B, Gao X, Gardner MJ, Gingle A, Grant G, Harb OS, Heiges M, Hertz-Fowler C, Houston R, Innamorato F, Iodice J, Kissinger JC, Kraemer E, Li W, Logan FJ, Miller JA, Mitra S, Myler PJ, Nayak V, Pennington C, Phan I, Pinney DF, Ramasamy G, Rogers MB, Roos DS, Ross C, Sivam D, Smith DF, Srinivasamoorthy G, Stoeckert CJ, Jr., Subramanian S, Thibodeau R, Tivey A, Treatman C, Velarde G, Wang H (2010) TriTrypDB: a functional genomic resource for the Trypanosomatidae. Nucleic Acids Res 38 (Database issue): D457-462. doi:10.1093/nar/gkp851

Borsani O, Zhu J, Verslues PE, Sunkar R, Zhu JK (2005) Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. Cell 123(7):1279–1291. doi:10.1016/j.cell.2005.11.035

Braun L, Cannella D, Ortet P, Barakat M, Sautel CF, Kieffer S, Garin J, Bastien O, Voinnet O, Hakimi M-A (2010) A complex small RNA repertoire is generated by a plant/fungal-like machinery and effected by a metazoan-like Argonaute in the single-cell human parasite *Toxoplasma gondii*. PLoS Pathog 6:e1000920. doi:10.1371/journal.ppat.1000920

Cavalier-Smith T (2003) Protist phylogeny and the high-level classification of Protozoa. Eur J Protistol 39(4):338–348. doi:10.1078/0932-4739-00002

Chen Q, Yan M, Cao Z, Li X, Zhang Y, Shi J, Feng GH, Peng H, Zhang X, Zhang Y, Qian J, Duan E, Zhai Q, Zhou Q (2016) Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder. Science 351(6271):397–400. doi:10.1126/science.aad7977

Chen XS, Collins LJ, Biggs PJ, Penny D (2009) High throughput genome-wide survey of small RNAs from the parasitic protists *Giardia intestinalis* and *Trichomonas vaginalis*. Genome Biol Evol 1:165–175. doi:10.1093/gbe/evp017

Chung WJ, Okamura K, Martin R, Lai EC (2008) Endogenous RNA interference provides a somatic defense against *Drosophila* transposons. Curr Biol CB 18(11):795–802. doi:10.1016/j.cub.2008.05.006

Czech B, Malone CD, Zhou R, Stark A, Schlingeheyde C, Dus M, Perrimon N, Kellis M, Wohlschlegel JA, Sachidanandam R, Hannon GJ, Brennecke J (2008) An endogenous small interfering RNA pathway in *Drosophila*. Nature 453(7196):798–802. doi:10.1038/nature07007

De S, Pal D, Ghosh SK (2006) Entamoeba histolytica: computational identification of putative microRNA candidates. Exp Parasitol 113(4):239–243. doi:10.1016/j.exppara.2006.01.009

Djikeng A, Shi HF, Tschudi C, Ullu E (2001) RNA interference in *Trypanosoma brucei*: Cloning of small interfering RNAs provides evidence for retroposon-derived 24-26-nucleotide RNAs. RNA 7(11):1522–1530

Ghildiyal M, Seitz H, Horwich MD, Li C, Du T, Lee S, Xu J, Kittler EL, Zapp ML, Weng Z, Zamore PD (2008) Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells. Science 320(5879):1077–1081. doi:10.1126/science.1157396

Goodarzi H, Liu X, Nguyen HC, Zhang S, Fish L, Tavazoie SF (2015) Endogenous tRNA-Derived Fragments Suppress Breast Cancer Progression via YBX1 Displacement. Cell 161(4):790–802. doi:10.1016/j.cell.2015.02.053

Gunasekera AM, Patankar S, Schug J, Eisen G, Kissinger J, Roos D, Wirth DF (2004) Widespread distribution of antisense transcripts in the *Plasmodium falciparum* genome. Mol Biochem Parasit 136(1):35–42. doi:10.1016/j.molbiopara.2004.02.007

Guo X, Zhang Z, Gerstein MB, Zheng D (2009) Small RNAs originated from pseudogenes: cis- or trans-acting? PLoS Comput Biol 5(7):e1000449. doi:10.1371/journal.pcbi.1000449

Ha M, Kim VN (2014) Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol 15(8):509–524. doi:10.1038/nrm3838

Hao L, Cai P, Jiang N, Wang H, Chen Q (2010) Identification and characterization of microRNAs and endogenous siRNAs in *Schistosoma japonicum*. BMC Genom 11:55. doi:10.1186/1471-2164-11-55

Huang PJ, Lin WC, Chen SC, Lin YH, Sun CH, Lyu PC, Tang P (2012) Identification of putative miRNAs from the deep-branching unicellular flagellates. Genomics 99(2):101–107. doi:10.1016/j.ygeno.2011.11.002

Kawamura Y, Saito K, Kin T, Ono Y, Asai K, Sunohara T, Okada TN, Siomi MC, Siomi H (2008) *Drosophila endogenous* small RNAs bind to Argonaute 2 in somatic cells. Nature 453 (7196):793–797. doi:10.1038/nature06938

Kolev NG, Tschudi C, Ullu E (2011) RNA interference in protozoan parasites: achievements and challenges. Eukaryot Cell 10(9):1156–1163. doi:10.1128/EC.05114-11

Li M, Belmonte JCI (2015) Roles for noncoding RNAs in cell-fate determination and regeneration. Nat Struct Mol Biol 22(1):2–4

Li ZH, Rana TM (2014) Therapeutic targeting of microRNAs: current status and future challenges. Nat Rev Drug Discovery 13(8):622–638. doi:10.1038/nrd4359

Li Y, Luo J, Zhou H, Liao JY, Ma LM, Chen YQ, Qu LH (2008) Stress-induced tRNA-derived RNAs: a novel class of small RNAs in the primitive eukaryote *Giardia lamblia*. Nucleic Acids Res 36(19):6048–6055. doi:10.1093/nar/gkn596

Li W, Saraiya AA, Wang CC (2011) Gene regulation in *Giardia lamblia* involves a putative microRNA derived from a small nucleolar RNA. PLoS Negl Trop Dis 5(10):e1338. doi:10.1371/journal.pntd.0001338

Li W, Saraiya AA, Wang CC (2012) The profile of snoRNA-derived microRNAs that regulate expression of variant surface proteins in *Giardia lamblia*. Cell Microbiol 14(9):1455–1473. doi:10.1111/j.1462-5822.2012.01811.x

Liao JY, Guo YH, Zheng LL, Li Y, Xu WL, Zhang YC, Zhou H, Lun ZR, Ayala FJ, Qu LH (2014) Both endo-siRNAs and tRNA-derived small RNAs are involved in the differentiation of primitive eukaryote *Giardia lamblia*. Proc Natl Acad Sci USA 111(39):14159–14164. doi:10.1073/pnas.1414394111

Lin WC, Li SC, Lin WC, Shin JW, Hu SN, Yu XM, Huang TY, Chen SC, Chen HC, Chen SJ, Huang PJ, Gan RRC, Chiu CH, Tang P (2009) Identification of microRNA in the protist *Trichomonas vaginalis*. Genomics 93(5):487–493. doi:10.1016/j.ygeno.2009.01.004

Lun ZR, Lai DH, Wen YZ, Zheng LL, Shen JL, Yang TB, Zhou WL, Qu LH, Hide G, Ayala FJ (2015) Cancer in the parasitic protozoans *Trypanosoma brucei* and *Toxoplasma gondii*. Proc Natl Acad Sci USA 112(29):8835–8842. doi:10.1073/pnas.1502599112

Macrae IJ, Zhou K, Li F, Repic A, Brooks AN, Cande WZ, Adams PD, Doudna JA (2006) Structural basis for double-stranded RNA processing by Dicer. Science 311(5758):195–198. doi:10.1126/science.1121638

Mallick B, Ghosh Z, Chakrabarti J (2008) MicroRNA switches in *Trypanosoma brucei*. Biochem Biophys Res Commun 372(3):459–463. doi:10.1016/j.bbrc.2008.05.084

Montagnes D, Roberts E, Lukes J, Lowe C (2012) The rise of model protozoa. Trends Microbiol 20(4):184–191. doi:10.1016/j.tim.2012.01.007

Nilsen TW (2008) Endo-siRNAs: yet another layer of complexity in RNA silencing. Nat Struct Mol Biol 15(6):546–548. doi:10.1038/nsmb0608-546

Obado SO, Taylor MC, Wilkinson SR, Bromley EV, Kelly JM (2005) Functional mapping of a *Trypanosome* centromere by chromosome fragmentation identifies a 16-kb GC-rich transcriptional "strand-switch" domain as a major feature. Genome Res 15(1):36–43. doi:10.1101/gr.2895105

Okamura K, Lai EC (2008) Endogenous small interfering RNAs in animals. Nat Rev Mol Cell Biol 9(9):673–678. doi:10.1038/nrm2479

Okamura K, Balla S, Martin R, Liu N, Lai EC (2008a) Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in *Drosophila melanogaster*. Nat Struct Mol Biol 15(6):581–590. doi:10.1038/nsmb.1438

Okamura K, Chung WJ, Ruby JG, Guo H, Bartel DP, Lai EC (2008b) The *Drosophila* hairpin RNA pathway generates endogenous short interfering RNAs. Nature 453(7196):803–806. doi:10.1038/nature07015

Prucca CG, Slavin I, Quiroga R, Elias EV, Rivero FD, Saura A, Carranza PG, Lujan HD (2008) Antigenic variation in *Giardia lamblia* is regulated by RNA interference. Nature 456 (7223):750–754. doi:10.1038/nature07585

Saraiya AA, Wang CC (2008) snoRNA, a novel precursor of microRNA in *Giardia lamblia*. PLoS Pathog 4(11):e1000224. doi:10.1371/journal.ppat.1000224

Saraiya AA, Li W, Wang CC (2011) A microRNA derived from an apparent canonical biogenesis pathway regulates variant surface protein gene expression in *Giardia lamblia*. RNA 17 (12):2152–2164. doi:10.1261/rna.028118.111

Shan SW, Fang L, Shatseva T, Rutnam ZJ, Yang X, Du W, Lu WY, Xuan JW, Deng Z, Yang BB (2013) Mature miR-17-5p and passenger miR-17-3p induce hepatocellular carcinoma by targeting PTEN, GalNT7 and vimentin in different signal pathways. J Cell Sci 126(Pt 6):1517–1530. doi:10.1242/jcs.122895

Sharma U, Conine CC, Shea JM, Boskovic A, Derr AG, Bing XY, Belleannee C, Kucukural A, Serra RW, Sun F, Song L, Carone BR, Ricci EP, Li XZ, Fauquier L, Moore MJ, Sullivan R, Mello CC, Garber M, Rando OJ (2016) Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals. Science 351(6271):391–396. doi:10.1126/science.aad6780

Sood P, Krek A, Zavolan M, Macino G, Rajewsky N (2006) Cell-type-specific signatures of microRNAs on target mRNA expression. Proc Natl Acad Sci USA 103(8):2746–2751. doi:10.1073/pnas.0511045103

Tam OH, Aravin AA, Stein P, Girard A, Murchison EP, Cheloufi S, Hodges E, Anger M, Sachidanandam R, Schultz RM, Hannon GJ (2008) Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature 453(7194):534–538. doi:10.1038/nature06904

Tarver JE, Donoghue PC, Peterson KJ (2012) Do miRNAs have a deep evolutionary history? BioEssays News Rev Mol Cell Dev Biol 34(10):857–866. doi:10.1002/bies.201200055

Thorvaldsdottir H, Robinson JT, Mesirov JP (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefs Bioinform 14(2):178–192. doi:10.1093/bib/bbs017

Tschudi C, Shi H, Franklin JB, Ullu E (2012a) Small interfering RNA-producing loci in the ancient parasitic eukaryote *Trypanosoma brucei*. BMC Genom 13:427. doi:10.1186/1471-2164-13-427

Tschudi C, Shi H, Franklin JB, Ullu E (2012b) Small interfering RNA-producing loci in the ancient parasitic eukaryote *Trypanosoma brucei*. BMC Genom 13:427. doi:10.1186/1471-2164-13-427

Watanabe T, Totoki Y, Toyoda A, Kaneda M, Kuramochi-Miyagawa S, Obata Y, Chiba H, Kohara Y, Kono T, Nakano T, Surani MA, Sakaki Y, Sasaki H (2008) Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature 453(7194):539–543. doi:10.1038/nature06908

Wen YZ, Zheng LL, Liao JY, Wang MH, Wei Y, Guo XM, Qu LH, Ayala FJ, Lun ZR (2011) Pseudogene-derived small interference RNAs regulate gene expression in African *Trypanosoma brucei*. Proc Natl Acad Sci USA 108(20):8345–8350. doi:10.1073/pnas.1103894108

Wheeler BM, Heimberg AM, Moy VN, Sperling EA, Holstein TW, Heber S, Peterson KJ (2009) The deep evolution of metazoan microRNAs. Evol Dev 11(1):50–68. doi:10.1111/j.1525-142X.2008.00302.x

Wightman B, Ha I, Ruvkun G (1993) Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*. Cell 75(5):855–862

Wu C, Arora P (2014) MicroRNA passenger strand: orchestral symphony of paracrine signaling. Circ Cardiovas Genet 7(4):567–568. doi:10.1161/CIRCGENETICS.114.000805

Yang X, Du WW, Li H, Liu F, Khorshidi A, Rutnam ZJ, Yang BB (2013) Both mature miR-17-5p and passenger strand miR-17-3p target TIMP3 and induce prostate tumor growth and invasion. Nucleic Acids Res 41(21):9688–9704. doi:10.1093/nar/gkt680

Zhang Y, Liu XS, Liu QR, Wei LP (2006) Genome-wide in silico identification and analysis of cis natural antisense transcripts (cis-NATs) in ten species. Nucleic Acids Res 34(12):3465–3475. doi:10.1093/nar/gkl473

Zhang YQ, Chen DL, Tian HF, Zhang BH, Wen JF (2009) Genome-wide computational identification of microRNAs and their targets in the deep-branching eukaryote *Giardia lamblia*. Comput Biol Chem 33(5):391–396. doi:10.1016/j.compbiolchem.2009.07.013

Zheng LL, Wen YZ, Yang JH, Liao JY, Shao P, Xu H, Zhou H, Wen JZ, Lun ZR, Ayala FJ, Qu LH (2013) Comparative transcriptome analysis of small noncoding RNAs in different stages of *Trypanosoma brucei*. RNA 19(7):863–875. doi:10.1261/rna.035683.112

Zuker M (1989) On finding all suboptimal foldings of an RNA molecule. Science 244(4900): 48–52

Chapter 12
Function of Non-coding RNA in *Toxoplasma gondii* Infection

Mariana Matrajt

Abstract The protozoan parasite *Toxoplasma gondii* is an important human and veterinary pathogen. Asexual replication of *T. gondii* in humans and intermediate hosts is characterized by two forms: rapidly growing 'tachyzoites' and latent 'bradyzoite' tissue cysts. Tachyzoites are responsible for acute illness and congenital neurological birth defects, while bradyzoites can remain latent within the tissues for many years, representing a threat to immunocompromised patients. In this chapter I discuss the function of a *T. gondii* non-coding RNA (ncRNA) that has a major role in the critical transition from tachyzoites to bradyzoites. In addition, I discuss how *T. gondii* infection affects the expression of human and mouse microRNAS (miRNAs), in particular, the dysregulation of host miRNA-132 by *T. gondii* infection, and it's implications for dopamine signaling and neuropathology.

12.1 Introduction

The human pathogen *Toxoplasma gondii* is one of the most widely distributed protozoan parasites, infecting approximately one-third of the world's population. *T. gondii* is an obligate intracellular parasite that can infect any nucleated animal cell and is rapidly controlled by the cellular immune response, leaving an unapparent infection. Because of the late development of the cellular immune response during fetal maturation, *T. gondii* is a leading cause of congenital neurological birth defects. In recent years, *Toxoplasma* has also achieved notoriety as a cause of life-threatening opportunistic disease in immunocompromised individuals, including patients with AIDS. Asexual replication of *T. gondii* in intermediate hosts, which includes humans, is characterized by two forms: rapidly growing 'tachyzoites' and 'bradyzoite' tissue cysts. Tachyzoites are responsible for acute illness

M. Matrajt (✉)
Natural and Health Science Department, Quincy College, 24 Saville Ave, Quincy, MA 02169, USA
e-mail: mmatrajt@quincycollege.edu

A.L. Leitão and F.J. Enguita (eds.), *Non-coding RNAs and Inter-kingdom Communication*, DOI 10.1007/978-3-319-39496-1_12

and congenital neurological birth defects, while the more slowly dividing bradyzoite form can remain within the tissues for many years, representing a threat to immunocompromised patients. These two developmental stages are essential for disease propagation and causation.

Congenital toxoplasmosis and toxoplasmic encephalitis are associated with neuropsychiatric symptoms, and the tropism of *T. gondii* for brain tissue has been linked to behavioral changes in humans and animals. *T. gondii* infection affects neurotransmitters levels in both acute and chronic infection (Xiao et al. 2014). Recently, it was shown that *T. gondii* infection dysregulates the expression of a specific host cell microRNA, which is associated with changes in dopamine neurotransmitter signaling (Xiao et al. 2014).

12.2 *T. gondii* Life Cycle and Differentiation

The *T. gondii* life cycle includes sexual and asexual stages. Sexual replication takes place only in cats, but is not obligatory for parasite proliferation. Asexual replication (in a wide variety of intermediate host species and tissues) consists of two inter-converting developmental stages: rapidly growing 'tachyzoites' that are sensitive to the immune system and several drugs, and the slowly dividing encysted 'bradyzoites' that hide from the host's immune response and are resistant to current drug treatment.

T. gondii has been found in almost every warm blooded animal (Hill et al. 2005), and many animals that are used for human food consumption contain bradyzoite cysts, which can lead to human infection. Upon infection, bradyzoite cysts differentiate to tachyzoites which disseminate throughout the body until the immune system is triggered, parasite growth is controlled, and tachyzoites differentiate to bradyzoites. Bradyzoites remain within most tissues for many years, and during the course of a chronic infection some tissue cysts differentiate back to tachyzoites. While this is not a problem for immunocompetent individuals that can mount the proper immune response needed to stop the unrestricted multiplication of tachyzoites (this being the reason why most healthy adults are asymptomatic), in immunocompromised individuals *T. gondii* can cause devastating effects such as fatal encephalitis. There is no effective treatment for chronic toxoplasmosis due to a lack of drugs capable of eliminating tissue cysts.

The asexual cycle bradyzoite-tachyzoite-bradyzoite can repeat for many rounds in intermediate hosts without the need for sexual transmission, however, to complete the life cycle and differentiate into sexual stages, *T. gondii* must pass through the gut of a cat. In the cat intestine, bradyzoites can differentiate into the sexual stage micro- and macrogametocytes that can fuse to form oocysts which are shed in the environment. Oocysts contain sporozoites, which are the product of meiosis. Oocysts in the environment can be ingested by intermediate hosts where they differentiate into tachyzoites completing the life cycle (Dubey 2009).

12.3 Non-coding RNA (ncRNA)

In recent years, genome sequencing has revealed that the genomes of all eukaryotes studied to date are nearly totally transcribed, but only ~2 % are translated into proteins, thus generating a vast number of non-coding RNAs. A growing number of studies have shown that many of these RNAs have important regulatory functions (Amaral et al. 2008; Mercer et al. 2009). This is based on the following observations: many loci produce transcripts that are developmentally regulated; some antisense and intergenic ncRNAs can function in *trans*; numerous ncRNAs expressed in the brain are associated with specific neuroanatomical regions, cell types, or subcellular compartments; some ncRNAs show strong purifying selection and clear evolutionary conservation; some ncRNAs are transcriptionally regulated by key transcription factors such as p53, NFkB, Sox2, Oct4 and Nanog; some ncRNAs are physically associated with chromatin-modifying complexes, and their depletion leads to changes in gene expression. The evidence that ncRNAs have central regulatory roles is growing very rapidly.

ncRNAs are classified as long ncRNAs (>200 nucleotides) (Mercer et al. 2009); or short ncRNAs (<200 nucleotides and typically 20–30 nt long). Short ncRNAs have been implicated in gene silencing, leading translational repression or messenger RNA (mRNA) degradation (Jinek and Doudna 2009). Long ncRNAs are typically associated with cellular differentiation and the development of complex organisms (Amaral et al. 2008; Mattick 2007).

Many long ncRNAs mediate epigenetic changes recruiting chromatin-remodeling complexes. For example, *S. Pombe*'s *fbp*1 locus chromatin is progressively converted to an open configuration by a cascade of ncRNAs (Hirota et al. 2008). Other long ncRNAs play a role regulating transcription, and some promoters can be transcribed into long ncRNAs. For example, the human DHFR promoter is transcribed into a ncRNA, which directs the interaction with transcription factor IIB and induces the dissociation of the preinitiation complex, repressing transcription (Martianov et al. 2007). Long ncRNA are very diverse, which corresponds with a vast diversity in their function (Mercer et al. 2009).

12.4 Non-coding RNA in Apicomplexan Parasites

The Apicomplexa are a diverse group of unicellular parasites that can cause important diseases, such as malaria (*Plasmodium falciparum*), toxoplasmosis (*Toxoplasma gondii*), or cryptosporidiosis (*Cryptosporidium parvum*).

Apicomplexan parasites are single cell lower eukaryotes with a simple morphology compared to higher eukaryotes. However, these parasites have very complex life cycles, and is very likely that ncRNAs play important regulatory roles. In addition, many apicomplexans are pathogens and ncRNA could regulate the expression of virulence factors.

Little is known about ncRNAs is apicomplexan parasites (reviewed in Matrajt 2010). *Plasmodium falciparum* evades human immunity thought antigenic variation. In *Plasmodium,* the virulence *var* genes are responsible for antigenic variation which results from switches in expression between members of this gene family. The chromatin surrounding the var genes generates long sense and antisense ncRNAs that associate with the chromatin, suggesting that ncRNA may be part of the structure that regulates antigenic variation (Epp et al. 2009). Recently, Amit-Avraham et al. (2015) have shown that antisense long ncRNAs initiating from var introns associate with the active var gene, and these antisense ncRNAs trigger the activation of a silent var gene *in trans*, which demonstrates a key regulatory role for these ncRNAs (Amit-Avraham et al. 2015).

In *P. falciparum*, 24 % of all open reading frames produce natural antisense transcripts (NATs), which may regulate the epigenome (Vembar et al. 2014). Broadbent et al. (2015) investigated the expression of long ncRNAs (lncRNAs) in blood stage development of *P. falciparum*, and showed that a subset of these lncRNAs peak during invasion (Broadbent et al. 2015). In *T. gondii*, RNA sequencing data recently identified 2690 putative long ncRNAs of unknown function (Hassan et al. 2012).

In this chapter, first I will discuss the function of a ncRNA produced by *T.gondii* parasites (*Tg-ncRNA*-1) that has a major role in the critical transition from tachyzoites to bradyzoites. Second, I will discuss how *T. gondii* infection affects the expression of human and mouse microRNA, which leads to subvert host gene expression and successful establishment of these parasites in mammalian hosts. In particular, I will discuss the dysregulation of host microRNA-132 by *T. gondii* infection, and it's implications for dopamine signaling and neuropathology.

12.5 Non-coding RNA and Bradyzoite Differentiation in *T. gondii*

Tachyzoites are eliminated by the host immune response, but they can escape immune attack by differentiating into bradyzoites. Bradyzoite cysts can persist for the life time of the host, representing not only a major source for the dissemination of the infection but also a potentially life-threatening opportunistic infection of particular significance to immunocompromised patients. In spite of the importance of stage conversion, little is known about the regulation of this process.

In an effort to identify genes that are essential for cyst formation we have developed a selection scheme for parasites that are unable to differentiate and we have isolated bradyzoite differentiation mutants (Lescault et al. 2010). Insertional mutagenesis was carried out to saturate the parasite genome with a vector that integrates at random, followed by bradyzoite induction in vitro and negative selection to identify mis-regulation mutants. One of these mutants is particularly interesting because the locus disrupted encodes a transcript, named *Tg-ncRNA*-1,

that is predicted to be a ncRNA (Patil et al. 2012). Microarray hybridizations show that this mutant has a global deficiency of bradyzoite gene induction strongly suggesting that the mutation involves a step very early in the cascade that triggers differentiation (Lescault et al. 2010).

Rapid amplification of cDNA ends (RACE) carried out in the disrupted locus showed that the insertion directly disrupted a 2.6 KB transcript, *Tg-ncRNA*-1, whose expression is up-regulated 24 fold during the transition from tachyzoites to bradyzoites (Fig. 12.1b).

All ORFs present in *Tg-ncRNA*-1 are smaller than 100 amino acids, and the algorithm CPC (Coding Potential Calculator) strongly indicates that *Tg-ncRNA*-1 corresponds to a ncRNA (Patil et al. 2012). The algorithm CPC evaluates the protein-coding potential of a transcript based on six biologically meaningful sequence features (Patil et al. 2012).

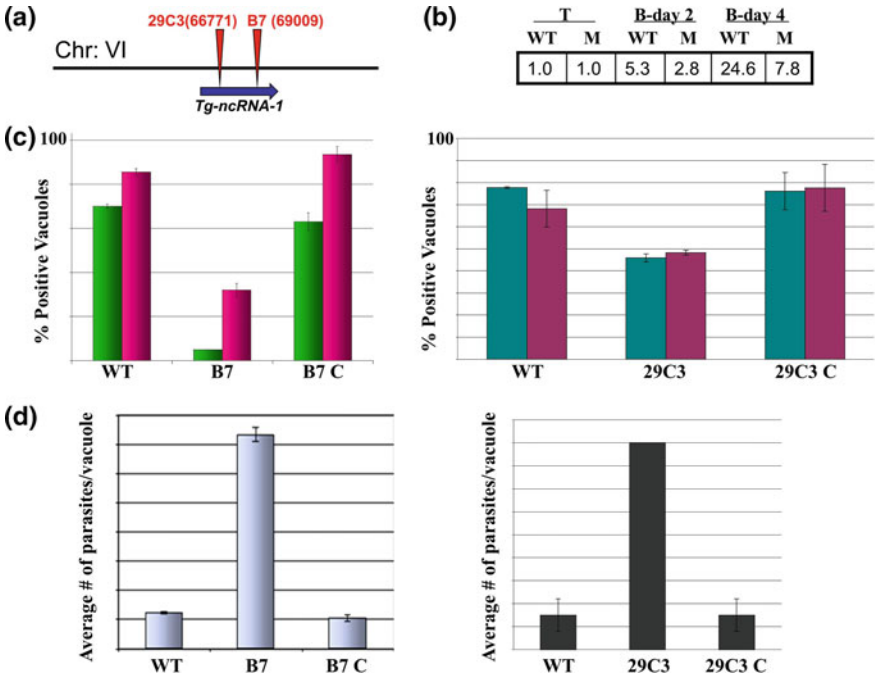


Fig. 12.1 **a** Locus disrupted in mutants B7 and 29C3. **b** Real-time PCR carried with wild-type (*WT*) and mutant B7 (*M*), 2 and 4 days post bradyzoite induction. The average of 3 experiments is shown (**c**) and (**d**) Human foreskin fibroblasts cells were infected with wild-type (*WT*), mutant (B7 and 29C3) and complemented mutant (B7C and 29C3C); and subjected to bradyzoite induction conditions for 72 h. **c** Immunofluorescnce assays were carried out for bradyzoite markers (Bag-1 and *Dolichos lectin*). The presence/absence of marker expression was counted for each parasite line, in triplicate experiments. **d** Proliferation after 72 h of induction was measured by counting the number of parasites/vacuole in triplicate experiments

Frankel et al. (2007) isolated 39 mutants with a ~10 fold reduction in the number of cysts per brain compared with infections with wild-type parasites. Interestingly, one of these mutants, mutant 29C3, has an insertion very close to the insertion observed in our mutant, named mutant B7 (Frankel et al. 2007). The same transcript (*Tg-ncRNA*-1) disrupted in mutant B7 is disrupted in mutant 29C3 at a different position (Fig. 12.1a). The fact that two different screens in two different laboratories isolated mutants with similar phenotype and the same transcript disrupted, strongly suggests that this locus is very important for cyst formation. Both mutants were complemented with the same cosmid sequence, which contains the wild-type *Tg-ncRNA*-1 transcript and flanking genomic regions. We quantified the expression of the bradyzoite markers (Bag-1 and *Dolichos lectin*) in wild-type, mutant and complemented mutant parasites (Fig. 12.1c). The complemented mutant expresses bradyzoite markers to the same level than wild-type parasites (Fig. 12.1c). Another hallmark of bradyzoite differentiation is the significant reduction of the replication rate. Quantification of the replication rate, shows that the complemented mutant slows down the replication to wild-type levels (Fig. 12.1d). These results show that wild-type *Tg-ncRNA*-1 restores the ability of the mutants B7 and 29C3 to differentiate into bradyzoites. This is the first time that a ncRNA has been implicated in the cellular differentiation of *T. gondii* (Patil et al. 2012). To date, many ncRNAs have been sequenced and characterized, but finding phenotypes associated to these ncRNAs has not been easy. In the future, more forward and reverse genetic screens designed to shed light on the function of ncRNAs will be needed to advance this field.

The mechanisms by which *Tg-ncRNA*-1 regulates tachyzoite to bradyzoite differentiation requires further investigation. Sequence analysis showed that *Tg-ncRNA*-1 contains a small RNA (sRNA). This sRNA is a 24-mer that belongs to a family named repeat-derived small RNAs (rdsRNAs) (Braun et al. 2010). This *T. gondii* family of rdsRNAs seems to be involved in transcriptional silencing (Braun et al. 2010). Therefore, one hypothesis to be tested is that members of the rdsRNA family mediate gene silencing during bradyzoite differentiation.

12.6 *T. gondii* Interferes with Host Cell ncRNA

T. gondii is an obligate intracellular parasite. Upon host cell invasion *T. gondii* secretes effectors into the host cell and subverts host gene expression (Boothroyd and Dubremetz 2008). In addition, this parasite interferes with host microRNA (miRNA) populations (Zeiner et al. 2010). As an obligate intracellular parasite, *T. gondii* profoundly reorganizes the host cell and therefore is not surprising that important gene regulatory molecules such as miRNAs are affected by these parasites. Zeiner et al. (2010) showed that *T. gondii* infected human foreskin fibroblasts

have increased expression of miR-17 ~ 92 and miR106b ~ 25 *loci*, however the impact of these changes in vivo remains to be studied (Zeiner et al. 2010).

Xiao et al. (2014) studied neural cells infected by *T. gondii* and showed that infection led to the increased expression of host cell miRNA-132. MiR-132 regulates dopamine receptor pathways (Xiao et al. 2014). Dopamine is a neurotransmitter that controls a diverse range of physiological processes, and disturbances of this pathway have been associated with many pathologies, such as Parkinson's disease, schizophrenia, attention-deficit/hyperactivity disorder, and addiction. Xiao et al. (2014) showed that infection of mice with *T. gondii* induces an increase in the host miRNA-132, which is linked to altered dopamine receptors and intracellular proteins associated with the transduction of dopamine mediated signaling. This leads to the alteration of the dopamine metabolism in *T. gondii* infected mice, resulting in elevated dopamine levels and elevated dopamine metabolites (Xiao et al. 2014). These results support the hypothesis that abnormal dopamine signaling may mediate some neuropsychiatric symptoms observed in congenital toxoplasmosis and toxoplasmic encephalitis patients.

The host miRNA-132 is not the only one affected by *T. gondii* infection in the brain. Cannella et al. (2014) showed that host cell miRNA-146a and miRNA-155 are significantly upregulated in the brains of mice infected with *T. gondii* type II strains. Both miRNA-146a and miRNA-155 have immunoregulatory roles, and this affects *T. gondii* long-term persistence in the host brain (Cannella et al. 2014).

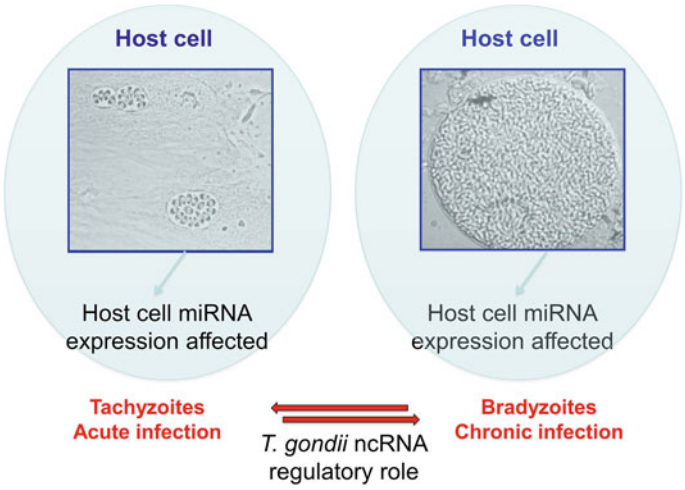


Fig. 12.2 Summary of ncRNA roles in *T. gondii* infection

12.7 Conclusions

The examples described in this chapter show that ncRNA plays central functions in the successful survival of *T. gondii* parasites in its animal host. This protozoan parasite produces its own ncRNA that helps evade the host immune response hiding as bradyzoite cysts, and in addition, *T. gondii* is able to subvert human and mouse ncRNA expression. A summary of ncRNA roles in *T. gondii* infection is shown in Fig. 12.2.

Bearing in mind that ncRNAs have central regulatory roles, and apicomplexan parasites are important pathogens, it is very likely that many ncRNAs regulate progression through the infectious cycle of these parasites, or regulate the expression of virulence factors, such as the var genes in *P. falciparum*. As intracellular parasites they need to interact with their host, and it will not be surprising to find many more host ncRNAs affected by these parasites.

Recently, sequencing data has revealed a large number of ncRNAs transcribed from apicomplexan genomes. The next challenge is to elucidate their function and mechanism of action.

References

Amaral PP, Dinger ME, Mercer TR, Mattick JS (2008) The eukaryotic genome as an RNA machine. Science 319(5871):1787–1789. doi:10.1126/science.1155472

Amit-Avraham I, Pozner G, Eshar S, Fastman Y, Kolevzon N, Yavin E, Dzikowski R (2015) Antisense long noncoding RNAs regulate var gene activation in the malaria parasite *Plasmodium falciparum*. Proc Natl Acad Sci U S A 112(9):E982–991. doi:10.1073/pnas.1420855112

Boothroyd JC, Dubremetz JF (2008) Kiss and spit: the dual roles of *Toxoplasma* rhoptries. Nat Rev Microbiol 6(1):79–88. doi:10.1038/nrmicro1800

Braun L, Cannella D, Ortet P, Barakat M, Sautel CF, Kieffer S, Garin J, Bastien O, Voinnet O, Hakimi MA (2010) A complex small RNA repertoire is generated by a plant/fungal-like machinery and effected by a metazoan-like Argonaute in the single-cell human parasite *Toxoplasma gondii*. PLoS Pathog 6(5):e1000920. doi:10.1371/journal.ppat.1000920

Broadbent KM, Broadbent JC, Ribacke U, Wirth D, Rinn JL, Sabeti PC (2015) Strand-specific RNA sequencing in *Plasmodium falciparum* malaria identifies developmentally regulated long non-coding RNA and circular RNA. BMC Genom 16:454. doi:10.1186/s12864-015-1603-4

Cannella F, Scagnolari C, Selvaggi C, Stentella P, Recine N, Antonelli G, Pierangeli A (2014) Interferon lambda 1 expression in cervical cells differs between low-risk and high-risk human papillomavirus-positive women. Med Microbiol Immunol 203(3):177–184. doi:10.1007/s00430-014-0330-9

Dubey JP (2009) History of the discovery of the life cycle of *Toxoplasma gondii*. Int J Parasitol 39 (8):877–882

Epp C, Li F, Howitt CA, Chookajorn T, Deitsch KW (2009) Chromatin associated sense and antisense noncoding RNAs are transcribed from the var gene family of virulence genes of the malaria parasite *Plasmodium falciparum*. RNA 15(1):116–127. doi:10.1261/rna.1080109

Frankel MB, Mordue DG, Knoll LJ (2007) Discovery of parasite virulence genes reveals a unique regulator of chromosome condensation 1 ortholog critical for efficient nuclear trafficking. Proc Natl Acad Sci U S A 104(24):10181–10186. doi:10.1073/pnas.0701893104

Hassan MG, Emara KM, Mohamed HA, Abdel-Wadood HM, Ikeda R, Wada M, Kuroda N, Nakashima K (2012) Determination of memantine in rat plasma by HPLC-fluorescence method and its application to study of the pharmacokinetic interaction between memantine and methazolamide. Biomed Chromatogr 26(2):214–219. doi:10.1002/bmc.1648

Hill DE, Chirukandoth S, Dubey JP (2005) Biology and epidemiology of *Toxoplasma gondii* in man and animals. Anim Health Res Rev 6(1):41–61

Hirota K, Miyoshi T, Kugou K, Hoffman CS, Shibata T, Ohta K (2008) Stepwise chromatin remodelling by a cascade of transcription initiation of non-coding RNAs. Nature 456 (7218):130–134. doi:10.1038/nature07348

Jinek M, Doudna JA (2009) A three-dimensional view of the molecular machinery of RNA interference. Nature 457(7228):405–412. doi:10.1038/nature07755

Lescault PJ, Thompson AB, Patil V, Lirussi D, Burton A, Margarit J, Bond J, Matrajt M (2010) Genomic data reveal *Toxoplasma gondii* differentiation mutants are also impaired with respect to switching into a novel extracellular tachyzoite state. PLoS ONE 5(12):e14463. doi:10.1371/journal.pone.0014463

Martianov I, Ramadass A, Serra Barros A, Chow N, Akoulitchev A (2007) Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. Nature 445 (7128):666–670. doi:10.1038/nature05519

Matrajt M (2010) Non-coding RNA in apicomplexan parasites. Mol Biochem Parasitol 174(1):1–7. doi:10.1016/j.molbiopara.2010.06.002

Mattick JS (2007) A new paradigm for developmental biology. J Exp Biol 210(Pt 9):1526–1547. doi:10.1242/jeb.005017

Mercer TR, Dinger ME, Mattick JS (2009) Long non-coding RNAs: insights into functions. Nat Rev Genet 10(3):155–159. doi:10.1038/nrg2521

Patil V, Lescault PJ, Lirussi D, Thompson AB, Matrajt M (2012) Disruption of the expression of a non-coding RNA significantly impairs cellular differentiation in *Toxoplasma gondii*. Int J Mol Sci 14(1):611–624. doi:10.3390/ijms14010611

Vembar SS, Scherf A, Siegel TN (2014) Noncoding RNAs as emerging regulators of *Plasmodium falciparum* virulence gene expression. Curr Opin Microbiol 20:153–161. doi:10.1016/j.mib.2014.06.013

Xiao J, Li Y, Prandovszky E, Karuppagounder SS, Talbot CC Jr, Dawson VL, Dawson TM, Yolken RH (2014) MicroRNA-132 dysregulation in *Toxoplasma gondii* infection has implications for dopamine signaling pathway. Neuroscience 268:128–138. doi:10.1016/j.neuroscience.2014.03.015

Zeiner GM, Norman KL, Thomson JM, Hammond SM, Boothroyd JC (2010) *Toxoplasma gondii* infection specifically increases the levels of key host microRNAs. PLoS ONE 5(1):e8742. doi:10.1371/journal.pone.0008742

Chapter 13
MicroRNAs of Filarial Nematodes: A New Frontier in Host-Pathogen Interactions

Lucienne Tritten and Timothy G. Geary

Abstract Metazoan parasites, especially nematodes, are a highly diverse group of large organisms that typically sustain an infection for long periods of time with relatively modest pathology. This state is achieved by the release of parasite-derived immunomodulatory molecules which coerce the host into providing a relatively safe niche in which the parasite is able to carry out the host-housed aspects of its life cycle. It has recently been recognized that parasitic nematodes release microRNAs (miRNAs) in culture and in mammalian hosts, primarily in exosome-like vesicles, and that these parasite-derived miRNAs may target host genes involved in the immune response. This review focused primarily on data from filarial nematodes, which occupy tissue niches in humans and other animals, and provides a perspective on possible biological roles of these molecules and their therapeutic and evolutionary implications.

13.1 Nematodes and Parasitism

Nematodes are one of the three phyla that constitute the group of animals commonly called parasitic helminths or worms. About 25,000 species of nematodes have been described (Zhang 2013), although over a million may exist (Kumar et al. 2012). Nematodes present a remarkable diversity in habitat and life histories. Some are found in water, others are terrestrial; living freely, like the model organism *Caenorhabditis elegans*, while other species have adopted parasitic lifestyles in plants or animals. Parasitism can be defined as a relationship between one (parasitic)

L. Tritten · T.G. Geary (✉)
Institute of Parasitology, Centre for Host-Parasite Interactions,
McGill University, 21,111 Lakeshore Road, Sainte-Anne-de-Bellevue,
Montreal, QC H9X 3V9, Canada
e-mail: Timothy.g.geary@mcgill.ca

L. Tritten
Institute of Parasitology, University of Zürich, Winterthurerstrasse 266,
CH-8057 Zurich, Switzerland

A.L. Leitão and F.J. Enguita (eds.), *Non-coding RNAs and Inter-kingdom Communication*, DOI 10.1007/978-3-319-39496-1_13

species diverting resources from another (host) species, the interaction benefiting the former species but being detrimental to the latter (Sorci and Garnier 2008). To reproduce, a parasite must spend all or part of its lifecycle in or on another organism. Adoption of parasitic lifestyles by nematodes is an ancient and widespread evolutionary phenomenon and presents a great diversity of life strategies (Poulin 2011; Sorci and Garnier 2008).

The origin of the phylum Nematoda is likely to be marine, but parasitism is not uncommon in these organisms and may have emerged at least 15 times independently (Blaxter and Koutsovoulos 2015). Five separate events are likely to have occurred to give rise to parasitism of mammals (Blaxter and Koutsovoulos 2015), in which they cause a range of diseases. In many cases, the involvement of an intermediate host (e.g., in an indirect life cycle) is required to transmit or facilitate the dissemination of the infective stage to the definitive host (Poulin 2011). Intermediate hosts are often arthropods (e.g., mosquitoes and flies), or gastropods (e.g., snails and slugs) (Anderson 2000).

Parasitic nematodes face multiple challenges in their lives, most notably physical barriers and immune defenses associated with host species. An evolutionary arms race underlies all host-parasite interactions in which hosts have constrained the detrimental impact of parasites, while the latter have learned how to overcome host responses to permit transmission of the next generation of parasites. This negotiation has resulted in highly specialized interspecific relationships, characterized by complex molecular dialogues between hosts and parasites and in a narrow window for an infection to be successful but not rapidly fatal. The often chronic and non-fatal nature of nematode infections, which persist for months to many years in their hosts, is due to their highly evolved abilities to hide from the host immune system and to modulate or suppress it to their own advantage (Allen and Maizels 2011).

The most well-known immunomodulatory molecules are excretory/secretory (E/S) proteins, released by helminths into their environment. These proteins facilitate the establishment of parasites in their hosts, creating a suitable environment for them to thrive. Classically, helminths evoke a skewed type 2 and regulatory immune response in the host that prevents excessive parasite-induced toxicity (inflammatory self-damage) while permitting reproductive success by the parasite. Th2 immune responses classically involve interleukin (IL)-3, IL-4, IL-5, IL-9, IL10, and IL-13, and the immunoglobulins IgG1, IgG4, and IgE. CD4 + T helper 2 cells, eosinophils, basophils, mast cells and alteRNAtively activated macrophages are cell populations that characterize Th2 responses (Allen and Maizels 2011). In parallel, helminths induce the proliferation of a regulatory T cell subset (Treg), promoting the secretion of IL-10 and TGF-β, dampening the inflammatory component of Th2 reactions (Diaz and Allen 2007; Siles-Lucas et al. 2015). For example, secretions from intestinal nematodes inhibit T cell proliferation and macrophage production of nitric oxide, harmful to the worm (Rzepecka et al. 2006). Similarly, filarial nematodes release cytokine homologs (i.e., MIF-1 of *Brugia malayi*), protease inhibitors (serpin-2, cysteine proteases, etc., of *B. malayi*), or other sorts of molecules with potent anti-inflammatory effects (i.e., ES-62 of *Acanthocheilonema viteae*), to cite only a

few examples (Hewitson et al. 2009; Pastrana et al. 1998; Pineda et al. 2014; Stepek et al. 2004; Zang et al. 1999).

In addition to proteins, there is growing evidence for a role for non-coding RNAs (ncRNA) in host-pathogen interactions and immune regulation, especially microRNAs (miRNAs), on which this chapter is focused. miRNAs are short (≈20 nucleotides) non-protein coding RNA molecules ubiquitously present in eukaryotes with important and widespread regulatory functions in gene expression. The translation of genes targeted by miRNAs is usually repressed; typically, the stability of the targeted messenger RNA (mRNA) is compromised (Bartel 2004; Griffiths-Jones et al. 2008; Rana 2007). Most miRNAs are of intergenic origin. A transcribed long primary miRNA is processed in the parent cell nucleus into a precursor miRNA. Precursor miRNAs adopt a characteristic hairpin conformation and are exported to the cytoplasm, where they are subjected to further enzymatic trimming. Mature miRNAs occur as single-stranded entities incorporated in the RNA-induced silencing complex (RISC) (Kim et al. 2009). miRNAs bind to mRNA targets, usually in the 3′ untranslated region, with partial complementarity, and can interact with multiple genes (Kim et al. 2009). The miRNA mechanism of action presents some parallels to RNA interference, relying on the same protein machinery (Rana 2007). miRNAs are cell or tissue-specific, and in humans, may reflect a pathophysiological status. miRNAs fulfill a multitude of functions by targeting many genes; it is nonetheless worth mentioning that miRNAs play key roles in the development of innate and adaptive immunity events, regulating inflammation following pathogen recognition and the development of B and T cells (Gracias and Katsikis 2011; Mobergslien and Sioud 2014).

13.2 miRNAs of Nematodes

miRNAs are not restricted to the inside of cells, as they have also been found circulating in cell-free biofluids. Extracellular vesicles (EVs), comprising exosomes (30–100 nm Ø) and microvesicles, have been identified as major vehicles for miRNA transport, together with proteins and other molecules, protecting them from degradation. miRNAs incorporated into EVs are selectively chosen for export, and do not reflect the proportions observed in the parent cell (Villarroya-Beltri et al. 2014). EVs mediate cell-to-cell communication, sometimes reaching distant recipient cells (Raposo and Stoorvogel 2013). Upon uptake and integration by a recipient cell, the miRNA cargo is functional (Chen et al. 2012; Ramachandran and Palanisamy 2012). Understandably, these discoveries raised questions about a potential role for miRNAs in mediating informational exchange between a host and parasite species. The fate of EVs has been studied in a number of host-parasite associations, summarized elsewhere by Coakley and colleagues (Coakley et al. 2015). Helminths, too, release EVs, and miRNAs found in helminth-derived EVs have been characterized (BeRNAl et al. 2014; Buck et al. 2014).

The first miRNA discovered was *lin*-4 in the early 1990s, followed a few years later by *let*-7, both initially reported in the free-living nematode *C. elegans* (Lee et al. 1993). *C. elegans* is an exceptionally valuable experimental model because its developmental life history has been extensively well-characterized, it has a known, fixed number of somatic cells with known origin, a fully sequenced and highly annotated genome with an abundant toolkit for functional genomics, and a sophisticated capacity to adapt to its environment, by for instance, arresting larval development and reducing metabolic activity (dauer state) if conditions are unfavorable (Brenner 1974; Felix and Braendle 2010). *lin*-4 and *let*-7, and many others discovered subsequently, regulate temporal development of the worm. Together with *lin*-14 and *lin*-28, *lin*-4 controls the cell lineage organization in the lateral hypodermis in the transitions from first- through third-stage larvae (L3). Similarly, *let*-7, in conjunction with *lin*-41 and *lin*-29, ensures normal developmental timing from the late L3 through the adult stage (Rougvie 2001).

Research in *C. elegans* has illuminated the general importance of miRNAs in development and physiology, and has provided mechanistic insights into the genesis and function of these key regulatory molecules (Hoogstrate et al. 2014; Lima and Pasquinelli 2014). These advances stimulated research on parasitic nematodes which analyzed miRNA populations in extracts of whole organisms, including the filarial parasites *B. malayi* (Poole et al. 2010, 2014), *B. pahangi* (Winter et al. 2012) and *Dirofilaria immitis* (Fu et al. 2013), the tissue parasite *Angiostrongylus cantonensis* (Chang et al. 2013; Chen et al. 2011; Li et al. 2014), several gastrointestinal parasites (Ahmed et al. 2013; Liu et al. 2011; Shao et al. 2014; Xu et al. 2013; Zhao et al. 2013) and plant parasitic nematodes (Ding et al. 2015; Wang et al. 2015). Assembly of the menu of parasitic nematode miRNAs is an important first step that enables characterization of their biological functions. Initial efforts in this direction have identified a *B. malayi* miRNA that is highly expressed in L3 larvae and may play a role in regulating transmission of this parasite (Winter et al. 2015). Comparative analyses of miRNAs in parasitic species also provide insight into adaptations that may be associated with parasitism; in this regard, Wang et al. (2015) identified 4 miRNAs that are found in a very diverse set of parasitic species, but are absent from *C. elegans*. The functional roles of these miRNAs should be a high priority for research.

13.3 miRNAs Released by Filarial Nematodes

Although it is important to analyze the developmental and physiological roles of miRNAs in parasitic nematodes, little progress has yet been reported in this area. Most interest to date has been devoted to the potential roles of miRNAs released by parasites in regulating the host-parasite interaction. Evidence that parasitic nematodes release miRNA-containing EVs in culture and in the host is beginning to accumulate (Buck et al. 2014; Hansen et al. 2015; Zamanian et al. 2015). Based on their life style, which includes larval and adult stage residence in inteRNAl tissues

of the host, we propose that filariae provide particularly fertile models in which to investigate the roles of miRNAs secreted by parasitic nematodes.

Filarial nematodes are important pathogens of humans and animals. They belong to the Clade III group of nematodes, in the Order Spirurida. The causative agents of lymphatic filariasis (*Wuchereria bancrofti*, *B. malayi* and *B. timori*), river blindness (*Onchocerca volvulus*) and the eye worm (*Loa loa*) are prominent examples of filarial parasitic nematodes of humans. Species of veterinary importance include the dog heartworm (*Dirofilaria immitis*), and the cow tissue parasite *O. ochengi*, among many others.

The availability of genome data is an essential prerequisite for the study of small RNAs. Over 100 nematode genomes are available, although in various stages of completion and annotation, either having been reported or as subjects of an ongoing sequencing project (www.nematodes.org). Among filarial species, the genomes of *B. malayi*, *D. immitis*, and *L. loa* are published (Desjardins et al. 2013; Ghedin et al. 2007; Godel et al. 2012), while genomes of another 11 filariae are currently in progress.

miRNAs have been sequenced in a rapidly increasing number of filarial organisms. To date, the sequence repository miRBase (release 21; (Griffiths-Jones et al. 2008)) contains 115 precursor and 107 mature high confidence miRNA sequences of *B. malayi*. miRNAs found in other filarial species are summarized in Table 13.1.

The numbers of miRNA candidates identified in the worm's environment (host blood, nodule fluid or culture media) have differed substantially among studies. For example, we identified over 200 putative parasite miRNAs in plasma obtained from dogs infected with *D. immitis*, but only 21 in plasma obtained from *O. volvulus*-infected humans (Tritten et al. 2014b). A separate study identified 6 *O. volvulus* miRNAs in human blood (Quintana et al. 2015). Circulating miRNAs in *O. ochengi*-infected cows were described twice independently, from different host biofluids, reporting 62 and 10 sequences (Quintana et al. 2015; Tritten et al. 2014a),

Table 13.1 Reported miRNAs of filarial nematodes

Species	Worm extract	Secretions	Source
B. malayi	**145** (Poole et al. 2014)	**Yes** (Zamanian et al. 2015)	Culture media
B. pahangi	**104** (Winter et al. 2012)	N/A	–
L. loa	N/A	**Yes** (Tritten et al. 2014a)	Host blood
O. volvulus	N/A	**Yes** (Quintana et al. 2015; Tritten et al. 2014b)	Host blood
O. ochengi	N/A	**Yes** (Quintana et al. 2015; Tritten et al. 2014a)	Host blood/nodule fluid
L. sigmodontis	N/A	**Yes** (Buck et al. 2014)	Host blood
D. immitis	**1063** candidates (Fu et al. 2013)	**Yes** (Tritten et al. 2014b)	Host blood

Sequences were obtained from culture media and/or host biofluids
N/A: data not available

while 22 predicted nematode miRNAs were found in plasma from *L. loa*-infected baboons (Tritten et al. 2014b) and 16 from plasma of mice infected with *Litomosoides sigmodontis* (Buck et al. 2014). Due to variations in methodological aspects, sample storage conditions and infection burdens (often incompletely known), direct comparison of these studies is difficult. However, parasite localization may contribute significantly to the variation in miRNA sequence diversity and abundance: all *D. immitis* dog stages reside in blood (adults and microfilariae), as do microfilariae of *L. sigmodontis*, while others occupy host tissues (*O. volvulus* and *O. ochengi*), with only indirect access to blood.

13.3.1 Stage-Specific miRNA Release

miRNAs are well-known to be differentially expressed across tissues and developmental stages. Several studies have reported variation of miRNA expression (up to >twentyfold) between embryonic, larval and adult stages of *C. elegans* (Karp et al. 2011; Kato et al. 2009). Larval stages in particular show an enrichment of many miRNAs. Studies in *C. elegans* have revealed variation in the expression of miRNAs related to ageing, stress-resistance, dauer state, etc. (Abbott 2011; de Lencastre et al. 2010; Karp et al. 2011; Pincus et al. 2011). With some exceptions, miRNAs in the same family (e.g., sharing a common seed sequence) tend to display similar expression patterns (Guo et al. 2014). Based on experiments with *B. pahangi* tissue extracts from L3 larvae and adults, Winter and colleagues identified 69 miRNA sequences in common to both stages, while 6 were unique to L3 and 10 to adults (Winter et al. 2012).

EVs containing small RNAs by *B. malayi* are released much more abundantly by larvae than adult males or females (Zamanian et al. 2015), despite the greatly differing amounts of worm tissue incubated in culture media to collect EVs (adults being 15–50 times larger than L3s). That study also confirmed in nematodes that miRNAs exported through EVs are present in proportions quite distinct from those in whole worm extracts. For example, Bma-*let-7* was significantly enriched in EVs released by L3s compared to its overall abundance at this stage. In support of this, *let-7* is known to regulate the L3-L4 larval transition in *C. elegans* (Rougvie 2001), while in *B. pahangi*, 3 members of the *let-7* family were shown to be up-regulated in L3 larvae as a result of invasion of the definitive host, from insect vector to mammal (Winter et al. 2015). Hence, it is hypothesized that EVs and their contents are biologically relevant and specific to infective stages (Zamanian et al. 2015).

In *D. immitis* E/S miRNA profiles established from culture media, we observed that miRNAs are abundantly produced by adult worms (especially females) in comparison to microfilariae (L1 larvae), and overlap substantially with the profile obtained from infected dog blood (manuscript in preparation). None of the most abundant miRNAs was exclusively released by one stage (not shown). The miR-100 family was among the most abundant sequences. In *C. elegans*, the conserved *mir-51/mir-100* family is an essential player in embryonic development,

specifically in the early morphogenesis of the pharynx, but also in growth and male mating (Shaw et al. 2010). Understanding the role of the miR-100 miRNAs released by parasites may not be resolvable through studies in a free-living species, but may have to await the development of more robust techniques to alter gene expression in parasites.

13.3.2 Species-Specific and Shared Filarial miRNA Sequences

miRNAs are generally thought to be highly conserved across animal species. However, a large proportion of miRNAs released from filariae appear to be species-specific (Fig. 13.1). Interestingly, only miR-71 and miR-100a were found in the secretions of all 6 examined species, miR-100c and miR-34 have been reported from all but *O. volvulus*, and miR-100d and *lin-4* from all but *L. sigmodontis*. Due to the large number of candidate miRNAs from other filariids that do not have a known homolog in the *B. malayi* secretome, its profile may be substantially different from the other 5 species. Confirming this possibility would require a detailed genomic search to determine if the other filariid miRNAs are truly absent in this species. In this context, it must be recognized that other explanations, including differences in

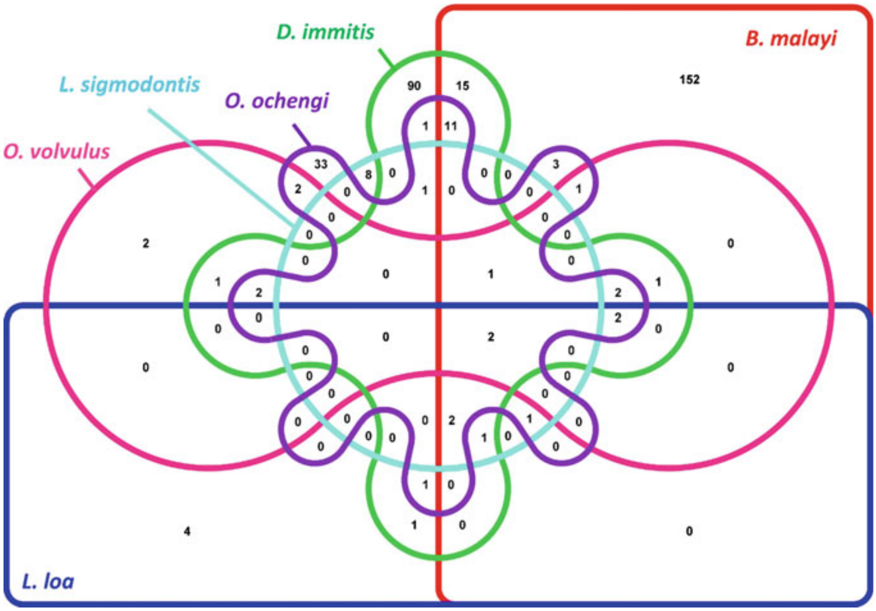


Fig. 13.1 Venn diagram of E/S miRNAs from filarial nematodes. Data are derived from the studies cited in Table 13.1. Only sequences of 18–24 nucleotides were included

the experimental milieu (i.e., host blood versus culture media), the generally unknown worm burden in vivo, the localization of the parasite in the host, and methods used for analysis and annotation may underlie this result.

Although abundant evidence supports the hypothesis that the E/S products of parasitic nematodes exert potent and profound effects on host immune response pathways, it is not yet known to what extent these molecules act at a local vs. systemic level. In this regard, parasite-derived miRNAs exported into the environment near the worm may fulfill functions that reflect the host-parasite interface at the site of infection, and diffusion of miRNAs from that site to the general circulation may be unrepresentative of the composition of the local milieu. In addition, due to the artificial nature of parasite culture systems, the miRNAs released in vitro may fail to represent accurately the miRNA profile secreted in vivo. We have begun to investigate this possibility through experiments with *D. immitis*; the population of miRNAs detected in culture media from adult males, adult females and microfilariae overlapped but were not identical to the population of *D. immitis* miRNAs detected in plasma from infected dogs (manuscript in preparation). Importantly, the most abundant miRNAs found in infected dog plasma were generally also detected in culture supeRNAtants. However, as all stages of this parasite examined so far in culture exist in the canine host bloodstream, this dataset does not directly address the local-systemic question proposed above; resolving this possible source of variation will require additional experiments in animal models of non-bloodstream filarial species. Based on currently available data, it would be premature to conclude that the suite of miRNAs so far reported from filarial parasites represents significant phylogenetic or functional differences; this is clearly an area that warrants further systematic research.

13.3.3 Predicted Host Targets of Filarial miRNAs and Cross-Species Transfer

In *C. elegans*, miRNAs possess many known functions as endogenous gene regulators, most notably fine-tuning critical transition steps in its complex life history. As noted, similar studies in parasitic species are only now beginning (Winter et al. 2015). Research in this area will be guided by the availability of bioinformatics tools that enable the identification of candidate miRNA targets, searching for perfect complementarity between the miRNA seed sequence and a mRNA, and considering aspects such as partial complementarity of the rest of the miRNA sequence, the free energy of the duplex formed by the short and the long RNA, site accessibility, position, etc. (Zheng et al. 2013). It is also common practice to further narrow down target searches using comparative analyses of sequences in orthologous species, reinforcing the statistical significance (Friedman et al. 2009; Rehmsmeier 2006; Zheng et al. 2013). A given miRNA can regulate multiple targets (Friedman et al. 2009), and clustered miRNAs are often co-expressed. There is also some evidence for a functional relationship between targeted transcripts,

i.e., in a biological pathway (Hausser and Zavolan 2014). Nevertheless, it must be kept in mind that miRNAs act in concert with transcription factors to regulate genes, and function in multi-level regulatory networks (Le Bechec et al. 2011).

It is reasonable to hypothesize that miRNAs released by worms into the host environment have been evolutionarily selected for export and function. Given the well-known immunomodulatory properties of E/S proteins, the research community rapidly envisioned that miRNAs of parasitic origin must have biological significance at the host-parasite interface. EVs from parasitic flatworms and nematodes filled with E/S products have been reported to be inteRNAlized by mammalian cells in culture (Buck et al. 2014; Marcilla et al. 2012; Zamanian et al. 2015). Of interest in this regard is that a number of nematode secreted miRNAs show identical sequences, or at least identical seed sequences, to host miRNAs. For instance, bma-*let-7*, along with bma-miR-1, bma-miR-9, bma-miR-92 and bma-miR-100b, found in EVs from *B. malayi* L3, are sequences with perfect sequence identity with human miRNAs (Zamanian et al. 2015), and hence, at least theoretically, could regulate the expression of host genes. Bma-*let*-7 was strongly enriched in EVs (Zamanian et al. 2015). Among other, diverse targets in humans, endogenous *let*-7 acts as a repressor of RAS and other oncogenes, and negatively influences cell proliferation (Johnson et al. 2007; Zamanian et al. 2015). Importantly, *let*-7 regulates several aspects of the immune system including macrophage polarization (Almanza et al. 2010; Banerjee et al. 2013; Chen et al. 2007). As algorithms that can predict miRNA–mRNA matches improve, it will be increasingly possible to identify high-quality host candidate targets for parasite-derived miRNAs (Fromm et al. 2015). Proof that parasite-derived miRNAs affect host gene regulation in situ is the next challenge facing this rapidly burgeoning field.

Recently, it has been reported that horizontal transfer of functional miRNAs can occur between parasitic nematodes and host cells in culture, with the postulate that parasites use miRNA release to mediate host immune suppression. Indeed, EVs from nematode parasites were observed to be taken up by cell lines in vitro. *B. malayi*-derived EVs were inteRNAlized by murine macrophages, via the phagocytic route, which elicited the classical activation of these immune cells (Zamanian et al. 2015), a response that was previously observed following exposure to dead or moribund worms, but which is somewhat opposed to typical immune responses to live worms or purified E/S proteins (Allen and MacDonald 1998; Osborne et al. 1996; Taylor et al. 2000). Similarly, miRNAs (miR-200, *let*-7, and miR-425; all 3 are homologous to mouse miRNAs) found in culture supeRNAtants of the murine gastrointestinal nematode *Heligmosomoides polygyrus* were predicted to interact with the 3′ UTR of the mouse gene *Dusp1* (Buck et al. 2014). DUSP1 is thought to attenuate acute inflammatory responses by inducing production of arginase by macrophages, a process shown to be protective in *H. polygyrus* infections (Anthony et al. 2006; Buck et al. 2014; Nelin et al. 2007). The study also showed the uptake of parasite EVs containing RNA and proteins, including an Argonaute homolog, by intestinal cells in vitro and provides the first evidence of immune suppression in vivo. Mice treated with *H. polygyrus* exosomes failed to mount the expected type 2 innate immune response following exposure to an allergenic fungus (Buck et al. 2014).

The immune suppression was observed to be local, as the parasite does not have direct contact with blood. In serum from mice infected with the filarial nematode *L. sigmodontis*, the adults of which reside in the pleural cavity, several putative miRNA sequences were detected, matching the parasite's genome (Buck et al. 2014). Hence, the localization of the parasite likely dictates the potential for a local versus systemic delivery of miRNAs to host cells.

13.4 Applications and Therapeutic Perspectives

Because miRNAs are present in biofluids (blood, urine, cerebrospinal fluid, sputum, etc.), stably packaged and protected in vesicles, and expressed in a tissue- or even cell-specific manner, they represent potentially valuable biomarkers for diagnostic or prognostic applications. In humans, miRNA profiles are dysregulated in all types of cancer and are tumor-specific, and thus have been widely tested for diagnostic purposes (Cortez et al. 2011). Some circulating pathogen-derived miRNAs are sufficiently distinct from host sequences to be candidate biomarkers of parasitic infections (Hoy et al. 2014; Manzano-Roman and Siles-Lucas 2012; Siles-Lucas et al. 2015). Previously, we investigated the diagnostic potential of two abundant *D. immitis* miRNAs found in dog blood, miR-71 and miR-34 (Tritten et al. 2014b). Using a stem-loop reverse transcription, quantitative PCR approach (Chen et al. 2005; Kramer 2011), we amplified *D. immitis* miR-71 and miR-34 from infected dog blood. These assays differentiated between infected and uninfected animals, but did not correlate well with the measured microfilariae counts (Tritten et al. 2014b); adult worm burdens in these dogs were unknown, and the relative proportion of these miRNAs released by microfilariae vs. adults is unknown. Nevertheless, these results suggest that an appropriately stage-specific (or stage-selective) miRNA may be of use for estimating worm burdens in infections for which such measures are difficult or impossible,

A wide range of miRNA-based therapies is currently being evaluated in human clinical trials. For example, miR-34 has been recognized as a tumor suppressor, and an injectable miRNA mimic in an ionizable liposome is undergoing clinical trials (Bader 2012). Another example is "Miravirsen", an antisense oligonucleotide that is used to sequester miR-122, a host miRNA required by the hepatitis C virus for its stability and propagation (Janssen et al. 2013).

Will miRNAs contribute one day to an improved control of nematode infections? The question has been raised by several authors (Britton et al. 2014; 2015; Schwab et al. 2015; Siles-Lucas et al. 2015). With the growing availability of parasitic nematode genomes, and affordable sequencing technologies, miRNA discovery has evolved at an increasingly rapid pace. In addition, technologies to engineer stable, chemically modified miRNA mimics, antagonists, sponges, etc., and administer them in suitable carriers are becoming increasingly available (Baumann and Winkler 2014). However, the fact that genes may be regulated by many different miRNAs raises issues of complexity in target validation, and for

efficacy in the implementation of miRNA-based treatments, as targeting a single miRNA may prove to be therapeutically insufficient (Baumann and Winkler 2014). Further substantial challenges reside in the precise targeting of administered miRNA-based treatment to the target tissue or cell population, and in obtaining a physiologically observable effect (Siles-Lucas et al. 2015).

For helminthologists, extra challenges and open questions apply. Given the parallel mode of action between RNAi and miRNA-induced repression, and the so far limited success of RNAi experiments in parasitic nematodes, miRNA-mediated strategies to interfere with parasite development or physiological function may be very challenging to develop (Britton et al. 2015; Maule et al. 2011). Besides this, proving that hosts recognize and respond to parasite EVs in situ, determining whether all parasite EVs are equally recognized, identifying how parasite miRNA populations are sorted and selected for export, and deriving a more quantitative understanding of miRNA-exosome stoichiometry, all are challenges that remain to be resolved (Chevillet et al. 2014; Coakley et al. 2015). However, for therapeutic utility, perhaps the most important challenge is to identify parasite-derived miRNAs that are essential and sufficient to achieve host immunomodulation. If multiple miRNAs contribute to the successful establishment of infection, it will be difficult to convert a normally permissible host into a non-permissible host by interfering with the functions of parasite miRNAs in host fluid compartments. If, however, a few parasite miRNAs can accomplish the necessary immunomodulation, targeting them through current approaches may enable the host to kill or eliminate the parasite immunologically.

Host miRNAs, too, have appeared to change in response to parasitic infections (El-Assaad et al. 2011; He et al. 2013). It will be crucial to define them, not only to better understand their possible role in parasite-induced pathology, but also to identify their possible roles in host defense mechanisms, and further elements in their targeted genes that may be required for the outcome of an immune response (Britton et al. 2015).

13.5 Evolutionary Considerations

It has long been known that host-parasite interactions are often quite specific or selective; this is especially true for filarial parasites. For example, the host ranges of parasites that cause onchocerciasis or lymphatic filariasis in humans is highly restricted, while the closely related parasite *D. immitis* successfully infects only canids (and felids to a limited extent). The biological bases underlying the success or failure of a parasitic nematode in a particular host remain almost entirely unknown, but it is reasonable to propose that parasite-induced immunomodulation is a key variable in determining this outcome. The molecular negotiations carried out by the parasite include the exchange of several types of molecules. While protein components of parasite E/S products have been the focus of much of this work (e.g., Hewitson et al. 2009; McSorley et al. 2013), the effects of parasite

secreted proteins have rarely been shown to be host species-specific. Since single nucleotide changes can profoundly alter miRNA-mRNA interactions, it may be more likely that selection for host-specific miRNA interactions is a key driver of the evolution of host-parasite specificity, a hypothesis that remains to be rigorously evaluated.

13.6 Considerations of Other ncRNA Species

miRNAs are arguably the most well-characterized (and the most abundant, with endogenous small-interfering (si)RNAs; (Winter et al. 2012)) type of ncRNA in filarial nematodes, but are not the only class of these molecules exported in EVs. It is worth noting that miRNAs may not be the only ncRNA species involved in host-parasite cross-talk. In *H. polygyrus*-derived exosomes, a set of Y RNAs was reported along with miRNAs (Buck et al. 2014). Also, endogenous siRNAs and piwi-interacting RNAs have been discovered in *B. pahangi* whole-worm extracts (Sarkies et al. 2015; Winter et al. 2012). There is yet no biological evidence for a role of other parasite ncRNAs in host modulation, but it would be unwise to ignore the possibility.

Acknowledgments We thank the Canadian Institutes of Health Research (Fellowship #320382 to LT), the Canada Research Chairs, the Natural Sciences and Engineering Research Council of Canada and the FRQNT-supported Centre for Host-Parasite Interactions for support.

References

Abbott AL (2011) Uncovering new functions for microRNAs in *Caenorhabditis elegans*. Curr Biol 21(17):R668–671. doi:10.1016/j.cub.2011.07.027

Ahmed R, Chang Z, Younis AE, Langnick C, Li N, Chen W, Brattig N, Dieterich C (2013) Conserved miRNAs are candidate post-transcriptional regulators of developmental arrest in free-living and parasitic nematodes. Genome Biol Evol 5(7):1246–1260. doi:10.1093/gbe/evt086

Allen JE, MacDonald AS (1998) Profound suppression of cellular proliferation mediated by the secretions of nematodes. Parasite Immunol 20(5):241–247

Allen JE, Maizels RM (2011) Diversity and dialogue in immunity to helminths. Nat Rev Immunol 11(6):375–388. doi:10.1038/nri2992

Almanza G, Fernandez A, Volinia S, Cortez-Gonzalez X, Croce CM, Zanetti M (2010) Selected microRNAs define cell fate determination of murine central memory CD8 T cells. PLoS ONE 5(6):e11243. doi:10.1371/journal.pone.0011243

Anderson RC (2000) Nematode parasites of vertebrates: their development and transmission. CABI Publishing, Wallingford

Anthony RM, Urban JF Jr, Alem F, Hamed HA, Rozo CT, Boucher JL, Van Rooijen N, Gause WC (2006) Memory T(H)2 cells induce alteRNAtively activated macrophages to mediate protection against nematode parasites. Nat Med 12(8):955–960. doi:10.1038/nm1451

Bader AG (2012) miR-34—a microRNA replacement therapy is headed to the clinic. Front Genet 3:120. doi:10.3389/fgene.2012.00120

Banerjee S, Xie N, Cui H, Tan Z, Yang S, Icyuz M, Abraham E, Liu G (2013) MicroRNA let-7c regulates macrophage polarization. J Immunol 190(12):6542–6549. doi:10.4049/jimmunol.1202496

Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116(2):281–297

Baumann V, Winkler J (2014) miRNA-based therapies: strategies and delivery platforms for oligonucleotide and non-oligonucleotide agents. Future Med Chem 6(17):1967–1984. doi:10.4155/fmc.14.116

Bernal D, Trelis M, Montaner S, Cantalapiedra F, Galiano A, Hackenberg M, Marcilla A (2014) Surface analysis of *Dicrocoelium dendriticum*. The molecular characterization of exosomes reveals the presence of miRNAs. J Proteomics 105:232–241. doi:10.1016/j.jprot.2014.02.012

Blaxter M, Koutsovoulos G (2015) The evolution of parasitism in Nematoda. Parasitology 142 (Suppl 1):S26–39. doi:10.1017/S0031182014000791

Brenner S (1974) The genetics of *Caenorhabditis elegans*. Genetics 77(1):71–94

Britton C, Winter AD, Gillan V, Devaney E (2014) microRNAs of parasitic helminths—identification, characterization and potential as drug targets. Int J Parasitol Drugs Drug Resist 4 (2):85–94. doi:10.1016/j.ijpddr.2014.03.001

Britton C, Winter AD, Marks ND, Gu H, McNeilly TN, Gillan V, Devaney E (2015) Application of small RNA technology for improved control of parasitic helminths. Vet Parasitol 212(1–2):47–53. doi:10.1016/j.vetpar.2015.06.003

Buck AH, Coakley G, Simbari F, McSorley HJ, Quintana JF, Le Bihan T, Kumar S, Abreu-Goodger C, Lear M, Harcus Y, Ceroni A, Babayan SA, Blaxter M, Ivens A, Maizels RM (2014) Exosomes secreted by nematode parasites transfer small RNAs to mammalian cells and modulate innate immunity. Nat Commun 5:5488. doi:10.1038/ncomms6488

Chang SH, Tang P, Lai CH, Kuo ML, Wang LC (2013) Identification and characterisation of microRNAs in young adults of *Angiostrongylus cantonensis* via a deep-sequencing approach. Mem Inst Oswaldo Cruz 108(6):699–706. doi:10.1590/0074-0276108062013005

Chen C, Ridzon DA, Broomer AJ, Zhou Z, Lee DH, Nguyen JT, Barbisin M, Xu NL, Mahuvakar VR, Andersen MR, Lao KQ, Livak KJ, Guegler KJ (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. Nucleic Acids Res 33(20):e179. doi:10.1093/nar/gni178

Chen MX, Ai L, Xu MJ, Zhang RL, Chen SH, Zhang YN, Guo J, Cai YC, Tian LG, Zhang LL, Zhu XQ, Chen JX (2011) *Angiostrongylus cantonensis*: identification and characterization of microRNAs in male and female adults. Exp Parasitol 128(2):116–120. doi:10.1016/j.exppara.2011.02.019

Chen X, Liang H, Zhang J, Zen K, Zhang CY (2012) Horizontal transfer of microRNAs: molecular mechanisms and clinical applications. Protein Cell 3(1):28–37. doi:10.1007/s13238-012-2003-z

Chen XM, Splinter PL, O'Hara SP, LaRusso NF (2007) A cellular micro-RNA, let-7i, regulates Toll-like receptor 4 expression and contributes to cholangiocyte immune responses against *Cryptosporidium parvum* infection. J Biol Chem 282(39):28929–28938. doi:10.1074/jbc.M702633200

Chevillet JR, Kang Q, Ruf IK, Briggs HA, Vojtech LN, Hughes SM, Cheng HH, Arroyo JD, Meredith EK, Gallichotte EN, Pogosova-Agadjanyan EL, Morrissey C, Stirewalt DL, Hladik F, Yu EY, Higano CS, Tewari M (2014) Quantitative and stoichiometric analysis of the microRNA content of exosomes. Proc Natl Acad Sci USA 111(41):14888–14893. doi:10.1073/pnas.1408301111

Coakley G, Maizels RM, Buck AH (2015) Exosomes and other extracellular vesicles: the new communicators in parasite infections. Trends Parasitol 31(10):477–489. doi:10.1016/j.pt.2015.06.009

Cortez MA, Bueso-Ramos C, Ferdin J, Lopez-Berestein G, Sood AK, Calin GA (2011) MicroRNAs in body fluids–the mix of hormones and biomarkers. Nat Rev Clin Oncol 8 (8):467–477. doi:10.1038/nrclinonc.2011.76

de Lencastre A, Pincus Z, Zhou K, Kato M, Lee SS, Slack FJ (2010) MicroRNAs both promote and antagonize longevity in *C. elegans*. Curr Biol 20(24):2159–2168. doi:10.1016/j.cub.2010.11.015

Desjardins CA, Cerqueira GC, Goldberg JM, Dunning Hotopp JC, Haas BJ, Zucker J, Ribeiro JM, Saif S, Levin JZ, Fan L, Zeng Q, Russ C, Wortman JR, Fink DL, Birren BW, Nutman TB (2013) Genomics of Loa loa, a *Wolbachia*-free filarial parasite of humans. Nat Genet 45 (5):495–500. doi:10.1038/ng.2585

Diaz A, Allen JE (2007) Mapping immune response profiles: the emerging scenario from helminth immunology. Eur J Immunol 37(12):3319–3326. doi:10.1002/eji.200737765

Ding X, Ye J, Wu X, Huang L, Zhu L, Lin S (2015) Deep sequencing analyses of pine wood nematode *Bursaphelenchus xylophilus* microRNAs reveal distinct miRNA expression patterns during the pathological process of pine wilt disease. Gene 555(2):346–356. doi:10.1016/j.gene.2014.11.030

El-Assaad F, Hempel C, Combes V, Mitchell AJ, Ball HJ, Kurtzhals JA, Hunt NH, Mathys JM, Grau GE (2011) Differential microRNA expression in experimental cerebral and noncerebral malaria. Infect Immun 79(6):2379–2384. doi:10.1128/IAI.01136-10

Felix MA, Braendle C (2010) The natural history of *Caenorhabditis elegans*. Curr Biol 20(22): R965–969. doi:10.1016/j.cub.2010.09.050

Friedman RC, Farh KK, Burge CB, Bartel DP (2009) Most mammalian mRNAs are conserved targets of microRNAs. Genome Res 19(1):92–105. doi:10.1101/gr.082701.108

Fromm B, Trelis M, Hackenberg M, Cantalapiedra F, BeRNAl D, Marcilla A (2015) The revised microRNA complement of *Fasciola hepatica* reveals a plethora of overlooked microRNAs and evidence for enrichment of immuno-regulatory microRNAs in extracellular vesicles. Int J Parasitol 45(11):697–702. doi:10.1016/j.ijpara.2015.06.002

Fu Y, Lan J, Wu X, Yang D, Zhang Z, Nie H, Hou R, Zhang R, Zheng W, Xie Y, Yan N, Yang Z, Wang C, Luo L, Liu L, Gu X, Wang S, Peng X, Yang G (2013) Identification of *Dirofilaria immitis* miRNA using illumina deep sequencing. Vet Res 44:3. doi:10.1186/1297-9716-44-3

Ghedin E, Wang S, Spiro D, Caler E, Zhao Q, Crabtree J, Allen JE, Delcher AL, Guiliano DB, Miranda-Saavedra D, Angiuoli SV, Creasy T, Amedeo P, Haas B, El-Sayed NM, Wortman JR, Feldblyum T, Tallon L, Schatz M, Shumway M, Koo H, Salzberg SL, Schobel S, Pertea M, Pop M, White O, Barton GJ, Carlow CK, Crawford MJ, Daub J, Dimmic MW, Estes CF, Foster JM, Ganatra M, Gregory WF, Johnson NM, Jin J, Komuniecki R, Korf I, Kumar S, Laney S, Li BW, Li W, Lindblom TH, Lustigman S, Ma D, Maina CV, Martin DM, McCarter JP, McReynolds L, Mitreva M, Nutman TB, Parkinson J, Peregrin-Alvarez JM, Poole C, Ren Q, Saunders L, Sluder AE, Smith K, Stanke M, Unnasch TR, Ware J, Wei AD, Weil G, Williams DJ, Zhang Y, Williams SA, Fraser-Liggett C, Slatko B, Blaxter ML, Scott AL (2007) Draft genome of the filarial nematode parasite *Brugia malayi*. Science 317 (5845):1756–1760. doi:10.1126/science.1145406

Godel C, Kumar S, Koutsovoulos G, Ludin P, Nilsson D, Comandatore F, Wrobel N, Thompson M, Schmid CD, Goto S, Bringaud F, Wolstenholme A, Bandi C, Epe C, Kaminsky R, Blaxter M, Maser P (2012) The genome of the heartworm, *Dirofilaria immitis*, reveals drug and vaccine targets. FASEB J 26(11):4650–4661. doi:10.1096/fj.12-205096

Gracias DT, Katsikis PD (2011) MicroRNAs: key components of immune regulation. Adv Exp Med Biol 780:15–26. doi:10.1007/978-1-4419-5632-3_2

Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ (2008) miRBase: tools for microRNA genomics. Nucleic Acids Res 36(Database issue):D154–158. doi:10.1093/nar/gkm952

Guo L, Zhao Y, Yang S, Zhang H, Chen F (2014) Integrative analysis of miRNA-mRNA and miRNA-miRNA interactions. Biomed Res Int 2014:907420. doi:10.1155/2014/907420

Hansen EP, Kringel H, Williams AR, Nejsum P (2015) Secretion of RNA-containing extracellular vesicles by the porcine whipworm. Trichuris suis. J Parasitol 101(3):336–340. doi:10.1645/14-714.1

Hausser J, Zavolan M (2014) Identification and consequences of miRNA-target interactions–beyond repression of gene expression. Nat Rev Genet 15(9):599–612. doi:10.1038/nrg3765

He X, Sai X, Chen C, Zhang Y, Xu X, Zhang D, Pan W (2013) Host serum miR-223 is a potential new biomarker for *Schistosoma japonicum* infection and the response to chemotherapy. Parasit Vectors 6:272. doi:10.1186/1756-3305-6-272

Hewitson JP, Grainger JR, Maizels RM (2009) Helminth immunoregulation: the role of parasite secreted proteins in modulating host immunity. Mol Biochem Parasitol 167(1):1–11. doi:10.1016/j.molbiopara.2009.04.008

Hoogstrate SW, Volkers RJ, Sterken MG, Kammenga JE, Snoek LB (2014) Nematode endogenous small RNA pathways. Worm 3:e28234. doi:10.4161/worm.28234

Hoy AM, Lundie RJ, Ivens A, Quintana JF, Nausch N, Forster T, Jones F, Kabatereine NB, Dunne DW, Mutapi F, Macdonald AS, Buck AH (2014) Parasite-derived microRNAs in host serum as novel biomarkers of helminth infection. PLoS Negl Trop Dis 8(2):e2701. doi:10.1371/journal.pntd.0002701

Janssen HL, Reesink HW, Lawitz EJ, Zeuzem S, Rodriguez-Torres M, Patel K, van der Meer AJ, Patick AK, Chen A, Zhou Y, Persson R, King BD, Kauppinen S, Levin AA, Hodges MR (2013) Treatment of HCV infection by targeting microRNA. N Engl J Med 368(18):1685–1694. doi:10.1056/NEJMoa1209026

Johnson CD, Esquela-Kerscher A, Stefani G, Byrom M, Kelnar K, Ovcharenko D, Wilson M, Wang X, Shelton J, Shingara J, Chin L, Brown D, Slack FJ (2007) The let-7 microRNA represses cell proliferation pathways in human cells. Cancer Res 67(16):7713–7722. doi:10.1158/0008-5472.CAN-07-1083

Karp X, Hammell M, Ow MC, Ambros V (2011) Effect of life history on microRNA expression during C. elegans development. RNA 17(4):639–651. doi:10.1261/rna.2310111

Kato M, de Lencastre A, Pincus Z, Slack FJ (2009) Dynamic expression of small non-coding RNAs, including novel microRNAs and piRNAs/21U-RNAs, during *Caenorhabditis elegans* development. Genome Biol 10(5):R54. doi:10.1186/gb-2009-10-5-r54

Kim VN, Han J, Siomi MC (2009) Biogenesis of small RNAs in animals. Nat Rev Mol Cell Biol 10(2):126–139. doi:10.1038/nrm2632

Kramer MF (2011) Stem-loop RT-qPCR for miRNAs. Curr Protoc Mol Biol Chapter 15:Unit 15 10. doi:10.1002/0471142727.mb1510s95

Kumar S, Koutsovoulos G, Kaur G, Blaxter M (2012) Toward 959 nematode genomes. Worm 1 (1):42–50. doi:10.4161/worm.19046

Le Bechec A, Portales-Casamar E, Vetter G, Moes M, Zindy PJ, Saumet A, Arenillas D, Theillet C, Wasserman WW, Lecellier CH, Friederich E (2011) MIR@NT@N: a framework integrating transcription factors, microRNAs and their targets to identify sub-network motifs in a meta-regulation network model. BMC Bioinformatics 12:67. doi:10.1186/1471-2105-12-67

Lee RC, Feinbaum RL, Ambros V (1993) The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75(5):843–854

Li Z, Chen X, Zen X, Liang J, Wei J, Lv Z, Sun X, Wu ZD (2014) MicroRNA expression profile in the third- and fourth-stage larvae of *Angiostrongylus cantonensis*. Parasitol Res 113 (5):1883–1896. doi:10.1007/s00436-014-3836-6

Lima SA, Pasquinelli AE (2014) Identification of miRNAs and their targets in *C. elegans*. Adv Exp Med Biol 825:431–450. doi:10.1007/978-1-4939-1221-6_12

Liu X, Song Y, Lu H, Tang B, Piao X, Hou N, Peng S, Jiang N, Yin J, Liu M, Chen Q (2011) Transcriptome of small regulatory RNAs in the development of the zoonotic parasite *Trichinella spiralis*. PLoS ONE 6(11):e26448. doi:10.1371/journal.pone.0026448

Manzano-Roman R, Siles-Lucas M (2012) MicroRNAs in parasitic diseases: potential for diagnosis and targeting. Mol Biochem Parasitol 186(2):81–86. doi:10.1016/j.molbiopara.2012.10.001

Marcilla A, Trelis M, Cortes A, Sotillo J, Cantalapiedra F, Minguez MT, Valero ML, Sanchez del Pino MM, Munoz-Antoli C, Toledo R, BeRNAl D (2012) Extracellular vesicles from parasitic helminths contain specific excretory/secretory proteins and are inteRNAlized in intestinal host cells. PLoS ONE 7(9):e45974. doi:10.1371/journal.pone.0045974

Maule AG, McVeigh P, Dalzell JJ, Atkinson L, Mousley A, Marks NJ (2011) An eye on RNAi in nematode parasites. Trends Parasitol 27(11):505–513. doi:10.1016/j.pt.2011.07.004

McSorley HJ, Hewitson JP, Maizels RM (2013) Immunomodulation by helminth parasites: defining mechanisms and mediators. Int J Parasitol 43(3–4):301–310. doi:10.1016/j.ijpara.2012.11.011

Mobergslien A, Sioud M (2014) Exosome-derived miRNAs and cellular miRNAs activate innate immunity. J Innate Immun 6(1):105–110. doi:10.1159/000351460

Nelin LD, Wang X, Zhao Q, Chicoine LG, Young TL, Hatch DM, English BK, Liu Y (2007) MKP-1 switches arginine metabolism from nitric oxide synthase to arginase following endotoxin challenge. Am J Physiol Cell Physiol 293(2):C632–640. doi:10.1152/ajpcell.00137.2006

Osborne J, Hunter SJ, Devaney E (1996) Anti-interleukin-4 modulation of the Th2 polarized response to the parasitic nematode *Brugia pahangi*. Infect Immun 64(9):3461–3466

Pastrana DV, Raghavan N, FitzGerald P, Eisinger SW, Metz C, Bucala R, Schleimer RP, Bickel C, Scott AL (1998) Filarial nematode parasites secrete a homologue of the human cytokine macrophage migration inhibitory factor. Infect Immun 66(12):5955–5963

Pincus Z, Smith-Vikos T, Slack FJ (2011) MicroRNA predictors of longevity in *Caenorhabditis elegans*. PLoS Genet 7(9):e1002306. doi:10.1371/journal.pgen.1002306

Pineda MA, Lumb F, Harnett MM, Harnett W (2014) ES-62, a therapeutic anti-inflammatory agent evolved by the filarial nematode *Acanthocheilonema viteae*. Mol Biochem Parasitol 194(1–2):1–8. doi:10.1016/j.molbiopara.2014.03.003

Poole CB, Davis PJ, Jin J, McReynolds LA (2010) Cloning and bioinformatic identification of small RNAs in the filarial nematode, Brugia malayi. Mol Biochem Parasitol 169(2):87–94. doi:10.1016/j.molbiopara.2009.10.004

Poole CB, Gu W, Kumar S, Jin J, Davis PJ, Bauche D, McReynolds LA (2014) Diversity and expression of microRNAs in the filarial parasite, Brugia malayi. PLoS One 9(5):e96498. doi:10.1371/journal.pone.0096498

Poulin R (2011) The many roads to parasitism: a tale of convergence. Adv Parasitol 74:1–40. doi:10.1016/B978-0-12-385897-9.00001-X

Quintana JF, Makepeace BL, Babayan SA, Ivens A, Pfarr KM, Blaxter M, Debrah A, Wanji S, Ngangyung HF, Bah GS, Tanya VN, Taylor DW, Hoerauf A, Buck AH (2015) Extracellular Onchocerca-derived small RNAs in host nodules and blood. Parasit Vectors 8:58. doi:10.1186/s13071-015-0656-1

Ramachandran S, Palanisamy V (2012) Horizontal transfer of RNAs: exosomes as mediators of intercellular communication. Wiley Interdiscip Rev RNA 3(2):286–293. doi:10.1002/wrna.115

Rana TM (2007) Illuminating the silence: understanding the structure and function of small RNAs. Nat Rev Mol Cell Biol 8(1):23–36. doi:10.1038/nrm2085

Raposo G, Stoorvogel W (2013) Extracellular vesicles: exosomes, microvesicles, and friends. J Cell Biol 200(4):373–383. doi:10.1083/jcb.201211138

Rehmsmeier M (2006) Prediction of microRNA targets. Methods Mol Biol 342:87–99. doi:10.1385/1-59745-123-1:87

Rougvie AE (2001) Control of developmental timing in animals. Nat Rev Genet 2(9):690–701. doi:10.1038/35088566

Rzepecka J, Lucius R, Doligalska M, Beck S, Rausch S, Hartmann S (2006) Screening for immunomodulatory proteins of the intestinal parasitic nematode *Heligmosomoides polygyrus*. Parasite Immunol 28(9):463–472. doi:10.1111/j.1365-3024.2006.00891.x

Sarkies P, Selkirk ME, Jones JT, Blok V, Boothby T, Goldstein B, Hanelt B, Ardila-Garcia A, Fast NM, Schiffer PM, Kraus C, Taylor MJ, Koutsovoulos G, Blaxter ML, Miska EA (2015) Ancient and novel small RNA pathways compensate for the loss of piRNAs in multiple independent nematode lineages. PLoS Biol 13(2):e1002061. doi:10.1371/journal.pbio.1002061

Schwab A, Meyering SS, Lepene B, Iordanskiy S, van Hoek ML, Hakami RM, Kashanchi F (2015) Extracellular vesicles from infected cells: potential for direct pathogenesis. Front Microbiol 6:1132. doi:10.3389/fmicb.2015.01132

Shao CC, Xu MJ, Alasaad S, Song HQ, Peng L, Tao JP, Zhu XQ (2014) Comparative analysis of microRNA profiles between adult *Ascaris lumbricoides* and *Ascaris suum*. BMC Vet Res 10:99. doi:10.1186/1746-6148-10-99

Shaw WR, Armisen J, Lehrbach NJ, Miska EA (2010) The conserved miR-51 microRNA family is redundantly required for embryonic development and pharynx attachment in *Caenorhabditis elegans*. Genetics 185(3):897–905. doi:10.1534/genetics.110.117515

Siles-Lucas M, Morchon R, Simon F, Manzano-Roman R (2015) Exosome-transported microRNAs of helminth origin: new tools for allergic and autoimmune diseases therapy? Parasite Immunol 37(4):208–214. doi:10.1111/pim.12182

Sorci G, Garnier S (2008) Parasitism. Elsevier, Encyclopedia in Ecology

Stepek G, Houston KM, Goodridge HS, Devaney E, Harnett W (2004) Stage-specific and species-specific differences in the production of the mRNA and protein for the filarial nematode secreted product, ES-62. Parasitology 128(Pt 1):91–98

Taylor MJ, Cross HF, Bilo K (2000) Inflammatory responses induced by the filarial nematode *Brugia malayi* are mediated by lipopolysaccharide-like activity from endosymbiotic *Wolbachia* bacteria. J Exp Med 191(8):1429–1436

Tritten L, Burkman E, Moorhead A, Satti M, Geary J, Mackenzie C, Geary T (2014a) Detection of circulating parasite-derived microRNAs in filarial infections. PLoS Negl Trop Dis 8(7):e2971. doi:10.1371/journal.pntd.0002971

Tritten L, O'Neill M, Nutting C, Wanji S, Njouendoui A, Fombad F, Kengne-Ouaffo J, Mackenzie C, Geary T (2014b) Loa loa and *Onchocerca ochengi* miRNAs detected in host circulation. Mol Biochem Parasitol 198(1):14–17. doi:10.1016/j.molbiopara.2014.11.001

Villarroya-Beltri C, Baixauli F, Gutierrez-Vazquez C, Sanchez-Madrid F, Mittelbrunn M (2014) Sorting it out: regulation of exosome loading. Semin Cancer Biol 28:3–13. doi:10.1016/j.semcancer.2014.04.009

Wang Y, Mao Z, Yan J, Cheng X, Liu F, Xiao L, Dai L, Luo F, Xie B (2015) Identification of MicroRNAs in *Meloidogyne incognita* Using Deep Sequencing. PLoS ONE 10(8):e0133491. doi:10.1371/journal.pone.0133491

Winter AD, Weir W, Hunt M, Berriman M, Gilleard JS, Devaney E, Britton C (2012) Diversity in parasitic nematode genomes: the microRNAs of *Brugia pahangi* and *Haemonchus contortus* are largely novel. BMC Genom 13:4. doi:10.1186/1471-2164-13-4

Winter AD, Gillan V, Maitland K, Emes RD, Roberts B, McCormack G, Weir W, Protasio AV, Holroyd N, Berriman M, Britton C, Devaney E (2015) A novel member of the let-7 microRNA family is associated with developmental transitions in filarial nematode parasites. BMC Genom 16:331. doi:10.1186/s12864-015-1536-y

Xu MJ, Fu JH, Nisbet AJ, Huang SY, Zhou DH, Lin RQ, Song HQ, Zhu XQ (2013) Comparative profiling of microRNAs in male and female adults of *Ascaris suum*. Parasitol Res 112(3):1189–1195. doi:10.1007/s00436-012-3250-x

Zamanian M, Fraser LM, Agbedanu PN, HaRISChandra H, Moorhead AR, Day TA, Bartholomay LC, Kimber MJ (2015) Release of small RNA-containing exosome-like vesicles from the human filarial parasite *Brugia malayi*. PLoS Negl Trop Dis 9(9):e0004069. doi:10.1371/journal.pntd.0004069

Zang X, Yazdanbakhsh M, Jiang H, Kanost MR, Maizels RM (1999) A novel serpin expressed by blood-borne microfilariae of the parasitic nematode *Brugia malayi* inhibits human neutrophil serine proteinases. Blood 94(4):1418–1428

Zhang ZQ (2013) Animal biodiversity: an outline of higher-level classification and survey of taxonomic richness (Addenda 2013). Zootaxa 3703:1–82. doi:10.11646/zootaxa.3703.1.1

Zhao GH, Xu MJ, Zhu XQ (2013) Identification and characterization of microRNAs in *Baylisascaris schroederi* of the giant panda. Parasit Vectors 6:216. doi:10.1186/1756-3305-6-216

Zheng H, Fu R, Wang JT, Liu Q, Chen H, Jiang SW (2013) Advances in the techniques for the prediction of microRNA targets. Int J Mol Sci 14(4):8179–8187. doi:10.3390/ijms14048179

Chapter 14
Human Fungal Infections: Emerging Role of Small Non-coding RNAs as Modulators of Host-pathogen Interactions

Ana Lúcia Leitão, Marina C. Costa and Francisco J. Enguita

Abstract Fungal infections are an increasing health problem. The intrinsic variability and complexity of pathogenic fungi and the unmet clinical need for new and more effective treatments requires a detailed knowledge of the physiology and molecular determinants of fungal infections. Pathogenic fungi are able to produce modulatory non-coding RNAs by using specific biosynthetic machinery that regulate several cellular processes as morphological differentiation or stress response. Small non-coding RNAs (sncRNAs) are key players in the regulation of fungal cell metabolism. They are typically generated from repetitive genome regions, genes encoding ribosomal RNA or DNA damaged loci, but they can be also produced from specific transcripts that resemble the animal and plant miRNAs. The importance of these tiny regulators in the host-pathogen interaction is starting to be unveiled, but the already acquired evidences showed a relevant role of sncRNAs in fungal biology, including the regulation of cell morphogenesis and phase transition, control of genome stability and even epigenetic mechanisms triggering resistance to antifungal, agents. Interestingly, there is a transcriptomic cross-talk between the pathogen and host, demonstrated by a specific transcriptional program triggered in the host-cell by the presence of the pathogenic agent.

A.L. Leitão (✉)
MEtRICs, Departamento de Ciências e Tecnologia da Biomassa, Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa, Campus de Caparica, 2829-516 Caparica, Portugal
e-mail: aldl@fct.unl.pt

M.C. Costa · F.J. Enguita (✉)
Universidade de Lisboa, Faculdade de Medicina, Av. Prof. Egas Moniz, 1649-028 Lisbon, Portugal
e-mail: fenguita@medicina.ulisboa.pt

A.L. Leitão and F.J. Enguita (eds.), *Non-coding RNAs and Inter-kingdom Communication*, DOI 10.1007/978-3-319-39496-1_14

14.1 Introduction

As infectious agents, fungi have been globally disregarded in comparison with bacterial and viral pathogens. Among the 1.5 million of fungal species, approximately 3000 are able to produce infections in humans and they belong to four main groups: zygomycetes, ascomycetes, deuteromycetes and basiomycetes (Fig. 14.1). These organisms are common inhabitants of natural environments or belong to the surface microbiota of animals and plants, and under the adequate conditions can colonize human tissues leading to infections. Their condition of pathogens contributes to their adaptability to several environments, including the human body, being difficult to detect and diagnose. Moreover, their eukaryotic nature has also conditioned the currently available therapeutic strategies, which are often toxic for the host, not reliable, and negatively influenced by the appearance of resistances. The growth of highly lethal opportunistic fungal infections in recent decades, including invasive mycoses, is attributed mainly to an increased number of patients with severe immunosuppression (Kriengkauykiat et al. 2011; Nucci and Marr 2005). When AIDS is considered as a death cause, associated fungal opportunistic infections such as coccidioidomycosis and cryptococcosis represent 50.9 % of cases, followed by candidiasis (30.2 %), and histoplasmosis (10.1 %). Among the pathogens that cause invasive fungal infections, *Candida* and *Cryptococcus* species are predominant in immunocompromised individuals (Colombo 2000; Tortorano et al. 2012).

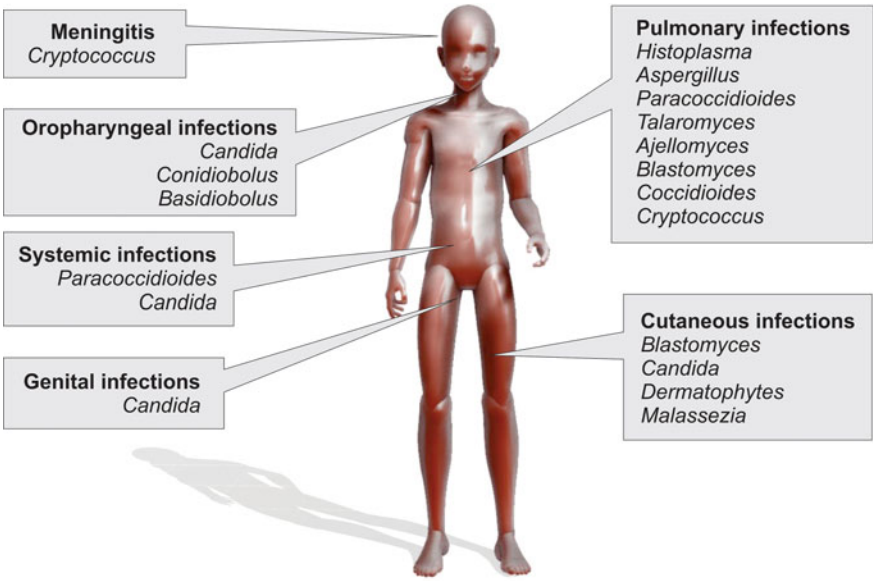


Fig. 14.1 Fungal geni producing infections in humans showing their main target organs or tissues and their infection capabilities

Prevention of a fungal disease depends on a deeper knowledge of the etiological agent, its virulence factors and its interaction with the host. Fungal infections can be considered as changes in the natural life cycle of certain fungi which in most cases require specific host determinants that allow its development within the cells (Casadevall 1998; Casadevall and Pirofski 1999). These infections have attracted increasing attention from the scientific community, due to the great inefficiency in the treatment of their systemic forms. The close functional similarity between the fungal and mammalian cells has classically limited the treatment strategies, which are still based on the use of inhibitors of the ergosterol biosynthesis (polyene and azole compounds, such as Amphotericin B), which are typically associated with the generation of significant side effects and the emergence of microbial resistance (Mikulska et al. 2011). Factors related to the immune status of the host also contribute directly to the development of fungal infections, evidenced by the direct association between immunodeficiency and occurrence of systemic mycoses (Pasqualotto et al. 2008). Thus, the search for alternative therapies and/or development of more specific antifungals has become increasingly urgent (Calderone et al. 2014). This could be achieved by a deeper knowledge of the biology of the infection with the introduction of new approaches and the search of new factors. Following this approach, host epigenetic factors and regulators can be a future target for the development of such strategies (Enguita et al. 2016).

14.2 Fungal Pathogens Causing Infections in Humans

The parasitic relationship between fungi and humans is a demanding process for the infecting microorganism. In many occasions, host colonization is depending on morphological transitions between unicellular and filamentous living forms. Filamentous fungi are typically more resistant to chemical and biological stress, but they fail to circumvent the immunological system of humans. On the other hand, the unicellular yeast-like forms are preferred by pathogenic fungi as an efficient way to colonize host tissues and spread the infection (Kohler et al. 2015).

Pathogenic fungi can be classified in two main groups, depending on their capabilities to infect healthy humans or only immunocompromised individuals. Fungi that infect healthy humans belong to three phyla: Entomophthoromycota, Ascomycota and Basidiomycota. As other obligatory pathogens, they have developed a myriad of strategies to circumvent the immune system and colonize first the surface and after the inner organs of their hosts. The opportunistic pathogens belong to species included in the phyla Basidiomycota, Mucorales and Ascomycota. They preferentially infect immunocompromised individuals, taking advantage of the immunological depression caused by other conditions like HIV infection or pharmacological treatments associated to xenotransplantation.

14.2.1 Fungal Pathogens Able to Infect Healthy Individuals

The geni *Basidiobolus* and *Conidiobolus*, belonging to the Entomophthoromycota phylum, are obligatory pathogens responsible for submucusal infections mainly localized in the nose and central face (Ubiali et al. 2013). Their colonization route starts from the inhalation of vegetative forms, that can colonize the nasal mucosa and upper respiratory tract, due to their ability to produce lytic enzymes such as proteases and lipases (Jayawickrama et al. 2012).

The phylum Ascomycota also includes several geni of pathogenic fungi able to infect healthy individuals, namely *Ajellomyces*, *Histoplasma*, *Paracoccidioides*, *Lacazia*, *Blastomyces*, and *Coccidioides*. The infection strategies of these organisms are similar to those observed in some protozoa, starting from a mucosal colonization and a rapid spread through the body via macrophage infection. *Histoplasma capsulatum*, is the responsible for histoplasmosis, a life-threatening condition which starts by the inhalation of microconidia or hyphal fragments of the fungus. After the colonization of lung epithelia, the evolution of respiratory disease depends on the *Histoplasma* ability to survive and replicate within alveolar macrophages. Subsequently to pulmonary infection and depending on the immune status of the host, vegetative fungal forms can spread to other organs belonging to the mononuclear phagocytic system, mainly the spleen and liver, causing the most severe form of histoplasmosis (Deepe and Gibbons 2009; Kroetz and Deepe 2012). The role of cellular immunity in protection against *H. capsulatum* has been well described by several authors (Kroetz and Deepe 2012). Thus, in immunocompromised patients, especially in HIV-positive ones, histoplasmosis is opportunistic, in most cases widespread and, if untreated, leads to death in almost 100 % of cases (Adenis et al. 2014; Bonifaz et al. 2011). Another fungus of extreme clinical relevance is *Paracoccidioides* spp. responsible for paracoccidioidomycosis, a disease of great social impact that primarily affects men and rural workers, and ranks as the seventh cause of death among infectious chronic diseases in countries as Brazil (Prado et al. 2009). *P. brasiliensis*, an endemic south American species, uses a colonization strategy based on the invasion of macrophages where fungal cells are able to survive and replicate being carried throughout the body and protected from the immune system (Lima Pde et al. 2014; Silva et al. 2008).

Fungi able to infect healthy humans included in the Basidiomycota phylum are represented by the genus *Cryptococcus*. Infection by *Cryptococci* occurs through inhalation of infectious particles (basidiospores) or desiccated yeasts present in the environment, which, thereafter colonize the alveolar tissue (Velagapudi et al. 2009). In lung tissue they are able to remain in a latency stage or manifest themselves by variable signs and symptoms, which can vary from asymptomatic patients to cases of severe pneumonia and respiratory failure (Brizendine et al. 2013). In healthy individuals, the infection is effectively counteracted by pro-inflammatory immune T-cell response, but in immunocompromised patients, the yeast spreads easily

through the blood, colonizing various organs, being central nervous system (CNS) its major target (Brizendine et al. 2013). The CNS infection is the main clinical manifestation of cryptococcosis, being fatal in the absence of treatment, with the development of frameworks of meningitis, encephalitis and meningoencephalitis. It is believed that one of the ways in which *Cryptococcus* reaches the CNS crossing the blood-brain barrier is within macrophages (Casadevall 2010; Vu et al. 2014).

14.2.2 Opportunistic Fungal Pathogens

These organisms belong to the natural saprophytic flora of animal and plants, and under the adequate conditions can colonize immunodepressed individuals. A relevant example of this group is constituted by the dimorphic yeast *Candida*, and its more prevalent species as opportunistic pathogens *C. albicans*, *C. parapsilosis* and *C. glabrata*. Long-term studies have demonstrated that these saprophytic organisms have co-evolved together with mammals and are extremely well adapted for the colonization of tissues, using several strategies ranging from the production of drilling hyphal structures and specific surface adhesins to the biosynthesis of lytic enzymes (Wang 2015; Brunke and Hube 2013). Systemic candidiasis is the most severe manifestation of *Candida* infection that only affects immunodepressed patients.

Aspergillosis is an infection mainly produced by *Aspergillus fumigatus*, a filamentous opportunistic fungus belonging also to the Ascomycota phylum. *A. fumigatus* is a versatile saprophytic organism inhabitant of plant surfaces. Its mycelium can give rise to reproductive structures called conidiophores, which contain small reproductive cells or conidiospores that can be easily transported and spreaded by the wind. The conidiospores can be inhalated and, depending on the immune status of the host, they can give rise to a mucosal colonization by interaction with the alveolar epithelial cells and a further systemic spreading, taking advantage of its thigmotropism (ability to sense and follow contours) that guide the hyphae throughout the blood vessels (Bowen et al. 2007). If not treated properly, systemic Aspergillosis can be a life-threatening condition, especially in environments with high-selective pressure as the intensive care units (Tortorano et al. 2012).

Other opportunistic fungi responsible for infections in humans include species from the Ascomycota geni *Fusarium* and *Talaromyces*, the basidiomycota geni *Malasezzia* and *Trychosporon*, and the Mucorales geni *Mucor* and *Rhizopus*. All of them are etiological agents of opportunistic infections with diverse severity depending on the immunological status of the host. Among them the *Fusarium* species *F. oxysporum* and *F. solanum* are the more frequent which often produce lethal infections in severe immunocompromised patients (Girmenia et al. 1992).

14.3 Small ncRNAs in Fungal Pathogens

14.3.1 Biosynthetic and Functional Machineries for Small ncRNAs in Fungi

The family of small non-coding RNAs (sncRNAs) comprises eukaryotic regulators of gene expression, often involved in the control of important cellular processes and globally connected to the RNA-interference phenomenon (RNAi). Endogenous sncRNAs include micro-RNAs (miRNAs), small interfering RNAs (siRNAs), PIWI-interacting RNAs (piRNAs) and primal-small RNAs (priRNAs) (Ghildiyal and Zamore 2009). The biosynthesis of all these non-coding species required the presence of devoted gene transcripts produced from specific genome locations and also the existence of enzymes designed to trim, select and functionalize the RNA precursors (Miyoshi et al. 2010). Globally, the origin of sncRNAs and RNAi appears to predate the emergence of eukaryotic organisms, but some of them such as miRNAs appeared to be specific of higher eukaryotes being absent in fungi (Guerra-Assuncao and Enright 2012). Interestingly, the enzymes involved in the biogenesis of sncRNAs have functional domains which are conserved throughout unicellular and multicellular organisms (Murphy et al. 2008). The proteins involved in the action of sncRNAs belong to two different groups: processors and effectors (Costa et al. 2012). Processing enzymes are RNA-dependent nucleases that can act in the nucleus or the cytoplasm, trimming precursor transcripts to generate sncRNAs. On the other hand, the effectors are proteins responsible for the final regulatory action of the sncRNAs, and are represented by the family of Argonaute proteins (Ago) (Costa et al. 2012).

In fungi, the proteins involved in the processing and action of sncRNAs were firstly characterized in *Neurospora crassa*, as responsible for a post-transcriptional gene silencing phenomenon called "quelling" (Romano and Macino 1992), a process closely related with plant co-supression (Napoli et al. 1990) and animal RNA interference (RNAi) (Fire et al. 1998). Quelling and other post-transcriptional silencing processes depend on the presence of sncRNAs that are targetted to a specific RNA transcript to facilitate its inactivation. Mutational screening allowed to identify three proteins (qde: quelling defective mutants), namely QDE-1, an RNA-dependent RNA-polymerase (RdRp), QDE-2, a dsRNA-binding protein similar to mammalian Argonautes, and QDE-3 a RecQ helicase required for productive quelling (Cogoni and Macino 1999a, b). Other required proteins for quelling in *N. crassa* were two ortologs of Dicer nucleases, designated as dicer-like proteins (DCL-1 and DCL-2) (Catalanotto et al. 2004).

This machinery has evolved from bacterial orthologs to ensure the maintenance of the genomic integrity by counteracting the action of viral pathogens, transposons or genomic DNA damage (Chicas et al. 2005). The transcription from transposon or damaged DNA loci produces the so called "aberrant RNA" (aRNA), which is detected by the fungal RNAi machinery. QDE1 helicase will recognize the aRNA and the RNA-dependent RNA polymerase QDE1 will use it as a template for the

production of a double-stranded RNA (dsRNA) (Lee et al. 2010a). Once produced, the dsRNA will be the substrate for the Dicer-like nucleases (DCL1 or DCL2) that will trim the dsRNA in small fragments to generate the final interference RNAs. The Argonaute protein QDE2 will be responsible for the final action of those sncRNAs, engaging them into an interference complex similar to the RISC (RNA interference silencing complex) from higher eukaryotes (Fig. 14.2).

The components of the biosynthetic machinery responsible for the production of sncRNAs are also present in pathogenic fungi. However the variability and complexity of fungal kingdom is patent in the organization of the genes encoding for the RNAi machinery in fungal pathogens (Fig. 14.3). We have analyzed the presence and protein sequences of the RNAi machinery components, namely Dicer-like proteins (DCL), Argonaute proteins (AGO) and RNA-dependent RNA-polymerases (RdRp) in ten species of pathogenic fungi using sequence data deposited in the funRNA database, a genome centered resource that compiles all the components of the RNAi machinery in the whole fungal kingdom (Choi et al. 2014). The presence of a complete biosynthetic machinery for RNAi is exclusive of filamentous complex fungi, being absent in yeast. However, recent studies have described a non-canonical enzymatic system in the yeast *Candida albicans* that includes a minimalist version of a Dicer protein able to trim dsRNA molecules (Bernstein et al. 2012). Filamentous pathogenic fungi have tipically two genes encoding Dicer-like proteins, with the exception of *Rhizopus oryzae* that has only one. Functional variability among pathogens is better illustrated with the number of Argonaute and RdRp proteins encoded in their genomes. *Fusarium oxysporum* and *R. oryzae* have five different RdRp proteins, harboring the most complex RNAi systems in all the analyzed genomes. Regarding the Argonaute proteins, the average number per genome is three. The variability in the number of components of the RNAi machinery suggests different levels of evolutionary and functional complexity among all the analyzed pathogenic fungi, but their physiological roles are far to be understood since there is no apparent correlation between the number of enzymatic components of the RNAi system and the variety of sncRNA produced by each selected strain. This fact is probably a cause and not a consequence of the evolutionary stage of fungi, which are in the interface between unicellular organisms and complex living forms (Lee et al. 2007; Richards 2011).

14.3.2 Families of Regulatory sncRNAs in Fungal Pathogens

The extensive use of next-generation sequencing techniques for the analysis of fungal genomes and transcriptomes has pointed out the existence of many classes of sncRNAs. Among them, the most abundant are the small ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), and small nucleolar RNAs (snoRNAs) which together constitutes 75–90 % of the total small RNAs depending on the analyzed

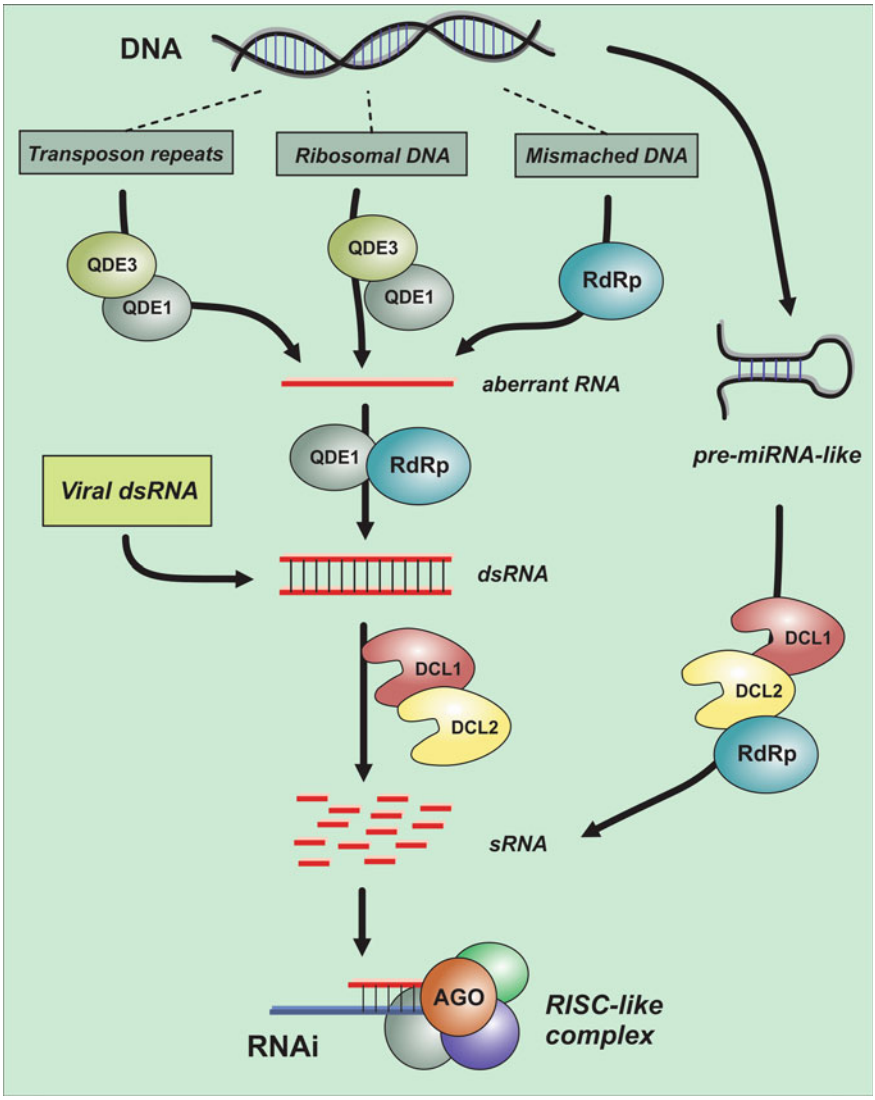


Fig. 14.2 Different flavors of sncRNAs and their biosynthetic pathways in pathogenic fungi. Some sncRNAs are generated from repetitive regions, rRNA genes or damaged DNA loci. In all these cases, the biosynthesis of sncRNAs involved the production of an aberrant RNA (aRNA). The action of RNA-dependent RNA-polymerases over the aRNA will produce a dsRNA (QDE1 or RdRp), which will be the main substrate for the Dicer-catalyzed trimming. In fungi there are two genes encoding Dicer-like proteins (DCL1 and DCL2). Short RNA species will be then guided to their targets by the action of Argonaute proteins (AGO or QDE2), and incorporated into a RISC-like complex (RNA Induced Silencing Complex). The production of small dsRNAs can be also a consequence of the presence of infecting viral RNA. In some pathogenic fungi, specific transcriptional units can generate RNAs harboring stem-loop structures that can be directly recognized by DLC1 or DCL2 to produce sncRNAs

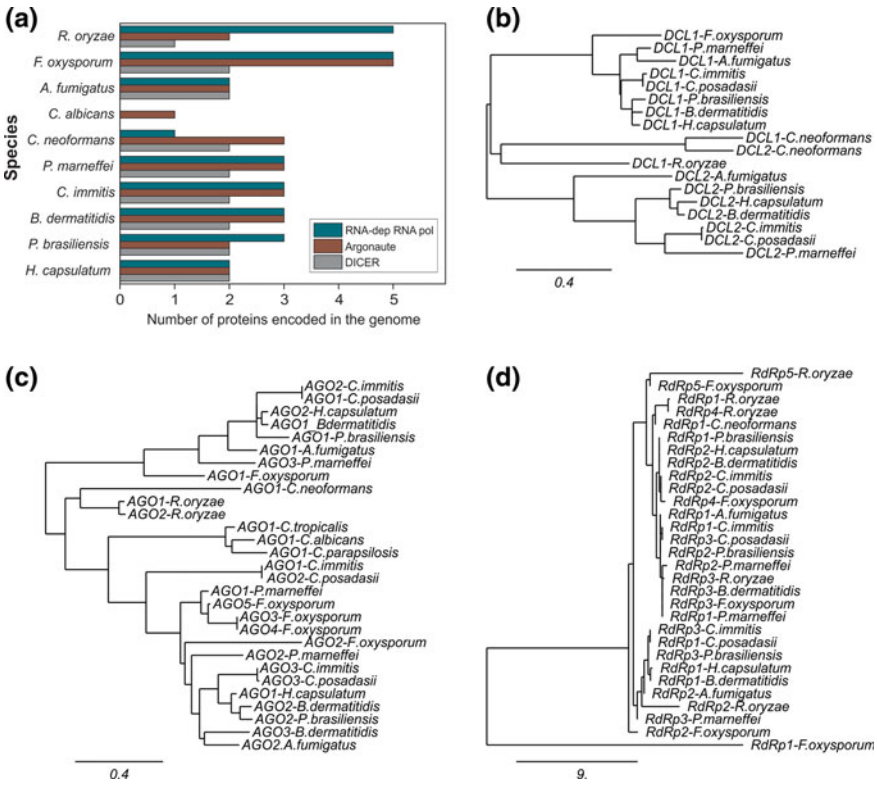


Fig. 14.3 Analysis of the RNAi machinery in selected genomes of pathogenic fungi. A, number of proteins related with RNAi within the fungal genomes; B, phylogenetic relationships among the Dicer-like proteins (DCL); C, phylogenetic relationships among the Argonaute proteins (AGO); D, phylogenetic relationships among the RNA-dependent RNA-polymerases (RdRp)

fungal transcriptome (Liu et al. 2013; Yin et al. 2012). The remaining small RNA transcriptome includes RNA families that are generated by enzymatic trimming of specialized transcripts, namely the small interfering RNAs (siRNAs) and the micro-RNAs (miRNAs) (Chen et al. 2014; Li et al. 2010).

14.3.2.1 Small Interfering RNAs (siRNAs)

As already discussed before, the generation of siRNAs is dependent on the production of aberrant RNA molecules (aRNAs) that can be generated from highly repetitive DNA regions or rRNA genes (Lee et al. 2010a) and also by the infection of a dsRNA virus (Segers et al. 2007). The generation of siRNAs from longer transcripts has been characterized in several fungal species as a mechanism to ensure genomic stability. In fact the main sources for siRNAs in fungi are

transposons and viral RNA, that can be detected and silenced by the generation of siRNAs (Dumesic et al. 2013). Moreover, the transcription of damaged DNA regions will also produce aRNAs, which will trigger the production of a specific class of sncRNAs called qiRNAs. The RNAi machinery together with the qiRNAs constitutes a primitive version of the eukaryote DNA damage response mechanism which links the presence of DNA lesions with the negative regulation of the transcription from the damaged DNA territories (Lee et al. 2009).

In pathogenic fungi the presence of a functional RNAi system, and the generation of siRNAs is also important for their viability, however its roles in fungal pathogenicity and communication with the infected host are still starting to be unveiled. Functional RNAi systems have been described in the pathogens *H. capsulatum* (Rappleye et al. 2004), *C. neoformans* (Janbon et al. 2010), *P. brasiliensis* (Goes et al. 2014), *Talaromyces marneffei* (formerly *Penicillium marneffei*) (Sun et al. 2014), *C. albicans* (Drinnenberg et al. 2009), *Mucor circinelloides* (Cervantes et al. 2013), and *R. oryzae* (Gheinani et al. 2011). However, only in *C. neoformans* and *M. circinelloides* a physiological role of siRNAs has been characterized. In the remaining species, the functionality of the RNAi system has been demonstrated only by the silencing of specific gene transcripts using synthetic dsRNA oligunucleotides (Sun et al. 2014).

C. neoformans is one of the most prevalent etiological agents of respiratory infections and severe meningitis in immunodepressed hosts. *C. neoformans* is a very peculiar dimorphic fungus, able to undergo a morphological transition from yeast to hyphae during opposite sex mating and unisexual reproduction (same sexual mating) (Wang et al. 2013). Initial studies of the possible role of the RNAi machinery in this fungal pathogen established a clear role of siRNAs in the genome maintenance and stability, especially against transposons, but neglected the possible role of these regulatory RNAs in the production of virulence factors, stress response or sexual differentiation (Janbon et al. 2010). However, a paper by Wang and coworkers published in the same year, described a detailed mechanism involving silencing of transposable elements and used by *C. neoformans* to control the sexual cycle and to ensure genome stability during this process (Wang et al. 2010). Using next-generation sequencing analysis, a cohort of siRNAs mapping transposable elements were identified, and the function of these siRNAs related with the dynamic silencing of specific transposons along the sexual cycle (Wang et al. 2010). Further studies completed the description of a sex-induced silencing (SIS) pathway in *C. neoformans* by siRNAs during sex mating (Wang et al. 2013). Since morphological transitions in *C. neoformans* are associated with the infection capabilities of the fungus, the detailed knowledge of this regulatory system opens new possibilities for advanced therapeutics against the infection.

Interestingly, for some specific genomic loci the transposon silencing mediated by siRNAs in *C. neoformans* is ensured by the presence of a protein complex designated as SCANR (spliceosome-coupled and nuclear RNAi complex) that it is physically associated with the spliceosome (Dumesic et al. 2013). The SCANR complex is the center of a genome-protecting mechanism against transposons that relies on the sub-optimal nature of the splicing signals present in the transposed loci

that induces the accumulation of spliceosomes in specific genome locations. The intramolecular competition between the spliceosome and the SCANR complex causes an accumulation of stalled spliceosomes unable to perform a productive splicing, and this situation favors the production of aberrant RNAs and the biosynthesis of siRNAs that will silence the locus from where they are generated (Dumesic et al. 2013).

M. circinelloides is an emerging human opportunistic pathogen, whose RNAi machinery has been extensively characterized (Cervantes et al. 2013). In this pathogen the generation of small RNAs is ensured by the presence of a minimalist processing system comprised by a Dicer-like protein, a single Argonaute protein and two RNA-dependent RNA polymerases. *M. circinelloides* is able to use siRNAs to regulate gene expression by two different mechanisms: one dicer-dependent and canonical pathway, and another dicer-independent mechanism which relies on the presence of a specific RNAse, R3B2. This RNAse III-like protein is able to degrade specific RNA transcripts by interacting with Argonaute and guided by siRNAs. The presence of both siRNA pathways have been considered as the first evolutionary step towards the development of advances RNAi machineries, and a genetic link between the mRNA degradation and post-transcriptional gene silencing (Trieu et al. 2015). In fact, some of the sncRNAs produced by *M. circinelloides* map to exons and regulate the expression of the genes from which they derive. Expression of 50 % of genes is dependent on one or more sncRNAs, which is in agreement with the existence of several classes of sncRNAs that regulate different cellular processes, including response to stress or sexual interaction and could be important for the infection capabilities of the microorganisms (Nicolas et al. 2015).

Moreover, the RNAi mechanism can be used by *M. circinelloides* to evoke resistance against antifungal drugs. In fact, it develops spontaneous resistance to the antifungal drug FK506 (tacrolimus) via two distinct mechanisms. One involves Mendelian mutations that confer stable drug resistance; the other occurs via an epigenetic RNA interference (RNAi)-mediated pathway resulting in unstable drug resistance (Calo et al. 2014). When exposed to tacrolimus, the peptidylprolyl isomerase FKBP12 binds to the antibiotic, forming a complex that inhibits the protein phosphatase calcineurin. This inhibition will block the transition from unicellular forms to hyphae, and increases the growth of the fungus in yeast form. Mutation of calcineurin or FKBP12 genes will confer resistance to the antibiotic and restore hyphal growth. Under these conditions, the fkbA gene locus encoding the FKBP12 enzyme will generate sncRNAs to silence the original fkbA gene and produce resistant epimutants, which can be reverted if the antibiotic pressure is released. The generation of siRNAs from fkbA gene is ensured by the use of the native mRNA and an antisense RNA (Calo et al. 2014). To date, this is the only demonstrated example of the generation of resistances to antifungals by epigenetic mechanisms involving siRNAs. However, we cannot disregard the possibility of finding similar

mechanisms in other fungi, since the RNAi machinery is widespread in this family of pathogens.

14.3.2.2 miRNA-like Small RNAs

Some highly evolved fungi contain specific transcriptional units that can give rise to transcripts containing stem-loop structures, similar to those observed in the miRNA precursors in mammals and plants (Guerra-Assuncao and Enright 2012). The stem-loops contain stretches of dsRNA that can be recognized and processed by the dicer-like enzymes DCL1 and DLC2 (Fig. 14.2), generating sncRNAs that have been designated as miRNA-like RNAs, since they are not produced by the canonical machinery present in mammals and plants (Lee et al. 2010b). It is assumed that the regulatory mechanism exerted by these miRNA-like RNAs is similar to the observed in higher eukaryotes, but the rules governing target recognition and pairing are still not completely understood.

The presence of putative precursors for miRNAs in fungi was firstly predicted by secondary structure analysis of specific conserved transcripts originated from the genome of several *Aspergillus* species (McGuire and Galagan 2008), and later demonstrated in other fungi. In *T. marneffei*, a thermal dimorphic fungus causing systemic mycoses mainly in Asia, high-throughput transcriptomic analysis allowed to identify two genomic loci responsible for the production of two miRNA-like transcripts designated as PM-milR-M1 and PM-milR-M2. The generation of mature miRNA-like RNAs is dependent of DLC2 enzyme, showing differential expression in the different growth phases of the microorganism (Lau et al. 2013). Also in *C. neoformans* recent studies identified two miRNA-like genes (miR1 and miR2) producing mature RNAs of 22 and 18 nt in length respectively. Bioinformatic analysis of *C. neoformans* genome sequence revealed a number of target sites for those miRNAs which are located within transposable elements and pseudogenes, suggesting the possible silencing activity of these sncRNAs over unstable genomic elements (Jiang et al. 2012).

Other pathogenic fungi such as *Aspergillus flavus* revealed a more complex regulatory network based on the presence of miRNA-like sncRNAs. Recent studies allowed characterizing more than 100 miRNA-like RNAs in *A. flavus*, that showed differential expression patterns in response to water activity and temperature. The stress-responsive expression of miRNAs strongly suggested their implications in mycelial growth and mycotoxin production, constituting an additional virulence factor (Bai et al. 2015). A combination of bioinformatics studies together with degradome sequencing has also allowed to characterize a family of apparently cryptic miRNA-like precursors in the genome of *F. oxysporum*. These miRNAs are expressed but apparently unable to induce a target degradation, probably because of their lack of proper complementarity with their targets (Chen et al. 2014).

14.4 Small ncRNAs in Host Cell Response to Fungal Infections

During an infection, the host response against the pathogenic agent is started by inflammatory and immune responses with the objective of the pathogenic agent elimination. On the other hand, the pathogenic agent has been evolved to reshape the host cell function to fulfill its objectives, promoting cell survival and generating an appropriate environment for the colonization of the host organism. The intricate networks involving host-pathogen interaction are not limited to the initial contact between both cells. Moreover, signaling pathways are activated upon infection, being linked to specific remodeling of the host-cell transcriptional state in response to the presence of the pathogenic organism (Baxt et al. 2013). In some cases, the triggered transcriptional program is mainly composed by genes involved in the innate immunity, but on the majority of occasions the host cell responds also in a pathogen-specific fashion (Chaussabel et al. 2003). Fungal infected cells typically trigger an acute inflammatory response mediated by cytokines, including interleukines and chemokines, but the interactions between fungal virulence factors and signaling pathways within the host cell would modulate the progression of the infection in a specific manner.

The expression profiles of non-coding RNAs, and more specifically of miRNAs, vary among different cell types and the literature has presented evidence of the role of these sncRNAs in several human diseases, including fungal, bacterial and viral infections (Liu et al. 2014; Pope and Lasser 2013; Schnitger et al. 2011). Interestingly, intracellular infectious agents are usually able to induce a transcriptional program within the host cells that frequently include modifications of the expression pattern of specific miRNAs (Das Gupta et al. 2014; Schnitger et al. 2011; Martinez-Nunez et al. 2009). The majority of the data already available supporting this evidence has been obtained by the analysis of viruses and bacteria, being fungal infections less studied. However, even isolated evidences obtained from well characterized fungal pathogens like *Candida* and *Aspergillus*, indicate a clear relationship between the non-coding transcriptome and fungal infections.

In fungal infections, recent studies revealed that miR-155 plays an important role during maturation of dendritic cells, which inhibits the expression of the transcription factor PU.1 and thus decreases the levels of DC-SIGN and the ability to bind *C. albicans* cells (Monk et al. 2010). Furthermore, miR-155 can be of importance in what concerns to several infectious diseases and may contribute to the susceptibility to infection, and invasion by a range of pathogens (Martinez-Nunez et al. 2009). Interestingly, the study of the infection of respiratory epithelial cells by invasive *C. albicans* strains showed a different miRNA expression pattern in response to infection when compared with dendritic cells, including a strong up-regulation of miR-16-1 and concomitant down-regulation of miR-17-3p (Muhammad et al. 2015). Additionally, it was shown that miR-155, miR-146, miR-146b, miR-455 and miR-125 expression can be regulated by LPS and TLR4

(Monk et al. 2010). In silico studies for miR-455 and miR-125a suggested that they may target some mRNA transcripts encoding proteins involved in signalling pathways of inflammatory response, suggesting that these miRNAs may also play a role in limiting inflammation. However, further studies are needed to delineate the role of these miRNAs in the innate immune system and its importance in response to fungal infection (Monk et al. 2010). Recent evidences also demonstrated that miR-132 and miR-155 are differentially expressed in monocytes and dendritic cells upon stimulation with *A. fumigatus* or bacterial lipopolysaccharide (LPS), while miR-132 was induced only by *A. fumigatus*, suggesting that miR-132 could be a relevant regulator of the immune response directed specifically against this fungus (Das Gupta et al. 2014).

14.5 Further Perspectives

The phenotypic change observed in a host cell when it is infected by a fungal pathogen is consequence of specific transcriptional programs induced by the pathogen. Although different pathogens tend to take advantage of similar pathways in the host, the way and methods used to hijack the host cell is usually pathogen-specific. Fungal pathogens are unique in their virulence characteristics and abilities to colonize the host cells, and comparatively less studied than bacteria. The accumulated experimental evidences in the last decade clearly showed distinct transcriptional patterns in host cells in response to fungal infections. The transcriptional response to infection in the host is often accompanied by a specific genomic expression of virulence genes and other non-coding transcripts by the pathogenic fungus. However, the functions and roles of the differentially expressed coding and non-coding genes in the context of the infection remain elusive.

The small ncRNAs produced by infectious fungi are starting to be considered as modulators of the colonization process that allow the pathogen to hijack the host cells to complete the infective cycle. Some examples have been studied in the literature, but the overall knowledge of these tiny modulators is still in its infancy. Several topics need to be further studied, namely the specific biosynthetic pathways for the production of small ncRNAs in pathogenic fungi and their regulation during the infection. The characterization of new RNA-based modulators would open a new antifungal therapeutic landscape based on the RNA interference of these regulatory pathways. Moreover, we cannot forget the transcriptional response of the host-cell after infection and the possible interaction between host and pathogen at the transcriptional level. For the successful completion of these studies, there is also an evident need for infection models, useful for the functional studies but also for the screening of new antifungal agents.

References

Adenis A, Nacher M, Hanf M, Vantilcke V, Boukhari R, Blachet D, Demar M, Aznar C, Carme B, Couppie P (2014) HIV-associated histoplasmosis early mortality and incidence trends: from neglect to priority. PLoS Negl Trop Dis 8(8):e3100. doi:10.1371/journal.pntd.0003100

Bai Y, Lan F, Yang W, Zhang F, Yang K, Li Z, Gao P, Wang S (2015) sRNA profiling in *Aspergillus flavus* reveals differentially expressed miRNA-like RNAs response to water activity and temperature. Fungal Genet Biol 81:113–119. doi:10.1016/j.fgb.2015.03.004

Baxt LA, Garza-Mayers AC, Goldberg MB (2013) Bacterial subversion of host innate immune pathways. Science 340(6133):697–701. doi:10.1126/science.1235771

Bernstein DA, Vyas VK, Weinberg DE, Drinnenberg IA, Bartel DP, Fink GR (2012) Candida albicans Dicer (CaDcr1) is required for efficient ribosomal and spliceosomal RNA maturation. Proc Natl Acad Sci USA 109(2):523–528. doi:10.1073/pnas.1118859109

Bonifaz A, Vazquez-Gonzalez D, Perusquia-Ortiz AM (2011) Endemic systemic mycoses: coccidioidomycosis, histoplasmosis, paracoccidioidomycosis and blastomycosis. J Dtsch Dermatol Ges 9 (9):705–714; quiz 715. doi:10.1111/j.1610-0387.2011.07731.x

Bowen AD, Davidson FA, Keatch R, Gadd GM (2007) Induction of contour sensing in *Aspergillus niger* by stress and its relevance to fungal growth mechanics and hyphal tip structure. Fungal Genet Biol 44(6):484–491. doi:10.1016/j.fgb.2006.11.012

Brizendine KD, Baddley JW, Pappas PG (2013) Predictors of mortality and differences in clinical features among patients with cryptococcosis according to immune status. PLoS ONE 8(3): e60431. doi:10.1371/journal.pone.0060431

Brunke S, Hube B (2013) Two unlike cousins: *Candida albicans* and *C. glabrata* infection strategies. Cell Microbiol 15(5):701–708. doi:10.1111/cmi.12091

Calderone R, Sun N, Gay-Andrieu F, Groutas W, Weerawarna P, Prasad S, Alex D, Li D (2014) Antifungal drug discovery: the process and outcomes. Future Microbiol 9(6):791–805. doi:10.2217/fmb.14.32

Calo S, Shertz-Wall C, Lee SC, Bastidas RJ, Nicolas FE, Granek JA, Mieczkowski P, Torres-Martinez S, Ruiz-Vazquez RM, Cardenas ME, Heitman J (2014) Antifungal drug resistance evoked via RNAi-dependent epimutations. Nature 513(7519):555–558. doi:10.1038/nature13575

Casadevall A (1998) Antibody-mediated protection against intracellular pathogens. Trends Microbiol 6(3):102–107. doi:10.1016/S0966-842X(98)01208-6

Casadevall A (2010) *Cryptococci* at the brain gate: break and enter or use a Trojan horse? J Clin Invest 120(5):1389–1392. doi:10.1172/JCI42949

Casadevall A, Pirofski LA (1999) Host-pathogen interactions: redefining the basic concepts of virulence and pathogenicity. Infect Immun 67(8):3703–3713

Catalanotto C, Pallotta M, ReFalo P, Sachs MS, Vayssie L, Macino G, Cogoni C (2004) Redundancy of the two dicer genes in transgene-induced posttranscriptional gene silencing in *Neurospora crassa*. Mol Cell Biol 24(6):2536–2545

Cervantes M, Vila A, Nicolas FE, Moxon S, de Haro JP, Dalmay T, Torres-Martinez S, Ruiz-Vazquez RM (2013) A single Argonaute gene participates in exogenous and endogenous RNAi and controls cellular functions in the basal fungus *Mucor circinelloides*. PLoS ONE 8 (7):e69283. doi:10.1371/journal.pone.0069283

Chaussabel D, Semnani RT, McDowell MA, Sacks D, Sher A, Nutman TB (2003) Unique gene expression profiles of human macrophages and dendritic cells to phylogenetically distinct parasites. Blood 102(2):672–681. doi:10.1182/blood-2002-10-3232

Chen R, Jiang N, Jiang Q, Sun X, Wang Y, Zhang H, Hu Z (2014) Exploring microRNA-like small RNAs in the filamentous fungus *Fusarium oxysporum*. PLoS ONE 9(8):e104956. doi:10.1371/journal.pone.0104956

Chicas A, Forrest EC, Sepich S, Cogoni C, Macino G (2005) Small interfering RNAs that trigger posttranscriptional gene silencing are not required for the histone H3 Lys9 methylation

necessary for transgenic tandem repeat stabilization in *Neurospora crassa*. Mol Cell Biol 25 (9):3793–3801. doi:10.1128/MCB.25.9.3793-3801.2005

Choi J, Kim KT, Jeon J, Wu J, Song H, Asiegbu FO, Lee YH (2014) funRNA: a fungi-centered genomics platform for genes encoding key components of RNAi. BMC Genom 15(Suppl 9): S14. doi:10.1186/1471-2164-15-S9-S14

Cogoni C, Macino G (1999a) Gene silencing in *Neurospora crassa* requires a protein homologous to RNA-dependent RNA polymerase. Nature 399(6732):166–169. doi:10.1038/20215

Cogoni C, Macino G (1999b) Posttranscriptional gene silencing in *Neurospora* by a RecQ DNA helicase. Science 286(5448):2342–2344

Colombo AL (2000) Epidemiology and treatment of hematogenous candidiasis: a Brazilian perspective. Braz J Infect Dis 4(3):113–118

Costa MC, Leitao AL, Enguita FJ (2012) Biogenesis and mechanism of action of small non-coding RNAs: insights from the point of view of structural biology. Int J Mol Sci 13(8):10268–10295. doi:10.3390/ijms130810268

Das Gupta M, Fliesser M, Springer J, Breitschopf T, Schlossnagel H, Schmitt AL, Kurzai O, Hunniger K, Einsele H, Loffler J (2014) *Aspergillus fumigatus* induces microRNA-132 in human monocytes and dendritic cells. Int J Med Microbiol 304(5–6):592–596. doi:10.1016/j.ijmm.2014.04.005

Deepe GS Jr, Gibbons RS (2009) Interleukins 17 and 23 influence the host response to *Histoplasma capsulatum*. J Infect Dis 200(1):142–151. doi:10.1086/599333

Drinnenberg IA, Weinberg DE, Xie KT, Mower JP, Wolfe KH, Fink GR, Bartel DP (2009) RNAi in budding yeast. Science 326(5952):544–550. doi:10.1126/science.1176945

Dumesic PA, Natarajan P, Chen C, Drinnenberg IA, Schiller BJ, Thompson J, Moresco JJ, Yates JR 3rd, Bartel DP, Madhani HD (2013) Stalled spliceosomes are a signal for RNAi-mediated genome defense. Cell 152(5):957–968. doi:10.1016/j.cell.2013.01.046

Enguita FJ, Costa MC, Fusco Almeida AM, Mendes-Giannini MJ, Leitao AL (2016) Transcriptomic crosstalk between fungal invasive pathogens and their host cells: opportunities and challenges for next-generation sequencing methods. J Fungi 2(1):7. doi:10.3390/jof2010007

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391 (6669):806–811. doi:10.1038/35888

Gheinani AH, Jahromi NH, Feuk-Lagerstedt E, Taherzadeh MJ (2011) RNA silencing of lactate dehydrogenase gene in *Rhizopus oryzae*. J RNAi Gene Silencing 7:443–448

Ghildiyal M, Zamore PD (2009) Small silencing RNAs: an expanding universe. Nat Rev Genet 10 (2):94–108. doi:10.1038/nrg2504

Girmenia C, Arcese W, Micozzi A, Martino P, Bianco P, Morace G (1992) Onychomycosis as a possible origin of disseminated *Fusarium solani* infection in a patient with severe aplastic anemia. Clin Infect Dis 14(5):1167

Goes T, Bailao EF, Correa CR, Bozzi A, Santos LI, Gomes DA, Soares CM, Goes AM (2014) New developments of RNAi in *Paracoccidioides brasiliensis*: prospects for high-throughput, genome-wide, functional genomics. PLoS Negl Trop Dis 8(10):e3173. doi:10.1371/journal.pntd.0003173

Guerra-Assuncao JA, Enright AJ (2012) Large-scale analysis of microRNA evolution. BMC Genom 13:218. doi:10.1186/1471-2164-13-218

Janbon G, Maeng S, Yang DH, Ko YJ, Jung KW, Moyrand F, Floyd A, Heitman J, Bahn YS (2010) Characterizing the role of RNA silencing components in *Cryptococcus neoformans*. Fungal Genet Biol 47(12):1070–1080. doi:10.1016/j.fgb.2010.10.005

Jayawickrama MM, Gamage HU, Perera PD (2012) *Conidiobolus coranatus* infection of the para-nasal sinuses masquerading as a sino-nasal tumour. Ceylon Med J 57(2):87–88. doi:10.4038/cmj.v57i2.4464

Jiang N, Yang Y, Janbon G, Pan J, Zhu X (2012) Identification and functional demonstration of miRNAs in the fungus *Cryptococcus neoformans*. PLoS ONE 7(12):e52734. doi:10.1371/journal.pone.0052734

Kohler JR, Casadevall A, Perfect J (2015) The spectrum of fungi that infects humans. Cold Spring Harb Perspect Med 5(1):a019273. doi:10.1101/cshperspect.a019273

Kriengkauykiat J, Ito JI, Dadwal SS (2011) Epidemiology and treatment approaches in management of invasive fungal infections. Clin Epidemiol 3:175–191. doi:10.2147/CLEP.S12502

Kroetz DN, Deepe GS (2012) The role of cytokines and chemokines in *Histoplasma capsulatum* infection. Cytokine 58(1):112–117. doi:10.1016/j.cyto.2011.07.430

Lau SK, Chow WN, Wong AY, Yeung JM, Bao J, Zhang N, Lok S, Woo PC, Yuen KY (2013) Identification of microRNA-like RNAs in mycelial and yeast phases of the thermal dimorphic fungus *Penicillium marneffei*. PLoS Negl Trop Dis 7(8):e2398. doi:10.1371/journal.pntd.0002398

Lee CT, Risom T, Strauss WM (2007) Evolutionary conservation of microRNA regulatory circuits: an examination of microRNA gene complexity and conserved microRNA-target interactions through metazoan phylogeny. DNA Cell Biol 26(4):209–218. doi:10.1089/dna.2006.0545

Lee HC, Chang SS, Choudhary S, Aalto AP, Maiti M, Bamford DH, Liu Y (2009) qiRNA is a new type of small interfering RNA induced by DNA damage. Nature 459(7244):274–277. doi:10.1038/nature08041

Lee HC, Aalto AP, Yang Q, Chang SS, Huang G, Fisher D, Cha J, Poranen MM, Bamford DH, Liu Y (2010a) The DNA/RNA-dependent RNA polymerase QDE-1 generates aberrant RNA and dsRNA for RNAi in a process requiring replication protein A and a DNA helicase. PLoS Biol 8(10). doi:10.1371/journal.pbio.1000496

Lee HC, Li L, Gu W, Xue Z, Crosthwaite SK, Pertsemlidis A, Lewis ZA, Freitag M, Selker EU, Mello CC, Liu Y (2010b) Diverse pathways generate microRNA-like RNAs and dicer-independent small interfering RNAs in fungi. Mol Cell 38(6):803–814. doi:10.1016/j.molcel.2010.04.005

Li L, Chang SS, Liu Y (2010) RNA interference pathways in filamentous fungi. Cell Mol Life Sci 67(22):3849–3863. doi:10.1007/s00018-010-0471-y

Lima P de S, Casaletti L, Bailao AM, de Vasconcelos AT, Fernandes Gda R, Soares CM (2014) Transcriptional and proteomic responses to carbon starvation in *Paracoccidioides*. PLoS Negl Trop Dis 8(5):e2855. doi:10.1371/journal.pntd.0002855

Liu T, Ren X, Xiao T, Yang J, Xu X, Dong J, Sun L, Chen R, Jin Q (2013) Identification and characterisation of non-coding small RNAs in the pathogenic filamentous fungus *Trichophyton rubrum*. BMC Genom 14:931. doi:10.1186/1471-2164-14-931

Liu Z, Zhou G, Deng X, Yu Q, Hu Y, Sun H, Wang Z, Chen H, Jia C, Wang D (2014) Analysis of miRNA expression profiling in human macrophages responding to *Mycobacterium infection*: induction of the immune regulator miR-146a. J Infect 68(6):553–561. doi:10.1016/j.jinf.2013.12.017

Martinez-Nunez RT, Louafi F, Friedmann PS, Sanchez-Elsner T (2009) MicroRNA-155 modulates the pathogen binding ability of dendritic cells (DCs) by down-regulation of DC-specific intercellular adhesion molecule-3 grabbing non-integrin (DC-SIGN). J Biol Chem 284(24):16334–16342. doi:10.1074/jbc.M109.011601

McGuire AM, Galagan JE (2008) Conserved secondary structures in *Aspergillus*. PLoS ONE 3(7):e2812. doi:10.1371/journal.pone.0002812

Mikulska M, Bassetti M, Ratto S, Viscoli C (2011) Invasive candidiasis in non-hematological patients. Mediterr J Hematol Infect Dis 3(1):e2011007. doi:10.4084/MJHID.2011.007

Miyoshi K, Miyoshi T, Siomi H (2010) Many ways to generate microRNA-like small RNAs: non-canonical pathways for microRNA production. Mol Genet Genomics 284(2):95–103. doi:10.1007/s00438-010-0556-1

Monk CE, Hutvagner G, Arthur JS (2010) Regulation of miRNA transcription in macrophages in response to *Candida albicans*. PLoS ONE 5(10):e13669. doi:10.1371/journal.pone.0013669

Muhammad SA, Fatima N, Syed NI, Wu X, Yang XF, Chen JY (2015) MicroRNA expression profiling of human respiratory epithelium affected by invasive *Candida* infection. PLoS ONE 10(8):e0136454. doi:10.1371/journal.pone.0136454

Murphy D, Dancis B, Brown JR (2008) The evolution of core proteins involved in microRNA biogenesis. BMC Evol Biol 8:92. doi:10.1186/1471-2148-8-92

Napoli C, Lemieux C, Jorgensen R (1990) Introduction of a chimeric chalcone synthase gene into *Petunia* results in reversible co-suppression of homologous genes in trans. Plant Cell 2(4):279–289. doi:10.1105/tpc.2.4.279

Nicolas FE, Vila A, Moxon S, Cascales MD, Torres-Martinez S, Ruiz-Vazquez RM, Garre V (2015) The RNAi machinery controls distinct responses to environmental signals in the basal fungus *Mucor circinelloides*. BMC Genom 16:237. doi:10.1186/s12864-015-1443-2

Nucci M, Marr KA (2005) Emerging fungal diseases. Clin Infect Dis 41(4):521–526. doi:10.1086/432060

Pasqualotto AC, Zimerman RA, Alves SH, Aquino VR, Branco D, Wiltgen D, do Amaral A, Cechinel R, Colares SM, da Rocha IG, Severo LC, Sukiennik TC (2008) Take control over your fluconazole prescriptions: the growing importance of *Candida glabrata* as an agent of candidemia in Brazil. Infect Control Hosp Epidemiol 29(9):898–899. doi:10.1086/590191

Pope SM, Lasser C (2013) *Toxoplasma gondii* infection of fibroblasts causes the production of exosome-like vesicles containing a unique array of mRNA and miRNA transcripts compared to serum starvation. J Extracell Vesicles 2. doi:10.3402/jev.v2i0.22484

Prado M, Silva MB, Laurenti R, Travassos LR, Taborda CP (2009) Mortality due to systemic mycoses as a primary cause of death or in association with AIDS in Brazil: a review from 1996 to 2006. Mem Inst Oswaldo Cruz 104(3):513–521. doi:10.1590/S0074-02762009000300019

Rappleye CA, Engle JT, Goldman WE (2004) RNA interference in *Histoplasma capsulatum* demonstrates a role for alpha-(1,3)-glucan in virulence. Mol Microbiol 53(1):153–165. doi:10.1111/j.1365-2958.2004.04131.x

Richards TA (2011) Genome evolution: horizontal movements in the fungi. Curr Biol 21(4):R166–168. doi:10.1016/j.cub.2011.01.028

Romano N, Macino G (1992) Quelling: transient inactivation of gene expression in *Neurospora crassa* by transformation with homologous sequences. Mol Microbiol 6(22):3343–3353

Schnitger AK, Machova A, Mueller RU, Androulidaki A, Schermer B, Pasparakis M, Kronke M, Papadopoulou N (2011) *Listeria monocytogenes* infection in macrophages induces vacuolar-dependent host miRNA response. PLoS ONE 6(11):e27435. doi:10.1371/journal.pone.0027435

Segers GC, Zhang X, Deng F, Sun Q, Nuss DL (2007) Evidence that RNA silencing functions as an antiviral defense mechanism in fungi. Proc Natl Acad Sci USA 104(31):12902–12906. doi:10.1073/pnas.0702500104

Silva SS, Tavares AH, Passos-Silva DG, Fachin AL, Teixeira SM, Soares CM, Carvalho MJ, Bocca AL, Silva-Pereira I, Passos GA, Felipe MS (2008) Transcriptional response of murine macrophages upon infection with opsonized *Paracoccidioides brasiliensis* yeast cells. Microbes Infect 10(1):12–20. doi:10.1016/j.micinf.2007.09.018

Sun J, Li X, Feng P, Zhang J, Xie Z, Song E, Xi L (2014) RNAi-mediated silencing of fungal acuD gene attenuates the virulence of *Penicillium marneffei*. Med Mycol 52(2):167–178. doi:10.1093/mmy/myt006

Tortorano AM, Dho G, Prigitano A, Breda G, Grancini A, Emmi V, Cavanna C, Marino G, Morero S, Ossi C, Delvecchio G, Passera M, Cusumano V, David A, Bonaccorso G, Corona A, Favaro M, Vismara C, Garau MG, Falchi S, Tejada MR (2012) Invasive fungal infections in the intensive care unit: a multicentre, prospective, observational study in Italy (2006-2008). Mycoses 55(1):73–79. doi:10.1111/j.1439-0507.2011.02044.x

Trieu TA, Calo S, Nicolas FE, Vila A, Moxon S, Dalmay T, Torres-Martinez S, Garre V, Ruiz-Vazquez RM (2015) A non-canonical RNA silencing pathway promotes mRNA degradation in basal Fungi. PLoS Genet 11(4):e1005168. doi:10.1371/journal.pgen.1005168

Ubiali DG, Cruz RA, De Paula DA, Silva MC, Mendonca FS, Dutra V, Nakazato L, Colodel EM, Pescador CA (2013) Pathology of nasal infection caused by *Conidiobolus lamprauges* and *Pythium insidiosum* in sheep. J Comp Pathol 149(2–3):137–145. doi:10.1016/j.jcpa.2012.12.002

Velagapudi R, Hsueh YP, Geunes-Boyer S, Wright JR, Heitman J (2009) Spores as infectious propagules of *Cryptococcus neoformans*. Infect Immun 77(10):4345–4355. doi:10.1128/IAI.00542-09

Vu K, Tham R, Uhrig JP, Thompson GR 3rd, Na Pombejra S, Jamklang M, Bautos JM, Gelli A (2014) Invasion of the central nervous system by *Cryptococcus neoformans* requires a secreted fungal metalloprotease. MBio 5(3):e01101–01114. doi:10.1128/mBio.01101-14

Wang Y (2015) Looking into *Candida albicans* infection, host response, and antifungal strategies. Virulence 6(4):307–308. doi:10.1080/21505594.2014.1000752

Wang X, Hsueh YP, Li W, Floyd A, Skalsky R, Heitman J (2010) Sex-induced silencing defends the genome of *Cryptococcus neoformans* via RNAi. Genes Dev 24(22):2566–2582. doi:10.1101/gad.1970910

Wang X, Darwiche S, Heitman J (2013) Sex-induced silencing operates during opposite-sex and unisexual reproduction in *Cryptococcus neoformans*. Genetics 193(4):1163–1174. doi:10.1534/genetics.113.149443

Yin Z, Li Y, Han X, Shen F (2012) Genome-wide profiling of miRNAs and other small non-coding RNAs in the verticillium dahliae-inoculated cotton roots. PLoS ONE 7(4):e35765. doi:10.1371/journal.pone.0035765

Index

A.L. Leitão and F.J. Enguita (eds.), *Non-coding RNAs and Inter-kingdom Communication*, DOI 10.1007/978-3-319-39496-1