High-throughput screening of enhancer elements and miRNA-target interactions

Differences in gene regulation are known to underlie phenotypic disparities between cell types in multicellular organisms. In particular, transcriptional regulation has often been assumed to provide the best approximation to the regulatory state of a cell and has, indeed, yielded numerous insights into a multitude of biological systems. Gene regulatory networks have been described that focus mostly on interactions involving promoter sequences of a few core genes. Mammalian transcriptional regulation, however, relies very heavily on distal regulatory elements also known as enhancers. The lack of a TATA box, a transcriptional start site or other very unique features has rendered the identification and subsequent incorporation of enhancers into gene regulatory networks very slow, inefficient and rather serendipitous. We have developed a comprehensive approach that combines computational and experimental methods with a high-throughput robotic platform to allow for the identification and functional testing of putative enhancer elements governing cell specification. More specifically, we have chosen to study the specification of hematopoietic cell fates and the transcriptional program induced upon stimulation with Lipo polysaccharide (LPS), which is meant to represent the response of immune cells to a bacterial infection.

The regulatory regions of genes of interest are exhaustively searched using multiple-genome alignment tools, obtaining an average of 20 clustered evolutionary conserved regions (CECRs) of 500bp to 1kb in length for each gene analyzed. Each element is subcloned into a bacterial vector and electroporated into mammalian cells to be tested for function in a luciferase assay. We have developed a mutagenesis approach to try to determine the identity of the transcription factors responsible for the activity of given enhancers. We plan to complement these latter experiments with Chromatin Immunoprecipitation (ChIP) and Electrophoretic Mobility Shift Assay (EMSA) to test for in vivo occupancy as well as in vitro binding, respectively.

Recently, I have started working on an experimental scheme that will take advantage of the robotic-based high-throughput approach already developed to functionally test miRNA-target interactions. I will discuss the strategies that we are currently considering and the design of the proof-of-principle experiments that will soon be carried out. By adding this new dimension to our efforts towards the construction of gene regulatory networks, we hope to gain deeper insight into the molecular mechanisms at work during cell fate specification and immune stimulation.

I have included two readings, both of them reviews.

I suggest starting with “The evolution of gene regulation by transcription factors and miRNAs”. Read all figures, diagrams and boxes carefully. Also read main text until “Rates of evolution”. Familiarize yourselves with transcription factors and microRNAs and generally understand how their mechanisms of action are different.

The second review, “microRNAs and the immune response” is to give you a sense of how much is known about the role of miRNAs in hematopoiesis and the immune response. The introduction and conclusion will be most helpful; there is no need to read the specifics known about each miRNA. Please, take a look at Fig. 2.
The emergence of complex, multicellular organisms was accompanied, and perhaps facilitated, by dramatic increases in the complexity of gene regulatory mechanisms\(^1\)\(^2\). At the level of transcriptional regulation, this can be clearly seen in the massive expansions of transcription-factor families and the pervasive combinatorial control of genes by multiple transcription factors in higher organisms\(^1\)\(^3\) (BOX 1). At the level of post-transcriptional control, entirely new mechanisms of gene regulation arose, typified by a large and growing class of \(\sim 22\)-nucleotide-long non-coding RNAs, known as microRNAs (miRNAs), which function as repressors in all known animal and plant genomes\(^4\)\(^5\) (BOX 1). Although transcription factors and miRNAs are two of the best-studied gene regulatory mechanisms, there are many other layers of gene regulation, including: cell signalling; mRNA splicing, polyadenylation and localization; chromatin modifications; and mechanisms of protein localization, modification and degradation (FIG. 1).

But why have higher plants and animals evolved such complex, multilayered gene regulatory systems? Is combinatorial transcriptional regulation alone insufficient to specify a developmental programme? What are the relative contributions of the various mechanisms of gene regulation to changes at the phenotypic level? Do all the different modes of gene regulation evolve in the same manner and at the same rate? Despite half a century of research on gene regulation, such questions have yet to be tackled seriously, because most of the effort in the field so far has been devoted to studying the evolution of transcriptional regulation. Although the primacy of transcription as the necessary first step in gene expression is undeniable, this does not imply that transcriptional regulation has the largest effect on the final concentration of the active gene product, which is the most relevant quantity to the phenotype.

An important goal for future research is to elucidate how complex gene regulatory networks evolve and how their evolution results in phenotypic change and speciation. However, a necessary first step towards this goal is to understand the basic principles that underlie the evolution of the individual regulators and their regulatory interactions with their target genes. Recent computational and experimental work has made it possible to begin to study the evolution of transcription factors, miRNAs and their binding sites, and to compare the rate and manner in which these two important regulatory mechanisms evolve. Here we mainly focus on animal evolution because most of the work on gene regulatory evolution has been carried out in animal systems, although we refer to plant evolution whenever possible.

Ultimately, a complete picture of the evolution of gene regulation will require a synthesis of information about all the diverse components of gene regulatory networks. As an initial step, we propose a simple model in which the evolution of transcriptional control of animal miRNA genes themselves is an important step in the successful acquisition of a novel miRNA. A corollary of this model is that there are miRNAs that are transcribed at low levels and in specific cell types, which might have little biological function in regulating target genes in trans.
Transcription factors and miRNAs in development

No discussion of the evolution of gene regulation is complete without a consideration of the developmental roles of the regulators themselves, because these roles imply certain constraints on the evolvability of the regulatory relationships\(^6,7\). Furthermore, if we accept that development is the consequence of the unfolding of precise and robust spatio-temporal patterns of gene expression, then it is in the context of development that the evolution of gene regulation is most closely related to the evolution of organismal form\(^3,7\).

In complex multicellular organisms, transcription factors generally do not work in isolation, but instead, together with co-regulators, they form large networks of cooperating and interacting transcription factors\(^3,8\). It is widely believed that the rate of evolution of regulatory relationships is not homogeneous over the entire network. For example, in one model that was proposed by Davidson and colleagues, animal developmental networks can be decomposed into subnetworks, including highly conserved ‘kernels’ that specify the spatial domain in which a particular body part will develop; ‘plug-in’

**Box 1 | Transcription factors and microRNA genes — an overview**

Transcription factors are proteins that either activate or repress transcription of other genes by binding to short cis-regulatory elements called transcription-factor binding sites that lie in the vicinity of the target genes. Transcription-factor binding sites, especially in developmental regulatory networks, are often organized into clusters called cis-regulatory modules (CRMs)\(^3\), which typically span a few hundred nucleotides and can contain dozens of binding sites for ~3–10 transcription factors. CRMs produce the initial spatio-temporal expression pattern of the target gene by ‘reading out’ the concentrations of multiple transcription factors that are present in a particular cell at a particular time. So, dependence on cellular context and combinatorial control are common themes in transcriptional regulation. Transcription factors are usually grouped into families on the basis of shared DNA-binding domains, which are an important determinant of transcription-factor binding specificity.

Mature microRNAs (miRNAs) are short (~22-nucleotide), non-coding ssRNAs that repress mRNAs post-transcriptionally by binding to partially complementary sites, called miRNA binding sites, in their target mRNAs. In animals, miRNA-mediated repression is often relatively weak, whereas transcription-factor-mediated repression can be much stronger. Mature miRNAs are cleaved from ~70-nucleotide hairpin structures, called precursor miRNAs (pre-miRNAs), by the enzyme dicer. Pre-miRNAs are in turn excised from a primary miRNA (pri-miRNA) transcript by the enzyme drosha. Pri-miRNAs are typically transcribed by RNA polymerase II\(^104\) and seem to possess promoter and enhancer elements that are similar to those of protein-coding genes (for example, \(^105,106\)). They can be thousands of nucleotides long and contain multiple pre-miRNAs. In some cases, however, pre-miRNAs are contained in introns of protein-coding genes and are excised by the splicing machinery. In metazoans, pre-miRNAs are exported into the cytoplasm where they are processed into mature miRNAs, whereas in plants, miRNA maturation occurs within the nucleus. miRNAs are grouped into families on the basis of their target recognition motifs (Box 3).

The predominant regulatory effect of miRNAs is to repress their target mRNAs; mechanisms for this include translational repression, mRNA cleavage, mRNA deadenylation or alteration of mRNA stability (reviewed in \(^107\)). miRNA-mediated cleavage of mRNAs seems to be the exception in animals. By contrast, it is believed to be the dominant regulatory mode in plants. DCL1, dicer-like 1; miRNA*, the rapidly decaying strand that is complementary to the mature miRNA.
components, such as signal transduction cassettes, which are re-used in multiple developmental contexts; terminal differentiation gene batteries that consist of all the genes that define a particular cell type; and ‘input–output switches’, such as Hox genes, which allow or disallow the action of particular developmental processes in specific spatio-temporal contexts. The authors proposed that each of these types of subnetwork has its own set of evolutionary constraints, and that changes in different subnetworks result in qualitatively different types of change at the phenotypic level. For example, kernels are by definition highly conserved, and their consistency over long periods of evolutionary time might provide an explanation for the high degree of conservation of body plans within animal phyla. By contrast, differentiation gene sets, being the least pleiotropic of all the regulatory relationships, are easiest to alter, and such changes might result in species-level phenotypic differences (for details, see REFs 3,7).

Various studies have demonstrated that miRNAs have important roles in animal and plant development (see REFs 10–14 for a consideration of the developmental roles of miRNAs). Some well known examples include miRNAs with switch-like roles, such as lin-4 and let-7 in Caenorhabditis elegans developmental timing or the miRNAs that are involved in plant leaf or flower development (reviewed in REF. 15), and miRNAs that confer more general tissue or temporal identity, such as miR-1 in Drosophila melanogaster muscle development and miR-430 in the zebrafish maternal–zygotic transition.

Much of the current evidence for an early developmental role for miRNAs is conflicting or difficult to interpret. First, zebrafish embryos that lack maternal dicer, a protein that is required for miRNA biogenesis (BOX 2), progress through axis formation and regionalization, a fact that strongly argues against a role for miRNAs in early zebrafish development. But dicer knockout mice and Arabidopsis thaliana that carry hypomorphic alleles of the dicer homolog DCL-1 die in early embryogenesis. Second, although miRNAs have not been detected in early zebrafish and medaka embryos, mature miRNAs have been detected in mice and D. melanogaster embryos. Primary miRNA transcripts are spatially regulated in early D. melanogaster embryogenesis, with the caveat that the processing of primary miRNAs into mature miRNAs (BOX 2) can be regulated. Third, on the basis of miRNA knockdowns using 2′-O-methyl antisense oligoribonucleotides, it has been reported that miRNAs are involved in patterning the D. melanogaster embryo, although a number of these results disagree with experimental data from genetic knockouts.

The prevailing opinion (for example, REFs 3,29 and the references therein) seems to be that miRNAs as a class tend to function as lock-down mechanisms for already-differentiated states, or confer an additional layer of robustness or ‘noise’ reduction on the developmental processes, rather than having fundamental roles in body-plan patterning. However, because the functions of only a few miRNAs have been dissected in detail, we believe that this point has not yet been proven unambiguously, at least not as a general principle.
Box 2 | MicroRNA gene discovery: bioinformatics and experimental methods

The two general approaches to microRNA (miRNA) gene discovery are bioinformatic and experimental methods (reviewed in REFs 108,109). Generally, bioinformatic methods use RNA-folding algorithms (for example, REFs 110,111) to search for approximately hairpin structures in non-coding and non-repetitive regions of the genome, and filter them using patterns of evolutionary conservation. Known examples of miRNA precursors are used as training examples for machine learning algorithms to discriminate between true predictions and false positives (reviewed in REFs 108,109). Predictions are generally verified by northern blots, PCR or microarray analysis. Naturally, bioinformatic predictions have false-positive rates and can miss species-specific miRNAs, as many, but not all112, of the current methods use evolutionary conservation as an indicator of biological function.

The traditional experimental approach to miRNA discovery is cloning and sequencing (for example, REF. 113). This approach successfully determines species-specific miRNAs, but tends to miss miRNAs that are expressed at low levels, in a small number of cells or only under particular cellular conditions. More recently, high-throughput sequencing methods, especially 454 sequencing114, have become popular for surveying small RNA populations (for example, REF. 115).

Currently, 328 miRNAs have been annotated in the human genome and 199 in Arabidopsis thaliana, the metazoan and plant species in which their small RNA complements have been well surveyed115. Several groups have shown that metazoan miRNAs probably regulate thousands of genes in mammals116,117,118, flies119,120, and nematodes121. For example, in humans, known miRNAs make up >1% of the gene repertoire and are thought to regulate >30% of all protein-coding genes.

Both computational and sequencing approaches indicate that there are likely to be many more miRNAs, many of which are lineage-specific (for example, REFs 86,92).

Evidence to the contrary) and miRNA biogenesis and the mechanism of miRNA-mediated gene repression are significantly different between plants and animals (BOX 1). Further evidence for this comes from the phylogenetic distribution of miRNAs, specifically the apparent lack of miRNAs in sponges and fungi, as discussed below. To a first approximation then, miRNA genes and most transcription-factor families have evolved independently in the animal and plant kingdoms.

Deep conservation of transcription factors and miRNAs

It is well known that many transcription factors are highly conserved across large evolutionary distances, and some have similar developmental roles in diverse species. Hox genes, which regulate development along the anterior–posterior axis in most animals, are the textbook example of this phenomenon. Other examples in vertebrates and many invertebrates include the paired box 6 (Pax6) genes, which direct eye and anterior nervous system development, and Csx/Nkx2-5/Tinman genes, which direct visceral mesoderm and heart development (reviewed in REF. 8). Although such broad similarities in function are intriguing, it is worth noting that they do not necessarily imply complete functional redundancy between distant homologues, or even evolutionarily conserved developmental roles122.

Likewise, many miRNAs seem to be extremely well conserved. The best known example is let-7, which is phylogenetically conserved in all bilaterians that have been tested so far, with the single exception of acoel flatworms123,124. Furthermore, the exact sequence of the mature form of let-7, its temporal expression pattern and in some cases its syntonic position in the genome are conserved125,126. Some other well known examples are muscle-specific mir-1, which is conserved in nematodes, mammals and flies (REF. 17 and the references therein) and miR-7, the mature form of which is perfectly conserved between mammals and flies, and which lies in an intron of the same host gene in both clades.

Several groups have used bioinformatic sequence comparisons or northern blots to study the conservation of miRNAs across many animal species127–31. They found that 18–30 miRNA families seem to be conserved in all bilaterians that have been studied, depending on the stringency of the bioinformatic methods used. Three of these miRNA families were also found in cnidarians, but none was found in sponges128,129. As previously discussed, this supports the idea of independent evolution of miRNAs in plants and animals. One of the cnidarian miRNAs, miR-10, has been detected by northern blots, and is particularly interesting because it is present in a Hox cluster and regulates Hox genes in Drosophila and vertebrates30,31. Overall, the level of sequence conservation of many miRNAs is generally high, although sequence conservation need not imply functional conservation130.

Further evidence for deep conservation of animal miRNA genes at a genome-wide level can be found in cross-clade comparisons of highly conserved motifs (HCMs) in 3’UTRs. Xie et al. showed that HCMs in vertebrates are highly enriched in miRNA-binding sites32. We and others have extended these results to nematodes and flies124,125, and have shown that HCMs are highly conserved between vertebrates, nematodes and flies (Supplementary information S1 (figure)). Because many HCMs are expected to represent miRNA recognition motifs, these analyses provide indirect evidence for deep conservation of miRNA genes.

The sampling of miRNAs in plants is less extensive than in animals, with most of the plant miRNAs that are known so far having been discovered in A. thaliana (for example, REF. 44). Nevertheless, several examples of deeply conserved plant miRNAs are known from sequencing of miRNAs in moss45–47. Some of these examples are particularly impressive because their regulatory relationships with their targets also seem to be conserved, a point we return to below.

Lineage-specific expansions of transcription factors and miRNAs

Lineage-specific expansions of transcription-factor families are common and are widely believed to have an important role in both plant and animal diversification and complexity48,49. Interestingly, there is evidence that the expansion of transcription-factor families is greater in plants than in animals48,50. In principle, this could reflect a fundamental difference between plant and animal biology. However, an alternative explanation is that A. thaliana has undergone a recent whole-genome duplication, whereas the putative whole-genome duplications in animal lineages are more ancient: immediately after a large-scale duplication event, a transcription-factor loss might be deleterious if it changes the relative concentrations of the set of transcription factors that are expressed in a cell131, leading to increased rates of transcription-factor retention in plants compared with animals.
Several detailed studies in plants\(^{52}\) and animals\(^{53,54}\) have shown that miRNA families can expand by the same processes of tandem, segmental and whole-genome duplication as protein-coding genes. Plant miRNA families tend to be larger and their members more similar to each other than animal miRNA families (reviewed in REF. 55). This indicates that the expansion of plant families is more recent, and that the main effect of having multiple paralogous copies of the same miRNA in plants is to increase dosage. By contrast, family members of animal miRNAs might tend to have synergistic but functionally distinct roles, as shown for the let-7 family in *C. elegans*\(^{6,67}\).

**cis-element evolution**

*Large-scale rewiring of miRNA-mediated regulatory relationships over large evolutionary distances in animals, but deep conservation in plants.* Although miRNAs themselves seem to be highly conserved, there are only a few miRNA-target regulatory relationships that are known to be conserved over large evolutionary distances in animals (for example, between vertebrates and *Drosophila* species). There are two cases, namely let-7:lin-41 and let-7:let-60–RAS, in which the target relationship has been experimentally dissected in *C. elegans*, and for which there is evidence for conservation in mammals that was obtained from computational target predictions and reciprocal expression patterns of the miRNA and mRNA in cell lines\(^{63,69}\). Another miRNA-target relationship, lin-4:lin-28, has been established experimentally in *C. elegans*, and there is computational and gene-expression evidence that it is conserved in a number of other animal species, including mammals\(^{66,68}\). However, even in these cases, caution in interpretation is warranted, as the existence of the same regulatory relationship in highly diverged species does not always imply that it is evolutionarily conserved as opposed to independently evolved. For example, the well known Pax6 transcription factor is involved in mammalian and *Drosophila* eye development, and ectopic expression of the mouse homologue in *Drosophila* can induce the development of ectopic *Drosophila* eyes, but it is not clear whether even this regulatory relationship represents conservation or convergent evolution\(^7\).

These isolated examples aside, the conservation of computationally predicted miRNA targets between vertebrates, *Drosophila* and nematodes seems to be close to what is expected by chance\(^{56,57,62}\). These results seem to be robust, even when accounting for the false-positive and false-negative rates of the target prediction algorithms and possible errors in assignments of homologous genes and miRNAs\(^6\). Together with the high level of conservation that is observed for the miRNA genes themselves, this indicates that miRNA regulatory networks have undergone extensive rewiring during animal evolution.

In plants, a number of miRNA regulatory relationships are known to be conserved between *A. thaliana* and moss\(^{66,67}\), which diverged over 400 million years ago. These relationships are supported by strong experimental evidence, with verification by S'–RACE (rapid amplification of cloned ends) of the cleavage products of the target mRNA (BOX 2). Because many of these regulatory relationships are involved in crucial biological processes (for example, auxin signalling) this indicates that at least some miRNAs have held central positions within plant developmental regulatory networks for a long time.

Because transcription-factor binding sites are difficult to predict computationally, it is harder to decipher the global dynamics of transcription-factor binding site turnover over large evolutionary distances using computational methods. However, the recent emergence of large-scale experimentally defined transcription-factor binding-site data (reviewed in REF. 63), particularly from ChIP-chip analysis, might make this problem tractable in the near future. Towards this end, an interesting study that compared the targets of the RNA-binding protein Pumilio in *D. melanogaster* and yeast\(^{64}\) found that, although the binding affinity of the regulator had remained virtually the same in these two species, its targets had diverged almost completely. Therefore, the high conservation of *trans*-acting factors and low overall conservation of *cis*-regulatory sites might be common to many regulatory mechanisms.

**High turnover of binding sites even over short evolutionary distances.** A number of groups have studied the turnover of experimentally verified transcription-factor binding sites between humans and mice\(^6\), and between various *Drosophila* species\(^{66–69}\). The general conclusion is that sequence conservation for known binding sites is surprisingly low (for example, \(~50\%\) for *D. melanogaster* and *Drosophila pseudoobscura*\(^6\)), although this picture is complicated by the fact that selection is more likely to work at the level of an entire *cis*-regulatory module than on an individual binding site (BOX 1). In particular, *cis*-regulatory modules often contain redundant binding sites and, as elegantly demonstrated by Ludwig et al.\(^{60,61}\), there can be compensatory mutations that maintain the function of the enhancer despite the loss of individual binding sites. Similar studies have not been carried out for miRNAs due to the paucity of experimentally verified miRNA binding sites. However, several recent microarray-based studies have indicated that the rate of binding-site conservation is also around 50% (between humans and mice or between zebrafish, *Tetraodon* and *Fugu*)\(^{14,72,73}\). Furthermore, computational miRNA target predictions indicate that many lineage-specific miRNA binding sites exist in *Drosophila* and vertebrates (N.R., unpublished observations). Finally, we have used human SNP data and population-genetics techniques to show that 30–50% of non-conserved miRNA binding sites in the human genome might be functional when the mRNA and miRNA are expressed in the same tissue\(^7\). It should be noted that terms such as ‘non-conserved binding sites’ generally refer to cases in which a binding site cannot be aligned to its homologous sequence. Because these alignments often suffer from technical problems (for example, almost all alignment algorithms are unable to deal with genomic rearrangements on various scales), ‘non-conserved’ sites might actually turn out to be conserved in a strict evolutionary sense. Nevertheless, these studies together indicate that both miRNA and transcription-factor binding sites are gained and lost quickly over short evolutionary distances.

---

**ChIP-chip analysis**

A method that combines chromatin immunoprecipitation with microarray technology to identify in vivo targets of a transcription factor.
Taking these observations to the species level, it has been shown that human promoter regions harbour a surprisingly high amount of variation that significantly affects expression levels, presumably by interfering with transcriptional control (reviewed in Ref. 75). Examples of functional polymorphism have also been identified recently in miRNA binding sites in humans and sheep, and sets of SNPs in predicted miRNA binding sites in the human and mouse genomes have been collected by several groups. By contrast, resequencing of miRNA genes in humans showed almost no polymorphism in the sequences of mature miRNA genes, consistent with the higher levels of constraint on trans-acting regulators compared with cis-regulatory sites.

**Rates of evolution**

It has been suggested that repressors should evolve faster than activators, as there are many ways to repress a gene but relatively few ways to activate it. As transcription factors can function as activators or repressors, but all known miRNAs work as repressors, one might expect miRNA binding sites to evolve faster than transcription-factor binding sites.

However, a more fundamental difference between transcription factors and miRNAs is that transcription-factor binding sites are typically ‘fuzzy’ (that is, the same transcription factor can bind to many similar DNA sequences, possibly with different binding affinities), whereas many miRNA binding sites exhibit almost exact Watson–Crick complementarity, either to the first 6–8 bases from the 5′ end of the mature miRNA in animals, or to the entire mature miRNA in plants (Box 3). Therefore, under neutral evolution, one would expect that it is more difficult to destroy a functional transcription-factor binding site than to create a new one, whereas the converse would be expected for miRNA binding sites.

Plant and animal miRNA binding sites might also evolve at different rates. Plant miRNA binding sites are typically found in coding regions, and if we assume that non-synonymous sites are highly constrained and synchronous sites are neutrally evolving, then approximately one-third of the ~22 bases in a plant miRNA binding site can accommodate a substitution without highly deleterious consequences for the organism. Therefore, the sizes of the mutational target for plant and animal miRNA binding sites are comparable, implying that the probabilities of losing a plant or animal miRNA binding site are similar under simple neutral evolutionary models. On the other hand, the length of a plant miRNA binding site means that in theory it would be virtually impossible for a plant gene to gain a new miRNA binding site by point mutation, whereas the same is not true for animals.

---

**Box 3 | MicroRNA target prediction: bioinformatics and experimental methods**

Most microRNA (miRNA) targets have been identified using bioinformatics methods (reviewed in Refs 88,109,121). In fact, miRNAs are one of the few classes of trans-acting regulatory factors for which computational approaches can, with reasonable confidence, successfully predict a large number of cis-regulatory binding sites. This is primarily owing to the fact that miRNAs recognize their targets at least partly on the basis of simple sequence complementarity between the miRNA and its binding site. In other words, pure knowledge of the sequence of a miRNA is sufficient to predict many targets. This is typically not yet possible for transcription factors, for which large training data sets or other experimental information are needed to accurately identify targets computationally.

In plants, miRNA binding sites are usually contained in coding regions and have extensive complementarity to the mature miRNA. Therefore, many plant miRNA binding sites have been successfully discovered using relatively straightforward bioinformatics screens. Because the usual mechanism of repression in plants is mRNA cleavage, predicted binding sites can be verified with confidence using 5′-RACE (rapid amplification of cloned ends) to identify the cleavage products. In addition, smaller off-target effects are also possible.

Animal miRNA binding sites usually lie in 3′ UTRs of target mRNAs and exhibit imperfect complementarity to the mature miRNA (reviewed in Refs 88,109,121). Many target prediction methods are based on a model in which miRNA–mRNA binding is nucleated by an exact Watson–Crick complementary match to the first 6–8 bases from the 5′ end of the mature miRNA. Estimates of false-positive rates on the basis of comparative genomics (see Ref. 88 and references therein), population genetics and experimental assays all indicate that the accuracy of these algorithms is high. For example, the estimated accuracy for targets that are conserved in humans, chimpanzees, mice, rats and dogs is 50–85% (Ref. 74,88). However, other classes of miRNA sites have also been predicted, such as imperfect miRNA sites without an exact Watson–Crick match to the first 6–8 bases of the miRNA (reviewed in Refs 88,109,121).

Various assays have been developed to verify predicted animal miRNA targets, for example, by expressing the miRNA in a localized domain in vivo while simultaneously expressing and monitoring the target mRNA in a broader domain (for example, Ref. 119). However, the number of target sites that have been validated in vivo under endogenous conditions and by mutagenesis of the predicted sites is extremely small. An entirely different approach from computational methods is to assay the expression levels of mRNAs directly, using the observation that miRNAs can not only downregulate protein levels, but also downregulate mRNA levels of their targets (for example, Ref. 125). Therefore, a direct approach to target discovery is to knockout or overexpress the miRNA and use microarrays to identify the genes that show expression changes (for example, Refs 18,72,73), and to correlate these changes with 3′ UTR sequence motifs, for example, by a linear regression model. Because of the additional information, evolutionary conservation filters on the target-site predictions can be relaxed. Therefore, this method is expected to have a higher sensitivity than approaches that are purely sequence-based, although it suffers from the possibility of indirect effects.

Several databases and web resources have been developed to store and maintain predicted target sites and sites with experimental support (reviewed in Ref. 88,121).
Quantitative models of transcription-factor and miRNA binding-site evolution. One promising way of making the qualitative arguments of the previous section more quantitative involves developing models of binding-site evolution on the basis of point substitutions. Note that this approach does not accommodate the acquisition of new binding sites by large-scale rearrangements such as transpositions (for example, REF 80), or small-scale rearrangements such as tandem duplications81.

Stone and Wray82 calculated the expected time for the appearance of a new transcription-factor binding site by neutral point substitution and concluded that this time was short (for example, 55,000 years for two 6-bp sites to evolve within a 200-bp region in D. melanogaster82), assuming that all members of the population evolve independently. The problem was revisited under the more realistic assumption that the members of the population are related by descent83–85. Although the methods and models differed, the three studies arrived at similar conclusions. First, neutral mutation is too slow to efficiently evolve new binding sites by point substitutions. Second, positive selection on partial binding sites can effectively speed up the rate of evolution, making it feasible to evolve new binding sites. Third, the time taken to create new sites scales linearly with the length of the regulatory region (for example, the promoter or enhancer regions for a transcription-factor binding site, or the 3′ UTR for a miRNA binding site) but exponentially with the length of the binding site. Fourth, the base composition of the region is important, particularly the presence of ‘pre-sites’ that are a single point mutation away from being a functional binding site.

As a specific example, consider a 1-kb region of non-coding DNA with equal base composition in the human genome. Assume that the binding sites for a miRNA and a transcription factor are each 8 bp long, but a miRNA requires exactly eight matches whereas a transcription factor requires any seven matches out of eight. Durrett and Schmidt calculated that, given neutral point substitutions, it would take ~650 million years for the miRNA binding site to appear in the absence of a pre-site, and ~375,000 years in the presence of a pre-site. By contrast, the transcription-factor binding site would take ~60,000 years to appear85.

The rate of acquisition of miRNA genes versus transcription factors. As more genome sequences are completed, it emerges that few novel transcription-factor families have arisen since the divergence between animals and plants83–85 (although the number of transcription factors in each family can be different in individual genomes). The situation seems to be different for miRNAs. As a result of a combination of bioinformatics and sequencing efforts, it is now apparent that the process of miRNA creation is both active and ongoing (for example, REF. 34, 75, 86, 87) (BOX 2). For example, it seems that the human genome alone might contain more than 1,000 miRNAs86, of which many have been proposed to be primate-specific or even human-specific87. It should be noted that this global picture of transcription-factor-gene versus miRNA-gene acquisition ignores important issues such as combinatorial transcriptional control, co-factors and mutations outside of the DNA-binding domain that can affect transcription-factor binding specificity, as well as difficulties in computational identification of homologous miRNAs in different genomes. Nevertheless, even taking these points into account, the difference in the rates of creation of transcription-factor families and miRNA families remains striking, and it seems reasonable to propose that the speed of creation of new miRNA families has been faster in animal evolution than that of new transcription-factor families.

Do animals need so many miRNAs, and if so, why? Also, given that an animal miRNA can apparently target easily hundreds of genes (BOX 3) (reviewed in REF. 88), how can new miRNAs be acquired with such apparent ease without seriously disrupting the existing regulatory network of the organism? Even in plants in which the specificity of miRNA–mRNA binding is higher than in animals, significant off-target effects can occur and so similar issues can arise89. To address these questions, we next consider models of how new transcription factors and miRNAs evolve.

Creating new trans-factors

Transcription factors typically contain multiple functional domains, which mediate binding to DNA, interactions with other proteins and the subcellular localization of the transcription factor. A transcription factor with a new binding specificity can be created by duplication of an existing transcription factor followed by mutations, often, although not always, in the DNA-binding domain. Creation of new transcription factors can also evolve by the acquisition or loss of one of its other functional domains. For example, the loss of a transcriptional activation domain could turn an activator into a repressor, whereas the acquisition of a new protein–protein interaction domain that facilitates heterodimerization with a novel binding partner could significantly alter the targets of the transcription factor. It is often assumed that mutations in the coding sequence of the transcription factor itself are likely to be highly deleterious because they potentially affect the expression of many downstream target genes. However, recent examples of transcription factors that are important in Drosophila development that have significantly diverged in sequence and function indicate that this assumption might be worth revisiting (reviewed in REF. 34).

A duplication-mutation model accounts also for the evolution of at least some miRNA genes. For example, human miR-10a and miR-100 are homologues but differ by a single nucleotide insertion-deletion in the predicted target recognition region of the respective mature miRNAs. If this were the predominant mode of miRNA-family creation, one might expect that miRNA target recognition motifs would not be randomly distributed in sequence space, but would show a propensity to cluster in groups of similar sequences. Although this problem has not been fully studied, preliminary bioinformatic analyses indicate that the distribution of these sequences seems to be close to random (K.C. and N.R., unpublished observations). It is true, this would indicate that a duplication-mutation process accounts for a
relatively small fraction of miRNAs, and that instead most new miRNA families arise de novo.

One model of de novo miRNA acquisition, the inverted duplication model, suggests that new miRNAs evolve by inverted duplication of a stretch of coding sequence followed by subsequent erosion of the sequence into an imperfect hairpin structure. In some cases, part of the coding sequence itself might be duplicated before the creation of the miRNA, so the miRNA shows homology to two regions of the target gene. This model is attractive for plant miRNAs because it accounts for the long stretches of sequence similarity that are required between the miRNA and its target.

The inverted duplication model seems less appropriate for animal miRNAs, because the length of complementary sequence in miRNA binding sites is much smaller in animals than in plants, and no examples of newly formed miRNAs arising in this manner are known in animals. A second model, the random creation model, proposes that new miRNAs simply arise randomly from existing hairpin structures in the genome. Hairpin structures are generally abundant in eukaryotic genomes — for instance, Bentwich et al. identified 11 million hairpins in the human genome in a bioinformatic screen. Therefore, the problem of creating a new miRNA in animals might be less involved with creating a new hairpin structure, but rather with appropriately transcribing an existing hairpin structure in the genome and providing the requisite signals for biogenesis of the new miRNA (for example, signals for processing by the RNase III Drosha). (FIG. 2)

A model of transcriptional control of new miRNAs. Because the minimal binding site of an animal miRNA is short, a new miRNA should be able to target many mRNAs simply by chance, and many of these interactions are likely to be selectively deleterious, as is the case for all types of mutations. Indeed, the existence of many genes for which the presence of a miRNA binding site would be deleterious (‘anti-targets’) was proposed and later demonstrated by several groups in various species using microarray data and computational studies.

These observations raise the question of how a new miRNA could ever be acquired without seriously impairing the fitness of the organism. We propose that one way this could happen is if the miRNA were initially transcribed only weakly, and in a specific tissue or at a specific developmental stage. Current studies indicate that multiple sites for the same miRNA in the same target mRNA are needed to generate a strong regulatory effect of the miRNA on this target (REF 88 and references therein). If this is true, natural selection can eliminate slightly deleterious miRNA sites over time — recall that it is easier for an mRNA to lose a miRNA site than to gain one — while also maintaining or creating beneficial binding sites for the new miRNA. Once this elimination process is complete, the expression level of the miRNA can be increased and its tissue-specificity can be relaxed (FIG. 2).

Although this model is clearly a simplified picture of a complex process, it makes the testable prediction that more-recently evolved miRNAs should be expressed weakly and in specific spatio-temporal domains. Indeed, this prediction is generally supported by the available miRNA expression data (for example, REFs 87,96, which indicate that more-recently acquired human miRNAs are more likely to be weakly expressed than ancient conserved miRNAs. A recent study by Berezikov et al., in which miRNA expression in the human and chimpanzee brain was determined using deep sequencing (BOX 2), is also consistent with this prediction.

The model puts the enormous number of putative small RNA transcripts that are being revealed by deep sequencing efforts in perspective: it suggests that a number of them are randomly transcribed hairpins that might not have a significant biological role as trans-acting regulators, although it does not preclude the possibility that they could acquire such functionality in the future. The model is consistent with the hypothesis proposed by Sempere et al. and Prochnik et al. that the acquisition of new miRNAs contributed to the acquisition of novel tissue types and organs in animal development. It is also consistent with an intriguing idea put forward by Davidson that it is relatively easy to evolve a new miRNA gene that targets a specific sequence motif, but the same is not true of transcription factors, because the sequence specificity of a protein is a complex function of its amino-acid sequence. So, if a transcription factor were to acquire a new domain of expression (for example, in a new tissue), it would be expected to regulate genes that it already regulates in its original domain of expression, probably leading to deleterious effects. However, assuming that a miRNA with an arbitrary target sequence can evolve easily, such complications can be avoided, and consequently it might be easier to insert a miRNA into a developmental network than a transcription factor.

The expression and conservation of miRNAs are not as well understood in plants compared with animals. The correlation between the age of a miRNA and its expression level is expected to be weaker in plants than in animals for two reasons. First, a newly arising plant miRNA is expected to have relatively few targets compared with
an animal miRNA, and if these are either highly beneficial or highly deleterious, then selection can drive up or down the expression level of the miRNA more quickly. Second, a single plant miRNA molecule can target and cleave many mRNAs, whereas it seems that animal miRNAs must be bound to their targets to confer repression, a process that is reversible under specific conditions\(^\text{97,98}\). Nonetheless, it is likely that some plant miRNAs that are weakly expressed or tissue specific have little or no biological function as trans-acting regulators\(^\text{124}\).

**Conclusion**

With our rapidly advancing knowledge of the different mechanisms of gene regulation in higher eukaryotes, we can begin to consider the evolutionary implications of these different mechanisms within a unified framework. In the past, much work focused on a synthesis between transcriptional regulation and cell signalling mechanisms\(^\text{119,120}\); here we have concentrated on the evolution of transcription factors and miRNAs. Ultimately, all other mechanisms of gene regulation should be brought into the discussion in order to form a holistic picture of the evolution of gene regulation.

Many open questions and directions for future research remain. This Review gives a local view of the evolution of individual regulators and binding sites as a necessary first step to understanding the evolution of gene regulation as a whole. In the future, it will be necessary to move towards the broader view of the evolution of developmental regulatory networks, and from there, towards the even bigger picture of changes in organismal form. The model of Davidson and colleagues that is discussed above is one promising way of thinking about the evolution of network structure and body plans on a global scale. Our current knowledge of how transcription factors, miRNAs, signalling pathways and other regulators are wired together into developmental networks is much too rudimentary to make any sensible statements about the effect of different regulatory mechanisms on global network evolution. However, as our knowledge of the developmental roles of these regulatory mechanisms increases, it should be possible to extend the model to account for these different components. For example, if it indeed turns out that miRNAs tend to work at the periphery of developmental networks to confer additional layers of robustness, and not as the primary agents of developmental patterning, then this might lead us to postulate a certain amount of evolutionary pliability for miRNA-mediated regulation. It will also be interesting to investigate whether different eukaryotic lineages, particularly plants and animals, use different regulatory mechanisms in similar ways or not.

Such global trends in network evolution have a natural counterpart in local subnetwork motifs that exist within the overall network. One particularly interesting example of such a motif is a feedback loop that involves multiple transcription factors and miRNAs. Examples of this motif have been identified in recent work on neuron cell-fate determination\(^\text{100}\) and vulval development\(^\text{101}\) in *C. elegans*, and granulocytic differentiation in humans\(^\text{102}\). Whether multicomponent feedback loops and other complex subnetwork motifs are common features of developmental networks, and whether they have any significance for organismal traits and evolution, are intriguing questions for future research. Within the Davidson model, signalling cassettes function as plug-in components that are re-used repeatedly in regulatory networks. Likewise, certain subnetwork motifs could also potentially form re-usable plug-in components.

Since the work of Mary-Claire King and Allan Wilson three decades ago\(^\text{103}\), scientists have asked whether changes in gene regulation or protein sequence have made bigger contributions to phenotypic differences between species. Today, we are well positioned to broaden the question to ask about the relative contributions of the evolution of different mechanisms of gene regulation to the evolution of phenotypic diversity in animals and plants. The long journey towards a comprehensive understanding of the evolution of gene regulation is only beginning.
These authors studied the evolution of miRNAs in bilaterian, cnidarians and sponges, and proposed that the acquisition of new miRNAs has had an important role in the development of new animal organs.


Acknowledgements
We regret that due to space constraints, we are unable to cite the work of many colleagues who have made key contributions to the field. We thank K. Birnbaum for insightful discussions about gene regulation in plants, and V. Ambros, M. Hammell, S. Small and N. Sokol for helpful comments on a preliminary version of the manuscript.

Competing interests statement
The authors declare no competing financial interests.

DATABASES
The following terms in this article are linked online to:
UniProtKB: http://ca.expasy.org/sprot
FURTHER INFORMATION
The Rajewsky laboratory web site: http://www.mdc-berlin.de/rajewsky
NYU Department of Biology and Center for Comparative Functional Genomics: http://www.nyu.edu/locator/biology
SUPPLEMENTARY INFORMATION
See online article: S1 (figure)
Access to this links box is available online.
microRNAs and the immune response

Mark A. Lindsay¹,²

¹ Respiratory Research Group, Wythenshawe Hospital, School of Translational Sciences, University of Manchester M23 9LT, UK
² Biopharmaceutics Research Group, Airways Disease, National Heart and Lung Institute, Imperial College, London SW3 6LY, UK

microRNA (miRNA)-mediated RNA interference has been identified as a novel mechanism that regulates protein expression at the translational level. Recent publications have provided compelling evidence that a range of miRNAs are involved in the regulation of immunity, including the development and differentiation of B and T cells, proliferation of monocytes and neutrophils, antibody switching and the release of inflammatory mediators. In this review, we examine what is presently known of the function and mechanism of action of these miRNAs in the regulation of the innate and acquired immune response.

Introduction

The posttranscriptional regulation of gene expression via microRNA (miRNA)-mediated RNA interference (RNAi) was first identified in Caenorhabditis elegans in 1998 [1] (Box 1 summarizes the pathway of miRNA-mediated interference). Since this initial observation, >700 miRNAs (miRNA registry at www.sanger.ac.uk/Software/Rfam/miRNA/) have been identified in mammalian cells and have been shown to be involved in a range of physiological responses, including development, differentiation and homeostasis [2,3]. Interestingly, it is now clear that miRNAs regulate many aspects of the immune response. This has emerged from studies that revealed selective expression of miR-181a in the thymus and miR-223 in the bone marrow and indicated their involvement in the differentiation of pluripotent hematopoietic stem cells (HSCs) into the various blood cell lineages including B and T cells [4–7]. Subsequent reports have identified functions for individual miRNAs such as miR-150, miR-181a and miR-17–92 during T- and B-cell differentiation [4,8], whereas miR-17–92 and miR-223 are implicated in myeloid production [9–12]. Experimentation has also revealed roles for miRNAs during the activation of the innate and acquired immune response. Thus, miR-146, miR-155 and miR-223 are thought to regulate the acute inflammatory response after the recognition of pathogens by the Toll-like receptors (TLR) [13–16], whereas miR-155 and miR-181a are implicated in B- and T-cell responses [8,13–15]. Although not reviewed in this article (for review, see Refs. [17,18]), miRNAs have also been shown to have a direct antiviral action by targeting crucial proteins and pathways used by hepatitis C virus [19,20], primate foamy virus type 1 [21] and vesicular stomatitis virus [22] (Table 1).

In the following sections, we will review the role of miRNAs during the development and differentiation of immune cells and in the regulation of the innate and acquired immune response (for additional reviews, see Refs. [23–26]). Although expression profiling has indicated a role for a host of miRNAs (Table 1), we shall focus on those for which we have functional and mechanistic information that is predominantly derived from studies using transgenic or knockout mouse models. With regard to the evaluation of mechanism, it must be emphasized that the difficulty in predicting the mRNA targets of individual miRNAs means that this process is often problematic. Thus, miRNAs are believed to either block mRNA translation or reduce mRNA stability after imperfect binding of the guide strand to miRNA-recognition elements (MREs) within the 3’ untranslated region (UTR) of target genes. Originally, it was believed that the specificity of this response was mediated by the ‘seed’ region, which is localized at residues 2–8 at the 5’ end of the miRNA guide strand (see Box 1). However, it now seems that this was an oversimplification and that miRNA targeting is influenced by additional factors such as the presence and cooperation between multiple MREs [27,28], the spacing between MREs [28,29], proximity to the stop codon [28], position within the 3’ UTR [28], AU composition [28] and target mRNA secondary structure [30]. This problem is compounded by the usual problems associated with the determination of the role of a gene, in this case miRNAs, through overexpression or inhibition. This is particularly the case when undertaking overexpression studies in either cells or transgenic mice, because nonphysiological levels of miRNAs expression are likely to lead to downregulation of incorrect targets.

miRNA-16: constitutive modulation of inflammation

Expression profiling has shown that miRNA-16 is found at high levels in most cells and tissues including those involved in inflammation such as monocytes, neutrophils, B cells and both CD4+ and CD8+ T cells [31]. This ubiquitous expression suggests that miR-16 might have a generic function, possibly in preventing cell cycle progression [32]. Interestingly, a publication by Jing et al. [33] has shown that miR-16 is required for the rapid degradation of proteins that contain AU-rich elements (AREs) in their 3’ untranslated region, which includes most inflammatory mediators [33]. This process requires the miRNA processing components Dicer, Ago/eIF2C family members and the ARE binding protein, tristetraprolin and involves binding between the UAAAUAUU sequence in miR-16 and AU-rich sequences in cytokines and chemokines such as tumor necrosis factor α (TNFα), interleukin-8 (IL-8) and IL-6 [33]. Because there are no reports that miR-16 expression is changed in response to immune modulators, in addition to a potential role in cell cycle progression, one could
Box 1. Pathway of microRNA-mediated RNA interference

MicroRNAs (miRNAs) are short double-stranded RNA molecules of ~19–23 nucleotides in length. They are produced from full-length, RNA polymerase II transcripts called pri-miRNA after cleavage by two RNase III enzymes called Drosha [58,59] and Dicer [60], in association with a range of accessory proteins. Initial cleavage by Drosha within the nuclear compartment produces a hairpin RNA of ~65 nucleotides known as pre-miRNAs. These pre-miRNAs are transported into the cytoplasm by exportin 5 [61] and further processed by Dicer [60] to produce the mature miRNAs, which are double-stranded RNA molecules with a length of 19–23 nucleotides. The actions of miRNAs are mediated by the miRNA-induced silencing complex (miRISC), which is composed of multiple proteins including a member of the double-stranded RNA binding protein argonute family (Ago) [62,63]. Using one strand of the miRNA, called the guide strand (the red strand in Figure I), the miRISC is believed to either block mRNA translation, reduce mRNA stability or induce mRNA cleavage, after imperfect binding to miRNA recognition elements (MREs) within the 3′ and 5′ untranslated region (UTR) of target mRNA genes [64–66]. Significantly, this redundancy within the miRNA:mRNA binding region means that individual miRNAs are able to target multiple mRNA targets and has led to speculation that miRNAs might have a similar role to transcription factors [67,68].

Figure I. Pathway of microRNA-mediated RNA interference.

speculate that the high miR-16 levels in inflammatory cells might also restrict the production of inflammatory mediators under nonstimulated conditions and/or that miR-16 binding might cooperate with other miRNA to regulate immunity.

miR-17~92 cluster: a ubiquitous regulator of B-cell, T-cell and monocyte development

The miR-17~92 cluster is located on chromosome 13 and is composed of six miRNAs (miR-17, miR-18a, miR-19a, miR-20a, miR-19b-1 and miR-92-1). Interestingly, a process of duplication and deletion in early vertebrate evolution has resulted in the production of two additional paralog clusters, miR-106a~363 (miR-106a, miR-18b, miR-20b, miR-19b-2, miR-92-2 and miR-363) and miR-106b~25 (miR-106b, miR-93 and miR-25), which are located on chromosome X and chromosome 6, respectively.

Previous studies indicated that miR-17~92 might be involved in the regulation of B cells after reports showing that this cluster is located within a region of DNA that is amplified in B-cell lymphomas [34] and that overexpression of both miR-17~92 and the transcription factor c-myc produced by B-cell lymphomas in mice [35]. Subsequent studies using miR-17~92 and Dicer knockout mice have indicated that this cluster is required for the pro-B to pre-B-cell transition during B-cell development [36,37] (Figure 1). Indeed, expression of miR-17~92 was shown to peak in pre-B-cells, where it inhibited cell death through the downregulation of the proapoptotic protein Bim [36] (Figure 2). Significantly, miR-17~92 knockout mice were nonviable at birth and demonstrated severely hypoplastic lungs and a ventricular septal defect in the heart [36]. For this reason, it was necessary to perform these studies using liver-derived B cells that were grafted into lethally irradiated mice. This contrasted with transgenic mice in which miR-17~92 was selectively induced in both T and B lymphocytes that were viable but that developed lymphoproliferative disease and autoimmunity, ultimately culminating in premature death [38]. As might be expected if miR-17~92 positively regulated B-cell development, these animals were found to have increased numbers of activated B cells. However, the major expansion was in the number of activated CD4+ T cells and to a lesser extent, CD8+ T cells, which implied that miR-17~92 also regulated T-cell development [38]. Importantly, as with the knockout studies, the increase in B- and T-cell population was shown to result from enhanced proliferation and survival after the downregulation of Bim and the tumour enhancer phosphatase and tensin homology (PTEN) [38]. This thereby facilitated the pro- to pre-transition of both B and T cells. In summary, the authors of this report speculated that the phenotypic changes after transgenic expression of miR-17~92 resulted in breakdown in peripheral T-cell tolerance, expansion of the CD4+ T cells that activates B cells, leading to the development of autoimmunity.

In humans, investigations using cord blood CD34+ haematopoietic progenitor cells that had been induced to differentiate into monocytes after exposure to macrophage-colony stimulating factor (M-CSF) indicated that miR-17-5p, miR-20a and miR-106a expressions are involved in monocyte differentiation and maturation [12] (Figure 1). Examination of the mechanism showed that reduced miR-17~5p, miR-20a and miR-106a expression resulted in increased levels of their target protein, acute myeloid leukaemia-1 (AML-1). AML-1 is a DNA-binding subunit of the transcription factor core-binding factor (CBF), which has been shown to induce monocyte differentiation and maturation by increasing the expression of the receptor for M-CSF, as well as other cytokines such as IL-3 and granulocyte macrophage-colony stimulating factor (GM-CSF) [12] (Figure 2). Significantly, AML1 was also found to bind within the promoter region of miR-17~5p, miR-20a and miR-106a and inhibit their transcription. Thus, a reduction in miR-17~5p, miR-20a and miR-106a levels establishes a negative feedback loop involving increased AML1 and M-CSF receptor expression and the subsequent downregulation in miR-17~5p, miR-20a and miR-106a expression [12] (Figure 2).

miRNA-146a: negative regulator of the innate immune response

The miR-146 family is composed of two members, miR-146a and miR-146b, that are located on chromosomes 5 and...
but not miRNA-146b expression in lung epithelial cells because we also observed increased miRNA-146a response does not seem to be restricted to inflammatory bacterial nucleic acids (Figure 2) [14]. Interestingly, this activation of the TLR-3, -7 or -9 in response to viral and contrast, miR-146a expression is not increased after acti-

miRNA-146a identified several potential transcription factor bind-

miRNA Function Transcriptional regulation Targets Refs
miR-15a Decreased expression in chronic lymphocytic leukaemia Bcl-2 [43,77–79]
miR-16 Binds to UA rich elements in the 3’ UTR and induces TNFα mRNA degradation TNFα [33]
miR-21 Increased expression in B-cell lymphoma and chronic lymphocytic leukemia AML-1 [43,80]
mir-17–5p In combination with miR-20a and miR-106a inhibits monocyte proliferation, differentiation and maturation AML-1 [12]
mir-17–92 cluster Regulates pro- to pre- transition during B- and T-cell development Bim, PTEN [36–38]
mir-20a In combination with mir-17–5p and miR-106a inhibits monocyte proliferation, differentiation and maturation AML-1 [12]
mir-24 Inhibits replication of vesicular stomatitis virus [22]
mir-29a Down-regulated in B-cell chronic lymphocytic leukemia Tcl-1 [81]
mir-32 Inhibits replication of primate foamy virus type 1 [21]
mir-93 Inhibits replication of vesicular stomatitis virus [22]
mir-106a In combination with mir-17–5p and miR-20a inhibits monocyte proliferation, differentiation and maturation AML-1 [12]
mir-122 Required for hepatitis C proliferation in liver [19]
mir-125b Expression downregulated by LPS and oscillations in expression after exposure to TNFα [15]
mir-146a Expression induced in macrophages and alveolar/bronchial epithelial following activation of TLR-2, -4 and -5 or exposure to TNFα and IL-1α. NFκB, IRAK1, TRAF6 [14,16]
mir-146b LPS induced expression induced in macrophages IRAK1, TRAF6 [14]
mir-150 Increased expression leads to suppression of B-cell formation by blocking in pro- to pre-B cell transition. Decreased expression in chronic lymphocytic leukemia (CLL) [43,82]
mir-155 Increased expression in Hodgkin and non-Hodgkin lymphomas and chronic lymphocytic leukemia (CLL) [39–43]
mir-155 Required for normal production of isotype-switched, high-affinity IgG1 antibodies in B-cells, determines Th1 and Th2 differentiation and positive regulator of antigen induced responses in T-cells AP-1, PU.1, c-Maf [44–46,48]
mir-155 Increased expression following activation of the innate immune response. Inhibits inflammatory mediator release and stimulates granulocyte and monocyte proliferation AP-1 [13–15,47]
mir-181a Positive regulator of B-cell development and CD4+ T-cell selection, activation and sensitivity. SHP-2, PTPN22, DUSP5, DUSP6 [4,8,49]
mir-196 Induced by IFNβ and inhibits replication of hepatitis C virus [20]
mir-223 Negative regulator of neutrophil proliferation and activation PU.1, C/EBPα, NFI-A, Meff2c, IGFR [8,11,50]
mir-296 Induced by IFNβ and inhibits replication of hepatitis C virus [20]
mir-351 Induced by IFNβ and inhibits replication of hepatitis C virus [20]
mir-431 Induced by IFNβ and inhibits replication of hepatitis C virus [20]
mir-448 Induced by IFNβ and inhibits replication of hepatitis C virus [20]

This table describes all the miRNAs that have been implicated in the immune response. It describes their function, transcriptional factors that are known to regulate their expression and potential targets, including viral targets.

Abbreviations: AML, acute myeloid leukaemia; AP-1, activator protein; Bcl, B-cell lymphoma; Tc1, T-cell lymphoma; C/EBP, CCAAT-enhancer binding protein; DUSP, dual specificity phosphatase; IGFR, insulin-like growth factor receptor; IRAK, IL-1 receptor activated kinase; JNK, c-Jun N-terminal kinase; Maf, musculoaponeurotic fibrosarcoma; Meff, myeloid ELF-1 like factor; miRNA, microRNA; NF, nuclear factor; PTP, protein tyrosine phosphatase; SHP, Src homology 2 domain-containing protein-tyrosine phosphatase; TLR, Toll-like receptor; TRAF, TNF receptor-associated factor.

Table 1. Summary of miRNA involvement in the immune response

10, respectively. Evidence that miR-146a and miR-146b might be involved in the innate immune response was first provided by Taganov et al. [14], who showed increased expression in the human monocyteic THP-1 cell line in response to lipopolysaccharide (LPS) [14] (Figure 2). Subsequent studies have shown that this is a general response in myeloid cells activated through TLR-2, -4 or -5 by bacterial and fungal components or after exposure to the proinflammatory cytokines, TNFα or IL-1ß [13–15]. By contrast, miR-146a expression is not increased after activation of the TLR-3, -7 or -9 in response to viral and bacterial nucleic acids (Figure 2) [14]. Interestingly, this response does not seem to be restricted to inflammatory cells because we also observed increased miRNA-146a but not miRNA-146b expression in lung epithelial cells [16] and airway smooth muscle cells (H. Larner-Svensson et al., personal communication) after IL-1ß stimulation. Inspection of the upstream promoter region of miRNA-146a identified several potential transcription factor binding sites including interferon regulatory factor-3 (IRF3), IRF7 and CCAAT enhancer-binding protein-β (C/EBPβ), although it was shown that NF-κB is responsible for the LPS-induced response [14]. At the present time, the transcriptional regulators of miR-146b are unknown.

To determine the biological role of miRNA-146a and b, examination of the public databases predicted several potential targets for miR-146a and miR-146b including the IL-1 receptor associated kinase 1 (IRAK1) and TNF receptor-associated factor 6 (TRAF6). Significantly, IRAK1 and TRAF6 are known to be part of the common signalling pathway derived from TLR-2, -4 and -5 and the IL-1ß receptor, which led to speculation that increased miRNA-146a
or miR-146b expression might act in a negative feedback pathway. Indeed, this contention was supported by studies showing downregulation in the activity of a luciferase reporter fused to the 3' UTR of IRAK1 or TRAF6 and co-expressed with a miR-146a or miR-146b containing plasmid [14]. Increased miR-146a but not miR-146b expression, was recently confirmed in alveolar epithelial cells and shown to negatively regulate IL-1β–induced IL-8 and RANTES (regulated on activation, normal T-cell expressed and secreted)[16]. Interestingly, increased miRNA-146a expression was only seen at high IL-1β concentrations, which indicated that negative feedback is only activated during severe inflammation and that this might be crucial in preventing potentially dangerous inflammation from spiralling out of control. However, examination of the mechanism showed that this was not caused by downregulation of IRAK1 or TRAF6 but instead occurred at the translational level, through an as yet unidentified mechanism (Figure 2) [16].

**miRNA-155: a regulator of T- and B-cell maturation and the innate immune response**

A single miR-155 transcript has been identified that is processed from within the second exon of the non–protein-encoding gene, bic. Interestingly, the observation that miR-155 is overexpressed in human B-cell lymphomas and studies showing that transgenic expression of miR-155 in mouse B cells produced malignancies have led to speculation that this miRNA is involved in B-cell differentiation and proliferation [39–43].

A central role for miR-155 in the regulation of T- and B-cell responses during the acquired immune response has emerged from studies in knockout mice, because these were shown to be immunodeficient and failed to develop...
a protective response to bacteria after immunisation [44,45] (Figure 1). Interestingly, miR-155 knockout mice also undergo age-related lung airway remodelling, which is characterised by increased collagen deposition, smooth muscle mass and inflammatory cell infiltrate within the broncholaveolar lavage [44]. These phenotypic observations seem to be in part mediated through involvement of miR-155 in B-cell production of isotype-switched, high-affinity IgG1 antibodies and during the development of B-cell memory [45,46]. The reduction in IgG1 antibodies was not the result of impaired somatic hypermutation and class switch recombination. Instead, the observed reduction in the size of the germinal centers suggested that reduced antibody production and B-cell memory was the result of a failure to select high affinity plasma B cells [45,46]. Examination of the mechanism showed that miR-155 attenuated the expression of the transcription factor PU.1, which was shown to downregulate IgG1 levels [46] (Figure 2).

As with B cells, it seems that miR-155 is involved in T-cell differentiation [44,45]. Thus, naïve T cells derived from miR-155 knockout mice were shown to have an increased propensity to differentiate into Th2 rather than Th1 cells, with the concomitant production of Th2 cytokines such as IL-4, IL-5 and IL-10 [44,45]. It has been speculated that this bias results from miR-155 targeting of c-Maf (musculoaponeurotic fibrosarcoma), a transcription factor that is known to be a potent transactivator of the IL-4 promoter, a key cytokine in the development of Th2 cells [44]. With regard to the acute immune response, the T lymphocytes had an impaired response and showed attenuated IL-2 and interferon γ (IFNγ) release in response to antigens [44,45] (Figure 2).

A link between miR-155 and the innate immune response was suggested from studies showing increased expression after LPS (via TLR-4) and lipoprotein (via TLR-2) stimulation in monocytes or macrophages and in the splenocytes of mice that had been inoculated with Salmonella enteritidis–derived LPS [13–15] (Figure 2). In contrast to miR-146a, increased miR-155 expression was also seen after activation of the innate response by viral- and

Figure 2. Role of microRNAs (miRNAs) in the regulation of the immune response. The illustration summaries what is know of the mechanism by which individual miRNAs (red boxes) interact with transcription factors (green circles) and other proteins (yellow boxes) to regulate cell responses (blue circles) during immune cell development and the innate (monocytes, macrophages and neutrophils) and acquired (B and T cells) immune response. Stimulation (↑) or inhibition (↓) should be determined by following how the changes in the individual miRNA expression (red squares) impacts on the indicated biological response (blue circles). AML, acute myeloid leukaemia; AP-1, activator protein; BCR, B-cell receptor; C/EBP, CCAAT-enhancer binding protein; DUSP, dual specificity phosphatase; IGFR, insulin-like growth factor receptor; IRAK, IL-1 receptor activated kinase; JNK, c-jun N-terminal kinase; M-CSF, macrophage colony stimulating factor; Maf, musculoaponeurotic fibrosarcoma; Mef, myeloid ELF-1 like factor; NF, nuclear factor; PTEN, phosphatase and tensin homolog; PTP, protein tyrosine phosphatase; RA, retinoic acid; SHP, Src homology 2 domain–containing protein-tyrosine phosphatase; TCR, T-cell receptor; TLR, Toll-like receptor; TRAF, TNF receptor-associated factor.
bacterial-derived nucleotides including poly(I:C) (via TLR-3) and CpG (via TLR-9), although increased expression in response to IFNγ and IFNβ seemed to be secondary to the autocrine release of TNFα [13]. Tili et al. [15] showed that miR-155 expression oscillated after TNFα stimulation, with an initial drop at 30 min followed by an increase at 60 min [15]. This led them to speculate that miR-155 might exert both positive and negative actions on the expression of target proteins that included the NF-κB signalling proteins, IKKβ and IKKε, as well as FADD (Fas-associated death domain) and Ripk1 (receptor interacting serine-threonine kinase 1) [15,42]. In contrast to miR-146a, miRNA-155 seems to positively regulate the release of inflammatory mediators during the innate immune response because it was shown to increase TNF-α production in human embryonic kidney cells (HEK-293) cells by relieving the 3’ UTR–mediated posttranscriptional inhibition. This observation was supported by studies in Eu-miR-155 transgenic mice that overexpress miR-155 in B cells, which demonstrate an elevated level of serum TNFα and increased susceptibility to septic shock [15]. Interestingly, a recent report has shown that LPS induced strong but transient miR-155 expression in mouse bone marrow cells and indicated that this is likely to drive granulocyte/monocyte expansion [47]. The possible involvement of miR-155 in the development of acute myeloid leukaemia (AML) was revealed from studies of the effect of long-term miR-155 overexpression. In these studies, virus-mediated transfection of miR-155 into hematopoietic stem cells (HSCs) and engraftment into lethally irradiated mice produced some of the pathological features characteristic of myeloid neoplasia [47]. This led these authors to speculate that the well-established link between inflammation and cancer might involve chronic upregulation of miR-155, which could predispose these individuals to the development of myeloproliferative disorders.

The proinflammatory transcription factors, AP-1 and NF-κB, have both been reported to regulate miR-155 expression [13,45,48]. Thus, in macrophages, the action of TLR-3 and TNFα is mediated via AP-1 [13], whereas the response to LPS is via NF-κB [19]. Similarly, BCR cross-linking in a human B-cell line was shown to induce miR-155 expression via activation of the extracellular-regulated kinase (ERK) and c-Jun N-terminal kinase (JNK) pathway and the subsequent recruitment of FosB and JunB to AP-1 binding site [48] (Figure 2).

miRNA-181a: a regulator of B-cell differentiation and CD4+ T-cell selection, activation and sensitivity

Early studies by Chen et al. [4] demonstrated that miR-181a was selectively expressed in thymus-derived B cells and identified a positive role for miR-181a in B-cell differentiation after showing that ectopic expression in mouse hematopoietic precursor cells led to an increased fraction of B cells [4]. Significantly, these overexpression studies also showed a reduction in CD8+ killer T cells, which suggested a potential role in the regulation in T-cell differentiation (Figure 1). However, despite subsequent studies that showed dramatic changes in miR-181a expression during the various stages of T-cell differentiation, these observations were often contradictory, and the role of miR-181a during T-cell development is still inconclusive [8,49].

In contrast to T-cell differentiation, miR-181a has been shown to positively regulate the T-cell receptor (TCR)-mediated response to antigen in CD4+ T cells. Thus, by measuring changes in the intracellular Ca2+ transient and IL-2 release, these investigators showed that miR-181a overexpression amplified the strength and sensitivity of TCR-mediated activation [8] (Figure 2). In addition, these studies showed a qualitative change that permitted T cells to respond to antagonists as agonists and indicated that miR-181a, at least in part, regulated the positive and negative selection of T cells during thymic development [8]. Examination of the mechanism demonstrated that this did not result from changes in the expression of surface receptors but instead involved the coordinated downregulation of multiple phosphatases. These included the tyrosine phosphatases, Src homology 2 domain–containing protein-tyrosine phosphatase (SHP)-2 and protein tyrosine phosphatase (PTP)-N22 and the ERK-specific, dual specificity phosphatases (DUSP)-5 and -6, which are all known to negatively regulate the TCR signalling pathway. Specifically, it has been speculated that changes in miR-181a might act like a rheostat by regulating protein phosphorylation levels. In this model, increased miR-181a and reduced phosphatase levels would lead to increased basal phosphorylation of protein kinases such as Lck and ERK. Under these conditions, T cells would exhibit increased TCR signalling and a reduced T-cell activation threshold [8] (Figure 2).

miRNA-223: a regulator of neutrophil proliferation and activation

Expression profiling has shown that miR-223 expression is associated with myeloid cells in the bone marrow and that increased expression of miR-223 is involved in the differentiation of myeloid precursors into granulocytes such as neutrophils [4,9,11] (Figure 1). Interestingly, recent observations in miR-223 knockout mice, which were shown to be viable and healthy, revealed increased numbers of granulocyte progenitors in the bone marrow and hypermature neutrophils in the circulation. This implied that miR-223 was involved in the negative regulation of maturation but not differentiation of granulocytes [11]. Mechanistic studies showed that the action of miR-223 is probably mediated through downregulation of either myeloid ELF-1-like factor (Mef)-2c, a transcription factor that promotes myeloid progenitor proliferation, or the insulin-like growth factor receptor (IGFR) [11] (Figure 2). Paradoxically, an earlier report by Fazi et al. [9] reported that miR-223 was a positive regulator of granulocyte differentiation. However, these studies were performed in a retinoic acid (RA)-treated acute promyelocytic leukemic cell line and the role of miR-223 might be dependent on the timing of miRNA overexpression or miRNA inhibition [9]. Functional studies in circulating neutrophils obtained from the knockout animals indicated that miR-223 was not involved in extravasagation, migration or phagocytosis of the bacterium Escherichia coli but that reduced miR-223 resulted in enhanced oxidative burst and killing of the fungi, Candida albicans (Figure 2). In addition, these animals spontaneously developed inflammatory lung pathology and exhibited exaggerated tissue destruction after...
end of the endotoxin challenge. Of relevance to this observation, we have recently demonstrated a rapid and selective increase in miR-223 expression in lung and airway epithelia after exposure of mice to aerosolized endotoxin [10]. This, if miR-223 also acts as a negative modulator of the inflammatory response in these cell types, this might contribute to the lung pathologies observed in miR-223 knockout mice.

Multiple transcription factors have been implicated in regulating miR-223 expression during granulocyte differentiation and maturation. In the report showing positive regulation of RA-induced granulocyte differentiation by miR-223, Fazi et al. [9] showed that miR-223 expression was increased by C/EBP-α and attenuated by nuclear factor (NFJ-A). This investigation also showed that miR-223 targeted the downregulation of NFJ-A, which would thereby established a negative feedback loop that promoted granulopoiesis [9]. A role of C/EBPα in regulating miR-223 expression was confirmed by Fukao et al. [50], who also demonstrated increased miR-223 expression after activation of another myeloid specific transcription factor, PU.1 (Figure 2).

Conclusion

Although investigators have only recently identified a role for miRNA-mediated RNA interference in the immune response, several interesting trends have emerged. The first indication that miRNAs, including miR-155, miR-181a and miR-223, might regulate the development of immune cells was their differential expression in the thymus and bone marrow and reports that changes in expression were associated with the development of various leukaemias [4,39–43,51]. Significantly, subsequent studies have confirmed that miR-155 [44,45,47], miR-181a [4,8,49] and miR-223 [9,11,50], as well as those miRNAs derived from the miR-17–92 cluster [36–38], are indeed involved in the development and differentiation of lymphoid and myeloid cells. In addition, mechanistic studies have suggested that this developmental process is crucially dependent upon interactions between miRNAs and transcription factors (Figure 2), which has also been observed during development of other tissues such as skeletal muscle and heart [52–55]. Given that aberrant expression of miRNAs has been associated with development of leukaemias, how might this be linked to development of immune cells? Interestingly, increased expression of miR-155, miR-181a and miR-17–92 has been shown to promote the survival and proliferation of both lymphoid and myeloid cells. Following this observation, Xiao et al. [38] speculated that somatic mutation and amplification of these miRNAs might result in a reduction in the activation threshold, increased proliferation and enhanced survival. In turn, this could lead to expansion of this cellular population and a higher probability of accumulating subsequent mutations that are required to produce the malignant transformation [38].

At the present time, miRNAs seem to regulate the responses associated with acquired immunity, such as antibody production [45,46] and inflammatory mediator release [44,45] through their impact on the development and differentiation of immune cells. Thus, there are no reports that activation of TCR and BCR signalling influences T- and B-cell responses through changes in miRNA expression. This contrasts with activation of the innate immune response that has been shown to induce rapid changes in the levels of miR-146a [14,16] and miR-155 expression [13–15,47]. These miRNAs have been reported to regulate acute inflammatory responses such as the release of IL-8 and RANTES either through the targeting of proteins involved in the signalling pathway (IRAK1, TRAF6, IKKα, IKKβ) or by modulating the translation of the inflammatory mediators themselves (i.e. TNFα, IL-8 and RANTES). At the present time, the mechanism by which miR-146a and miR-155 influence translation is unknown, although, like miR-16 [33], these might regulate mRNA stability via AU-rich elements that are found in the 3′ UTR of many inflammatory mediators [34]. Of relevance, a recent in silico survey of the interactions between miRNAs and immune genes unexpectedly found preferential targeting of proteins involved in the regulation of AU-rich elements and miRNA metabolism [57]. Furthermore, this report indicated that major targets of miRNAs are transcription factors, cofactors and chromatin modifiers rather than receptors, their ligands or inflammatory mediators [57].

Overall, there is now compelling evidence that miRNAs are involved in the regulation of the immune response, and modulation of their activity might ultimately provide a novel therapeutic approach in the treatment of inflammatory disease and certain leukaemias (Box 2).

Acknowledgements

M.A.L. is supported by the Wellcome Trust University Award (076111).
67 Lewis, B.P. et al. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120, 15–20
72 Zeng, Y. et al. (2002) Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. Mol. Cell 9, 1327–1333
81 Pekarsky, Y. et al. (2006) Tcl1 expression in chronic lymphocytic leukemia is regulated by miR-29 and miR-181. Cancer Res. 66, 11590–11593