The great leap forward

Major evolutionary jumps might be caused by changes in gene regulation rather than the emergence of new genes

One of the great challenges in molecular and evolutionary biology is to explain the link between giant evolutionary leaps, such as the colonization of land by plants or the emergence of vertebrates, and the underlying genetic and genomic changes. For a long time, biologists thought that such profound changes in phenotype would be accompanied—if not driven by—equally dramatic upheavals at the genetic level. Similarly, the emergence of the flowering plants within the plant kingdom, or mammalian vertebrates, must surely have been marked by recognizable changes in the genome. Yet in fact, it turns out that the genomic changes that enabled these evolutionary developments were far more subtle—it is the regulation, rather than the modification or creation, of genes that has driven macroscopic events throughout evolution.

...it is the regulation, rather than the modification or creation, of genes that has driven macroscopic events throughout evolution

It was natural for scientists to assume that protein-encoding genes would be at the heart of evolutionary progress. After all, evolution has created organisms with radically different appearances and functional components. For example, the cranium and spinal column or the liver, which represent major improvements in structure and metabolic function, were first seen in vertebrates. However, these developments were not caused by the sudden appearance of new genes or through massive mutations of existing genes. In fact, protein-encoding genes, which account for less than 2% of the genome in many higher organisms, tend to be highly conserved and are thus unlikely to lead a phenotypic revolution on their own. Moreover, vertebrates have almost identical numbers of genes—slightly less than 30,000—with remarkable similarity across

diverse species. Mice and humans, for example, have many genes for which sequences are 90% identical, yielding protein products that are almost indistinguishable.

Similarly, the theory of genome duplication, which was first seriously developed in 1970 by the Japanese-born geneticist Susumu Ohno (1970), also cannot properly explain the great phenotypic changes that have occurred during the 75 million years since the human and mouse lineages split. Duplications, either of chromosomal regions or of whole genomes, occur naturally, for example, through errors in chromosomal recombination during reproduction. Historically, genome duplication was thought to be a mechanism of evolutionary change because it creates redundant copies of protein-encoding genes with impunity; these copies can have mutations, but the cell will continue to function as it retains an original version of the gene that was duplicated. Usually, mutations are deleterious, but the duplicated genes are not generally expressed and so do no harm with no selective pressure acting against them. Occasionally, however, a mutation can result in a potentially useful modification of the gene, which can then be activated and selected. This, it is argued, could create relatively rapidly new functions for genes and thus explain major leaps in evolution.

Major gene duplication events certainly occur frequently in plants and microorganisms. A famous example is the complete duplication of the yeast genome, which is thought to have occurred more than 100 million years ago (Rikke *et al.*, 2003), after which the two parts evolved independently. Consequently, gene or genome duplication could create the evolutionary templates for more sophisticated developments in multicellular eukaryotes. Indeed, genome-wide comparisons of duplications between humans and chimpanzees indicate that these might have had an important role in the evolution of higher primates that eventually led

to *Homo sapiens* (Cheng *et al*, 2005). Even so, there is not yet sufficient evidence that genome duplications alone can explain rapid evolutionary advances.

Instead, another idea is gaining momentum—namely, that changes in gene regulation, rather than in genes themselves, have driven some of the major evolutionary leaps. Recent work by Philip Donoghue and colleagues at the University of Bristol in the UK suggests that, although genome duplication is important for creating new genes on an ongoing basis, great functional or structural developments also require major regulatory changes that affect whole gene networks (Heimberg *et al*, 2008). Their work provides circumstantial evidence that small non-coding RNAs called microRNAs (miRNAs)—small molecules that consist of about 22 nucleotides—are a major causal factor for structural change and increasing complexity among organisms.

According to Gill Bejerano, Assistant Professor of Developmental Biology and of Computer Science at Stanford University (Palo Alto, CA, USA), miRNAs are now thought to be one of at least three crucial drivers of genomic evolution. The first drivers are protein-encoding genes themselves—which Bejerano described as the

‘big beasts’ of evolution—around which the rest of the genome revolves. The second drivers are miRNAs, which regulate the activity of messenger RNAs (mRNA) after transcription. The third are other non-coding regions of the genome that make up the so-called *cis*-regulatory network, which controls the production of mRNA and thereby of proteins according to factors such as cell type or environmental conditions.

Exactly how these three mechanisms interact remains to be fully elucidated but some of the basic principles have recently come to light, particularly concerning the role of miRNA in mediating the production of mRNA and therefore of proteins. Research conducted under Nam-Hai Chua, head of the Laboratory of Plant Molecular Biology at Rockefeller University (New York, NY, USA), has highlighted the importance of miRNAs for turning off a pathway when it is not wanted. Although Chua’s work focuses on the role of miRNAs in plants, the underlying principles of operation seem to be the same in all organisms.

“If you drive a car, you need a brake as well as an accelerator,” Chua explained. “The point we were trying to make is that when a hormone triggers a certain pathway, and messenger RNA is induced, then when you withdraw the hormone, you need some

desensitisation mechanism to switch off, to tell the cell the signal is no longer there” (Guo *et al*, 2005). This process of switching off a pathway must involve two coordinated steps: degradation of the already-present protein, and ceasing the manufacture of the mRNA that is driving production of the protein. As Chua pointed out, degrading a protein with one hand while manufacturing it with the other would mean wasting energy and resources, as well as failing to switch off the process efficiently.

The implication is that the emergence of miRNA coupled with *cis*-regulatory elements introduced a more effective mechanism for managing and coordinating pathways, which, in turn, allowed more complex systems and structures to emerge like the liver, pancreas or the brain. It remains to be established whether miRNAs enabled major evolutionary leaps, but there is growing evidence that they at least played a significant part, as Donoghue’s work has shown. “We know little about miRNAs, but the emergence of fundamental novelty in the miRNA repertoire correlates with fundamental novelty in organismal complexity,” Donoghue commented. “We also know that miRNA expression can be organ-specific and

that there is great evolutionary conservation of miRNAs. In addition we know that miRNAs add an entire level of regulation in development, hitherto unforeseen."

...protein-encoding genes [...] tend to be highly conserved and are thus unlikely to lead a phenotypic revolution on their own

Donoghue also pointed out that, although miRNAs already existed in invertebrate animals, they seem to have become much more numerous among vertebrates, which suggests that a high rate of miRNA acquisition coincided with rapid phenotypic evolution early in vertebrate history—in both cases far outstripping the progress during any other episode in the evolution of chordates.

There is also evidence that miRNAs are associated with major structural and functional changes in plants. The first relevant observation is that miRNAs are absent from fungi and marine plants, but are abundant in all land plants—which have acquired greater complexity such as leaf growth, flowering and exudation of roots. These phase changes are triggered and timed by miRNAs in response to environmental cues, according to Scott Poethig and colleagues at the University of Pennsylvania (Philadelphia, PA, USA; Poethig, 2003).

It is still not clear whether miRNAs evolved in a common ancestor of both plants and animals, or independently in line with a 'two-origin' hypothesis. However, the hypothesis that miRNAs only evolved in multicellular organisms—driving the evolution of complex pathways and organs—has been contradicted by two separate studies, both of which show that miRNAs also evolved in single-celled eukaryotes—although they have not been found in prokaryotes (Zhao *et al*, 2007; Molnár *et al*, 2007). This information opens the door either to a 'three-origin' hypothesis, in which miRNAs evolved independently in algae, plants and animals, or to a 'single-origin' hypothesis, in which they evolved just once in an ancestral unicellular organism. Either way, it seems that their incidence and sophistication increased enormously in land plants and vertebrates.

There is also some debate about whether it is miRNAs or other non-coding RNA or

DNA segments involved in *cis*-regulation that have played the most important role in driving evolution. There is some evidence from considering the phenomenon of pleiotropy, which occurs when a single gene affects several traits, that *cis*-regulation might be more readily available as raw material for major adaptive shifts.

Pleiotropic genes and their regulation are highly conserved because any change would almost certainly be catastrophic to one or more of their functions. Therefore, the paradox is that, if miRNAs are to be successfully used to coordinate the activity of several genes within a pathway—that is, they are pleiotropic in function—then their function must be conserved and they cannot therefore be invoked to facilitate further evolutionary change. Any alteration in the sequence of a particular miRNA would have to be coordinated with all the genes it regulates, which is almost impossible for evolutionary mechanisms to achieve. "The more genes each miRNA regulates simultaneously, the more constrained it becomes," Bejerano said.

Instead, Bejerano argues that the *cis*-regulatory network is less constrained than miRNAs, in the sense that if a given gene has several different functions, the expression in each case is controlled by a separate *cis*-regulatory element. "If you have a gene involved in brain, heart and limb development, you can very often find one or more individual *cis*-regulatory elements responsible for turning it on in the brain, another subset for turning it on in the heart, and yet a third for limbs," he said. "Therefore, if you want to experiment with limbs while not messing up brain and heart, *cis*-regulation seems the way to go."

Bejerano suggests that *cis*-regulation is likely to have been a far simpler evolutionary mechanism for driving adaptive change than miRNAs. But, although miRNAs might be constrained by pleiotropy, they still have a unique role in turning off gene expression after translation. They must therefore have had a significant role in evolutionary adaptation in the more advanced eukaryotes, as adaptation necessitates substantial changes in the ability of genetic machinery to 'apply the brakes', as Chua put it. The big question is whether miRNAs are triggers for such adaptation, or are merely led by them.

According to John Bowman at the School of Biological Sciences at Monash University in Melbourne, Australia, miRNAs are complementary to the entire target site on the

mRNA of the gene they are regulating in plants. In other words, all 22 or so nucleotides of the miRNA match a sequence of similar length in the target mRNA. But in the case of vertebrates, only about seven nucleotides of the miRNA complement their target. This would explain the observed differences in the roles of miRNA in plants and animals according to Bowman, as animal miRNAs have more freedom to evolve through mutation because fewer nucleotides need to change to match a new target on the genome. This flexibility has opened a larger number of genes up to miRNA regulation in animals—about one-third of the total set of genes in humans, for example. By contrast, plant miRNAs have only managed to converge on a small number of target sites—and thus fewer genes are subject to miRNA regulation. The result is that the relationship between the miRNA and the target gene in plants is highly conserved—meaning that, although miRNAs have acquired important roles in major changes over evolutionary history, they are now fairly stable in plants.

...degrading a protein with one hand while manufacturing it with the other would mean wasting energy and resources, as well as failing to switch off the process efficiently

In vertebrates, however, Bowman argues that miRNAs have continued to drive rapid evolutionary change. This view is supported by evidence of adaptations that have occurred in vertebrate classes and individual species, and that were accompanied by an expansion of miRNAs. For example, a team from the Chinese Academy of Sciences (Beijing, China) sequenced a cluster of several individual miRNAs linked to the X chromosome in mammals. Rodents and dogs have single copies of these miRNAs, but primates have several copies in some cases (Zhang *et al*, 2007) and some species of primate have more copies than others; in the case of one particular miRNA—known as miR-514—humans have four copies, chimpanzees three, and other primates just one. These miRNAs are expressed preferentially in the testis and have evolved rapidly in primates, according to the Chinese group, which suggests that they might have had an active role in the development of male reproduction. This could be the genetic

mechanism underlying the high level of reproductive competition among primate males, but until the actual function of these X-linked miRNAs is established, the case remains open to debate.

There is a growing understanding of how miRNAs might have been recruited and how they have proliferated, especially in vertebrates. miRNAs probably began life through a phenomenon known as exaptation, whereby regions of non-coding DNA acquire function by interacting with so-called 'trans-acting factors' within the genome. A trans-acting factor is a DNA sequence, normally containing a gene, which codes for RNA or a protein that, in turn, acts on another part of the genome. The RNA or protein thus acts as the intermediary, performing the trans-action. The target DNA might be a retroposon—an element that was inserted into the genome at some earlier point in evolution through reverse transcription, but is unable to perform any encoding itself. When the trans-acting protein or RNA encounters a retroposon, or a region of DNA with a similar sequence, it might interact with it and recruit it for some new regulatory function. In this way, a target region with no function previously can become a miRNA, or for that matter, a cis-regulatory element (Lowe *et al*, 2007).

Cis-regulators can also evolve by exaptation and might also have had an important role in major evolutionary developments. One such development probably occurred in the mammalian forebrain. A Japanese group from the Tokyo Institute of Technology has shown that AmnSINE1, a member of a group of retroposons called SINEs (short interspersed nuclear elements), is conserved among all mammals and has a specific and essential role in development of the forebrain (Sasaki *et al*, 2008). According to Norihiro Okada, the senior author of the study, this is an example of exaptation, as

...the emergence of miRNA coupled with cis-regulatory elements introduced a more effective mechanism for managing and coordinating pathways, which, in turn, allowed more complex systems and structures to emerge

"AmnSINE1s got function for the formation of mammalian brain in a common ancestor of mammals, probably because these SINEs originally had sequences very similar to those which can interact with trans-acting factors responsible for mammalian brain formation." Since then, AmnSINE1 has diverged among mammalian species, perhaps reflecting the differing selective pressures on forebrain development.

As Okada pointed out, all vertebrates have almost exactly the same number of genes, many of which are highly conserved. The evolution of higher organisms must therefore have required the generation of new gene-expression networks, particularly those involved in functional development. miRNAs and cis-regulatory elements could well be the regulatory partners that were involved in these developments, with non-functional, non-coding regions of the genome providing the raw material. If so, there is plenty of scope for further evolution, as most of the genome in higher organisms is made up of non-coding DNA—formerly known as 'junk DNA'.

Much research remains to be done and the benefits will not only be academic. Both miRNAs and cis-regulatory elements hold great therapeutic potential, in particular in terms of gene therapy. Rather than trying to target individual genes—to block the effect of the allele causing disease—a relevant miRNA could be used to turn off the offending gene altogether. Indeed,

therapies based on the manipulation of both miRNAs and cis-regulatory elements could be used to manipulate whole pathways in precise ways that would be impossible with approaches that target single protein-encoding genes.

REFERENCES

- Cheng Z *et al* (2005) A genome-wide comparison of recent chimpanzee and human segmental duplications. *Nature* **437**: 88–93
- Guo HS, Xie Q, Fei JF, Chua NH (2005) MicroRNA directs mRNA cleavage of the transcription factor NAC1 to downregulate auxin signals for *Arabidopsis* lateral root development. *Plant Cell* **17**: 1376–1386
- Heimberg AM, Sempere LF, Moy VN, Donoghue PCJ, Peterson KJ (2008) MicroRNAs and the advent of vertebrate morphological complexity. *Proc Natl Acad Sci USA* **105**: 2946–2950
- Lowe CB, Bejerano G, Haussler D (2007) Thousands of human mobile element fragments undergo strong purifying selection near developmental genes. *Proc Natl Acad Sci USA* **104**: 8005–8010
- Molnár A, Schwach F, Studholme DJ, Thuenemann EC, Baulcombe DC (2007) miRNAs control gene expression in the single-cell alga *Chlamydomonas reinhardtii*. *Nature* **447**: 1126–1129
- Ohno S (1970) *Evolution by Gene Duplication*. Heidelberg, Germany: Springer
- Poethig RS (2003) Phase change and the regulation of developmental timing in plants. *Science* **301**: 334–336
- Rikke BL, Clifton PF, Johnston K, Piskur J (2003) Yeast genome duplication was followed by asynchronous differentiation of duplicated genes. *Nature* **421**: 848–852
- Sasaki T *et al* (2008) Possible involvement of SINEs in mammalian-specific brain formation. *Proc Natl Acad Sci USA* **105**: 4220–4225
- Zhang R, Peng Y, Wang W, Su B (2007) Rapid evolution of an X-linked microRNA cluster in primates. *Genome Res* **17**: 612–617
- Zhao T, Li G, Mi S, Li S, Hannon GJ, Wang XJ, Qi Y (2007) A complex system of small RNAs in the unicellular green alga *Chlamydomonas reinhardtii*. *Genes Dev* **21**: 1190–1203

Philip Hunter

doi:10.1038/embor.2008.115