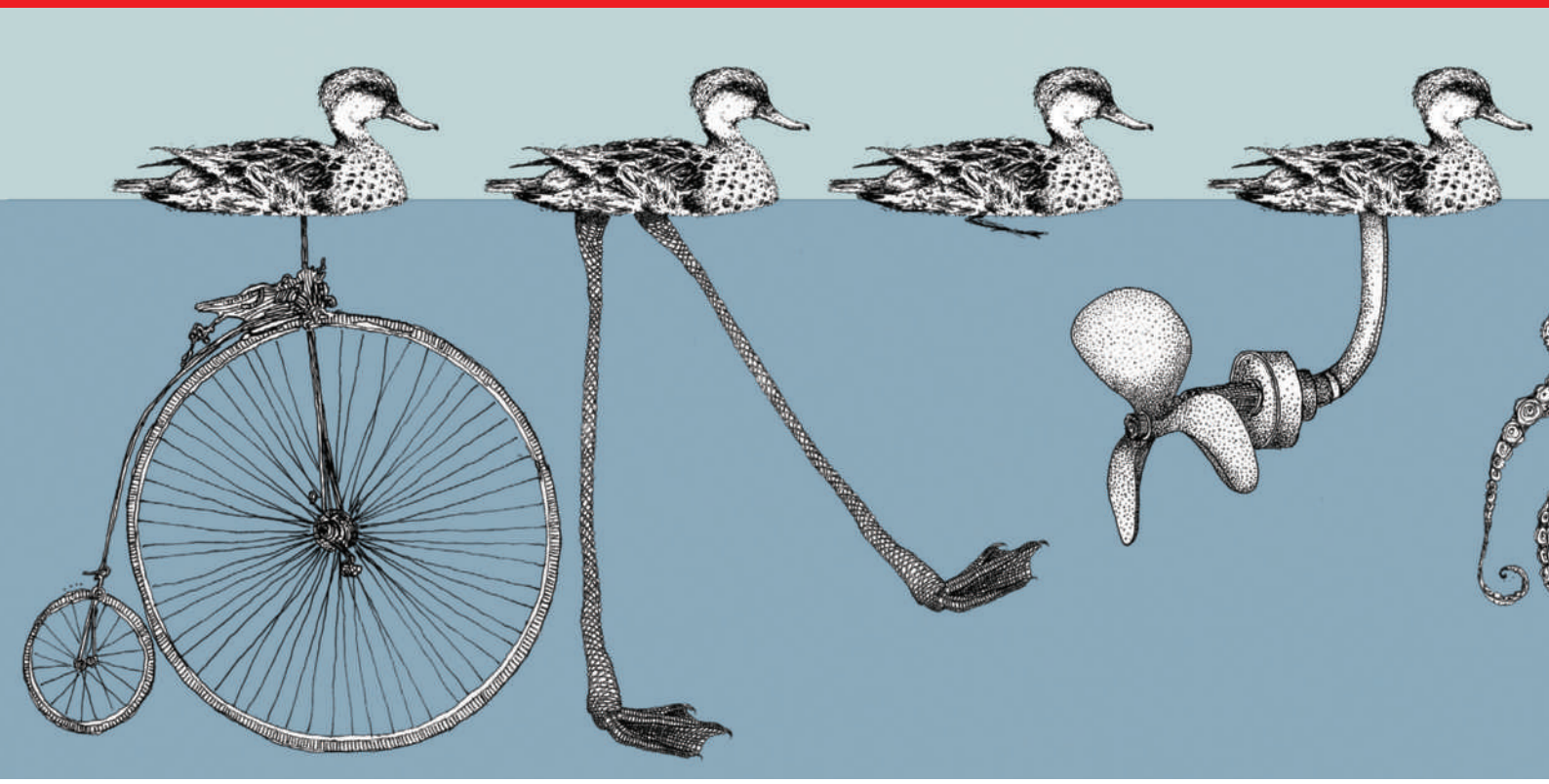


Beneath the surface

You might think that once evolution has found one way to get something done, it will stick with it. But similar physical forms can hide radically different wiring, finds **Tanguy Chouard**.

Darwin200

"It is not birth, marriage or death, but gastrulation, which is truly the most important time in your life," British embryologist Lewis Wolpert famously said. That's when the primordial sheets of embryonic cells are instructed to fold inwards on their way to becoming more specialized tissues. The process is governed in part by a group of cells called the Spemann organizer, and developmental biologists, echoing Wolpert's view, thought that you couldn't mess with something as important as the organizer.

So when, in the early 2000s, those biologists tried to find the Spemann organizer in tunicates they were in for a big surprise. Tunicates — also known as sea squirts — are humans' closest invertebrate cousins. They have tadpole-like larvae that closely resemble miniature vertebrate embryos and so were expected to build their bodies in the same way. But they don't. Most of the 'organizer genes' are there in the tunicate genome, but they are expressed elsewhere in the embryo and do dramatically different things[1]. It's as if you had found a car in which components of the engine were scattered all over the back seat — but the car still worked.

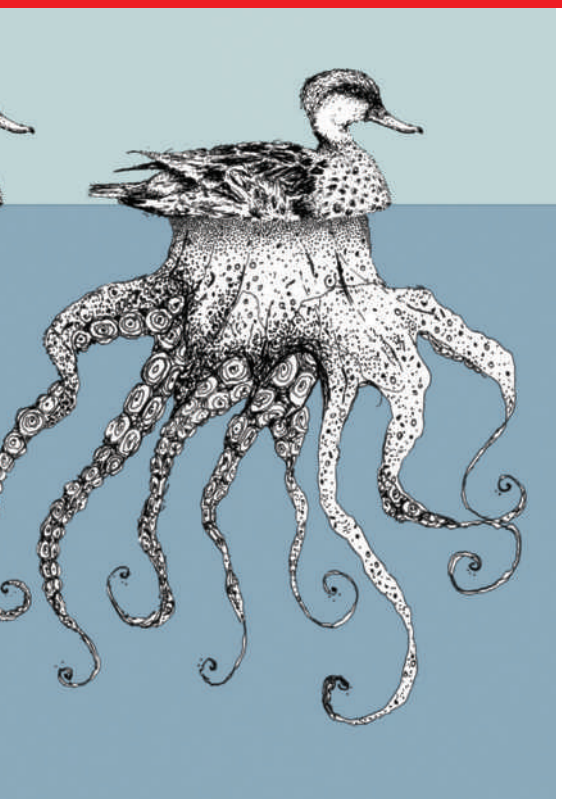
Many biologists, consciously or not, tend to see living systems as optimally tuned. If one species has a complex solution to a difficult problem (such as a Spemann organizer for building a swimming tadpole), it is often assumed that the way genes are orchestrating that process will be the same in related species. That idea has only been strengthened by the discovery that many genes, and the proteins they encode, display a stunning degree of conservation across hundreds of millions of years of evolutionary time.

What is becoming clear, however, is that although most genes might be conserved, the regulatory connections that control their expression might not be. Closely related species can connect up their genes in very different regulatory networks, while keeping the end result deceptively unchanged. "Problems with many solutions are the rule rather than the exception in living systems," says Andreas Wagner, a bioinformatics expert at the University of Zurich, Switzerland. Now researchers are trying to understand how evolution finds the solutions it does, and why. Some think that this 'underground' variation was selected for. Some think that it appeared by chance. And regardless of how it arose in the first place, some believe that a variety of regulatory networks may offer an evolutionary advantage in the future. Trying out many different designs 'underground' could provide a hidden source of evolutionary innovation, variation on which organisms might draw when faced with new challenges.

All these ideas present a major challenge to those who study 'evo-devo', the evolution of developmental processes. Researchers can no longer conclude that two organisms are built in the same way by considering one gene or even one signalling pathway at a time. They must consider the entire system with its inputs, outputs and the connections in between. Where evo-devo has met systems biology, a new discipline — systems evo-devo — has emerged. "I've never been too fond of any of those buzzwords," says Patrick Lemaire at the Developmental Biology Institute of Marseilles in France, "but if this means bringing more rigour to evo-devo, making it less descriptive and more mechanistic, then, yes, a systems framework seems useful."

Conservative ideology

It is hard to shake the belief that conservation runs deep. From the 1980s onwards, DNA sequencing and gene-knockout technology revealed the extent to which genes and their functions were conserved. These techniques showed that many of the genes that determine the animal body plan are virtually identical in both structure and function in creatures that, on the outside, have little in common. The expression patterns of *Hox* genes, for example, specify the same positional values in the head-to-toe body axis of fruitflies and mice. Such discoveries strengthened the intuitive idea that when



K. LEMON

gene sequence and organization are conserved, so too is gene function — and that conserved function implies conserved genetics. But in the past few years it has become apparent that the gene–function relationship is not so clear cut. A given function can result from diverse combinations of the same genes, or different genes, even in very closely related species.

Molecular dissections of the extremely conserved body plan of insects — head, thorax and abdomen — provide a striking example of how different genes and wiring can lead to very similar endpoints. In the fruitfly *Drosophila*, the *Hox*-related gene *Bicoid* is essential to establishing this form. Work starting in the 1980s has shown that messenger RNA coding for the Bicoid protein is glued to the front end of the egg. Here, the protein is translated after fertilization, diffuses from one end of the embryo to another and orchestrates the activity of several other key genes involved in distinguishing head from tail. In fact, Bicoid does so many wonderful things that it was widely assumed to be fundamental to insect development.

Not so. Starting in the late 1990s, sequence data made clear that the *Bicoid* gene, although crucial in *Drosophila*, is absent from most other insects. This apparent paradox is resolved only when considering gene regulatory networks as a whole. The labs of Reinhard Schroeder, of the University of Tübingen, Germany, and Claude Desplan, of New York University, for example, found that several genes that are activated by Bicoid after fertilization in fruitflies are deposited into the egg before fertilization in the flour beetle[2] and a parasitic wasp[3]. The lack of *Bicoid* is compensated by minor changes in the action of several other genes in the developmental network.

So whether you think about the Spemann organizer or the *Bicoid* gene, the same rule seems to apply: there are many combinations of contributing factors that can reach the same outcome. "You can't understand any of this if you think at the single-gene level," says Lemaire. The information that determines biological function lies at a higher, more abstract level, in the entire network of genes, proteins and other factors that each act on the others in a series of nonlinear feedback loops. The body plan or feature that results is what scientists who study complex systems call an 'emergent property' — one that is more than the sum of its parts.

"You can't understand any of this if you think at the single-gene level."
— Patrick Lemaire

It all feels very counter-intuitive. Maybe it is because man-made machines are sensitive to individual component failure, that biologists tend to assume that genetic networks are similarly constrained, and that changing their wiring piecemeal would mess up the whole system. So one of systems evo-devo's key questions is how the genes and their interactions can change while keeping the output of the system as good as, or better than, it was before. To find out, researchers need an experimental system simple enough that many species could be analysed in the finest molecular details. Sexual differentiation in yeast offers such a system, one conserved across species that diverged up to about a billion years ago.

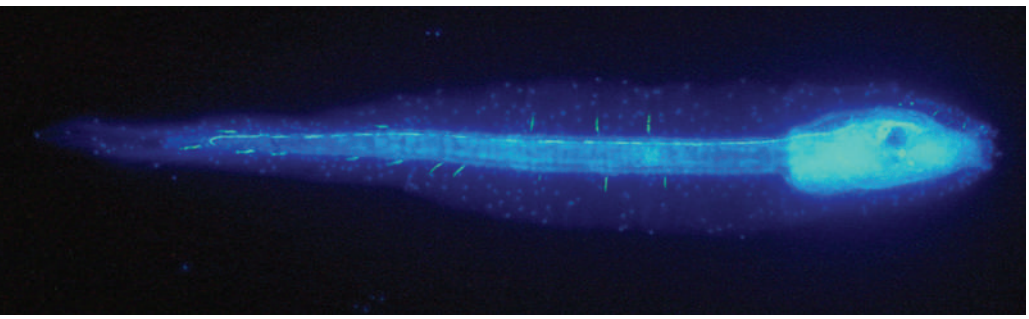
Most yeasts have two mating types: 'a' cells express 'a' genes, and 'α' cells express 'α' genes. In the human pathogen *Candida albicans*, a-specific genes must be actively turned on by a DNA-binding protein called a2 before they are expressed. But in the brewer's and baker's yeast *Saccharomyces cerevisiae*, a-specific genes have to be turned off by a repressor protein α2 — their default setting is on. The circuitry works in completely different ways and *S. cerevisiae* doesn't even have a copy of the a2 gene vital to *C. albicans*. But these two systems didn't just fall from the sky, they must have evolved from a common ancestral system.

To find out how, Alexander Johnson and his team at the University of California, San Francisco, teased out the way that a-genes are regulated in 16 yeast species whose genomes had been sequenced and could thus be reliably ordered on a phylogenetic tree[4]. They identified the regulatory DNA controlling many of the a-genes and compared the binding sequences for a2, α2 and a third DNA-binding protein called MCM1, which associates with either a2 or α2 depending on the yeast species. They also looked for changes in the way the proteins interact with each other. They reasoned that changes in the proteins and their binding sites might reveal how one regulatory system incrementally changes to give rise to the other.

And they did. *Kluyveromyces lactis* — a yeast used in the cheese industry — carries what is likely to have been the transition state between the ancestral *C. albicans* logic and the *S. cerevisiae* scheme. Some of its a-genes bind a2, some bind α2 and some bind both. At the same time, MCM1 is able to bind both a2 and α2. "This is absolutely terrific work," says Sean Carroll, a developmental biologist at the University of Wisconsin in Madison. "They took enough snapshots to show us the whole movie of the evolutionary process." The key to this overhaul is functional redundancy: by possessing more than one regulatory pathway to achieve the same function, it is possible for one to change while the other keeps the system running smoothly. It's like moving around a boat in rough weather: you always keep at least one of two clips hooked to the lifeline.

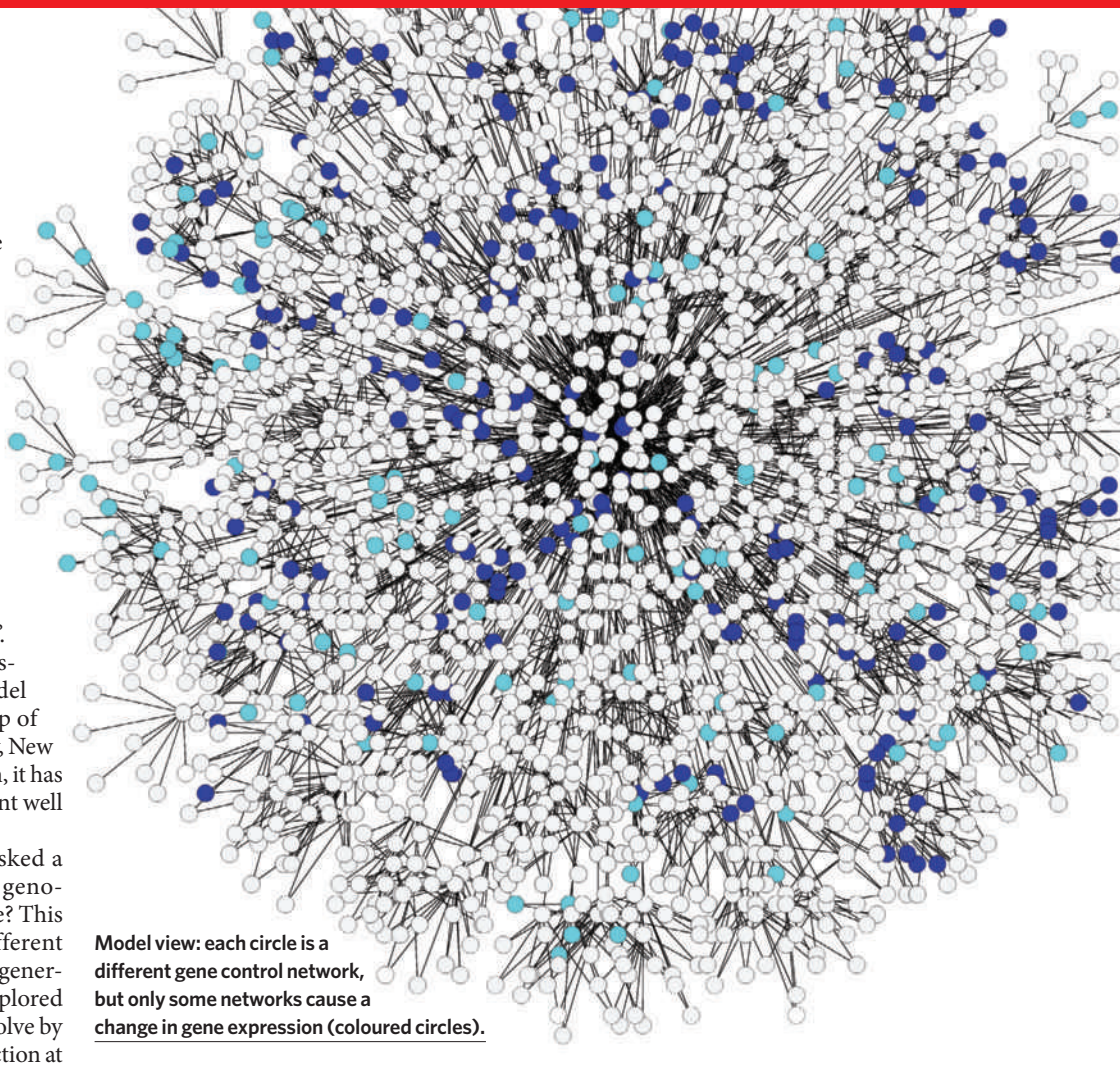
Johnson's yeast studies revealed how an organism might switch from one regulatory network to another without losing fitness in between. But such experimental studies can only deal with a couple of genes at once. So Wagner and two French collaborators turned to computer simulation to ask how far successive small steps can take a large gene-regulatory network away from its starting point[5].

Wagner's model consists of a number of genes, each of which codes for a regulatory protein that activates or represses one or more of the other genes. In the virtual embryo with



S. DARRAS

Tunicates such as *Ciona intestinalis* build a tadpole differently from all their vertebrate cousins.



Model view: each circle is a different gene control network, but only some networks cause a change in gene expression (coloured circles).

S. CILIBERTI AND A. WAGNER

which the model starts, every gene has a given level of expression. The model then calculates what that level of expression, for each gene, will do to all the other genes, and makes the necessary changes. That gives a new pattern of expression — and again the model sees what that means in terms of regulating each gene, until the system settles down to a steady state — the end point of the development of the 'organism'. Each 'genotype', the set of instructions dictating which gene does what to which other gene, results in a different final state, or 'phenotype'. (Systems that never settle down are discarded as unviable.) This type of model was originally developed by the group of John Reinitz, of Stony Brook University, New York, and, despite its level of abstraction, it has described early *Drosophila* development well enough to predict mutant phenotypes.

Wagner and his colleagues then asked a question of their model: how many genotypes can generate the same phenotype? This is equivalent to asking how many different species' developmental programs can generate a perfect insect body plan. They explored this by allowing a given genotype to evolve by small steps (modifying one gene interaction at a time), but keeping only the genotypes that produced the same phenotype.

By repeating such 'neutral' evolutionary steps hundreds of times they could measure the 'neutral space' of genotypes that all produce the same phenotype — itself a fraction of the total space of all possible genotypes. Even with a relatively small number of genes, says Wagner, "the number of genotypes with any given phenotype is astronomically large". Not only large, but spread out; the team found that it was possible to get from their starting genotype to something completely different — to move from one end of that imaginary space to the other — while keeping the corresponding phenotype unchanged.

If the real world works like the model, regulatory networks can undergo dramatic reshuffles with no outward sign of what is going on at all. And Wagner's studies hint that being able to build the same edifice on such different — and changing — foundations may confer an evolutionary advantage. The team altered its algorithm to look at novel phenotypes that are just one mutation away from a starting genotype. So imagine genotypes in two far-flung parts of that imaginary neutral space, creating the same phenotype. Make a small mutation to one of those and it will produce a second phenotype; make a small change to the other and it will produce something very different. Wagner and his colleagues were not trying to get their model to innovate; but because the genotypes originally explored such a range of the possibilities without changing phenotype, they could later mutate in a very wide range of directions. So it may be that invisible variation in networks better prepares organisms for new challenges to come, by arming them with more possible solutions.

The real deal

This year, a group led by Mark Isalan, of the Center for Genomic Regulation in Barcelona, Spain, rewired a real genetic network[6]. Isalan picked out 22 DNA-binding regulatory proteins in *Escherichia coli* and then created close to 600 genetic sequences in which the DNA coding for a given protein was put under the regulatory control of another. They then introduced each one of these synthetic genes into normal bacteria. In essence, they added a handful of new links to *E. coli*'s endogenous regulatory network.

The group found that roughly 95% of the rewired bacteria survived: the bacterial regulatory network seems remarkably robust to arbitrary rewiring. And some of these engineered bacteria could out-compete the original bacteria in their ability to survive culture conditions such as 'heat-shock'. If genetic systems were like the computer code that humans build, this would be like dumping arbitrary 'Go-To' instructions in the middle of some air-traffic-control software and finding that not only did it cause very few crashes, but it also got more planes to their destination on time. The researchers concluded that, for *E. coli* at least, new links in regulatory networks are rarely a problem, and that they can even confer improved fitness in some environments.

But some biologists question the evolutionary significance of lab experiments, in which organisms are subject to less challenging conditions than in the wild. Systematic gene deletion in yeast, for example, has similarly shown that close to 80% of all protein-coding genes are not essential in normal lab culture. And yet their persistence as functional sequences in the genome is a strong indication that they contribute to the organism's fitness in the wild. So even if computer simulations or lab experiments suggest that network rewiring can be innocuous or even advantageous, the situation in real life might be different.

Come back to the yeast mating type: although the various genetic logics of *S. cerevisiae*, *C. albicans* and *K. lactis* seem to have identical outputs, in that they all generate two mating types, these may provide the various organisms with some unknown adaptive features that are of benefit in their different ecological niches in humans, breweries or cheese factories. "We don't know what hidden

advantages if any might be associated with the various mating-type circuits found in those yeast species," Johnson says.

Sea-urchin development is one case in which alternative strategies for building the same adult body plan seem to have offered advantages in lifestyle. The two species in the genus *Heliocidaris*, for example, still share the shallow waters of southeastern Australia; their last common ancestor lived around the same time as the last common ancestor of chimps and humans. But they have dramatically different means of development. The *H. tuberculata* embryo develops first into a swimming and feeding larva called the 'pluteus', a form of indirect development that is shared by most of the 1,000 known species of sea urchins and probably represents the primitive mode for the clade. But *H. erythrogramma* develops direct from egg to adult in just a few days. Its entire developmental program has been completely scrambled to bring this about: its egg is 100 times bigger and the fates of the various cells that the egg divides into are completely different to those seen in *H. tuberculata*[7]. Even Wolpert's sacrosanct gastrulation is rearranged. It is as if, in a brief 4 million years, chimps had evolved to lay hard-shelled eggs while still developing into chimp-like creatures in the end.

Solving multiple problems

The sea urchins' developmental modes may represent equivalent solutions to one problem — how to build a spiny adult body that sticks to a rock — and yet be non-equivalent solutions to another — how to survive a youth in open water. Sea urchins with swimming and feeding larvae produce many small eggs, have broader geographic distributions, and show high mortality rates due to predators. Sea urchins that develop direct and fast produce fewer, bigger eggs and tend to stay put in remote niches; they have higher survival rates per embryo. In this case, changes that are 'neutral' with respect to one aspect of an organism's phenotype seem to present advantages in others.

Many are unconvinced by these evo-devo studies. The common view that "the guiding hand of natural selection" is behind the evolution of all biological sophistication is greatly misleading, says evolutionary biologist Michael Lynch of Indiana University in Bloomington. Lynch thinks that most biologists underestimate the influence of chance, particularly in small populations in which neutral genetic changes can become the norm without any selective pressure by the process of 'genetic drift'. "Nothing in evolution makes sense except in the light of population genetics," Lynch says.

This is particularly true for gene regulatory networks, which can evolve by chance much faster than protein-coding genes because they can tolerate changes by point mutation, duplication or deletion with little effect on overall function. New binding sites for regulatory proteins, for example, can appear or disappear and the change can quickly become fixed in a population without loss of fitness[8]. Such constant reshuffling makes regulatory networks more complex than they need to be. It also creates redundancy — several parts of the network can perform the same task — and this allows further reshuffling to eventually bring about dramatic change in the way that things are run. Lynch thinks that Johnson's yeast studies are a particularly ele gant demonstration of this.

> **"Nothing in evolution makes sense except in the light of population genetics."**
> — Michael Lynch

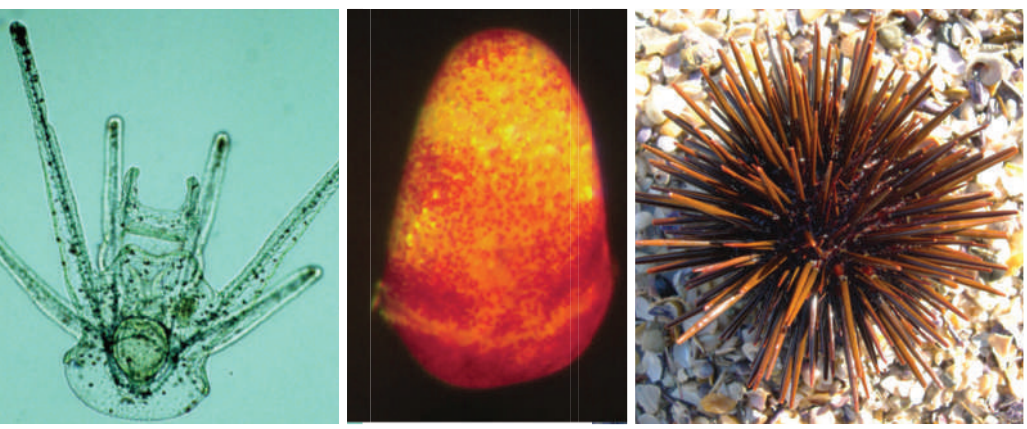
Such "seemingly baroque architecture" of biological networks, as Lynch puts it, leaves developmental biologists with a problem: how to understand whole systems while relying on single-gene experiments. Traditionally, they have knocked out a gene in a model organism and drawn conclusions from that about the gene's function in a range of species in which that gene's sequence is conserved. The evolutionary biologist Stephen Jay Gould long challenged this type of reductionist method, in which biological function of a gene is discussed 'all-other-things-being-equal' — the *ceteris paribus* paradigm[9]. Systems evo-devo makes it clear that all other things are not equal: the function of any gene cannot be defined outside its species-specific context.

Developmental biologists have to find a new way to describe the commonalities between organisms that were once ascribed to genes. The fact that tunicates and vertebrates both develop through a tadpole-like stage suggests that there must still be shared biological 'rules' governing the process. But if genes and regulatory links aren't conserved, then what exactly do they have in common? The shared features may actually lie in some other property of the regulatory networks.

It is this problem that Lemaire and others now hope to address in tunicates. Because of the relatively simple body plan and the ease with which tunicate genes can be manipulated, researchers have come closer to building a complete regulatory network for tunicates than for any other model organism, one that shows how each regulatory gene interacts with every other, in every cell, and at every stage of development[10]. Lemaire is hoping to extract a 'network signature' that would characterize the process of gastrulation, perhaps a mathematical expression that describes the density of links between genes or the types of feedback loops used. Then researchers can try to see whether the same features are shared by gene regulatory networks in humans and other species that do rely on a Spemann organizer to gastrulate.

If Wolpert is right, and gastrulation is truly the most important event in your life, then that seems an answer worth knowing. ■

Tanguy Chouard is a senior editor at *Nature*.



E. C. RAFF AND R. A. RAFF

***Heliocidaris* sea urchins have very different larval stages (left and centre), yet reach a similar adult form.**

1. Lemaire, P., Smith, W. C. & Nishida, H. *Curr. Biol.* **18,** R620–R631 (2008).
2. Schröder, R. *Nature* **422,** 621–625 (2003).
3. Lynch, J. A., Brent, A. E., Leaf, D. S., Pultz, M. A. & Desplan C. *Nature* **439,** 728–732 (2006).
4. Tsong, A. E., Tuch, B. B., Li, H. & Johnson, A. D. *Nature* **443,** 415–420 (2006).
5. Ciliberti, S., Martin, O. C. & Wagner, A. *Proc. Natl Acad. Sci. USA* **104,** 13591–13596 (2007).
6. Isalan, M. *et al. Nature* **452,** 840–845 (2008).
7. Wray, G. A. & Raff, R. A. *Dev. Biol.* **132,** 458–470 (1989).
8. Lynch, M. *Nature Rev. Genet.* **8,** 803–813 (2007).
9. Gould, S. J. *The Structure of Evolutionary Theory* (Harvard University Press, 2003).
10. Imai, K. S., Levine, M., Satoh, N. & Satou, Y. *Science* **312,** 1183–1187 (2006).

See pages 281, and 295, and the Darwin special at www.nature.com/darwin.