vertebrate processes of development, such as formation of the nervous system and heart. All of this suggests that the sea squirt retains a primitively simple complement of genes that would have been found in the last common ancestor of sea squirts and vertebrates, more than 550 million years ago.

The converse is also true — *Ciona* has had plenty of time to evolve its own, non-vertebrate peculiarities. Sea squirts have some interesting physiological quirks, the most bizarre of which is the manufacture of cellulose for their distinctive outer coverings, or 'tunics'. Some of the enzymes involved in cellulose metabolism are undoubtedly home-grown, but the report[1] of a cellulose synthase gene, previously found in plants, is unique for animals and might have been an import from nitrogen-fixing bacteria.

Genomes of other non-vertebrate deuterostomes will be required to fill in the fine details of the order in which the various distinctively vertebrate developmental and physiological features were acquired. In this context, the genome of the lancelet *Branchiostoma* — a vaguely fish-like creature even more closely related to vertebrates than is *Ciona* — would be especially welcome.

What we have is a splendid start — or rather restart. This story really began in the 1880s with the frustration of one young researcher at the general inability to get to grips with the issue of vertebrate origins. With nothing to go on but basic embryology and comparative anatomy, schools of researchers devised a variety of explanations for vertebrate origins, not all of which were compatible, and with no way of deciding which ones were more likely to be correct.

This was the last straw for our hero. "Out of the same facts of anatomy and development men of equal ability and repute have brought the most opposite conclusions," he wrote. "To take for instance the question of the ancestry of the Chordata, the problem on which I was myself engaged, even if we neglect fanciful suggestions, there remain two wholly incompatible views as to the lines of Vertebrate descent, each well supported and upheld by many. From the same facts opposite conclusions are drawn. Facts of the same kind will take us no further. The issue turns not on the facts but on the assumptions."

He then proceeded to burn his boats in spectacular fashion: "Surely we can do better than this. Need we waste more effort in these vain and sophistical disputes?"

Six years after publishing these lines[2], the author coined a word for an entirely new discipline that promised to put hard facts in place of vacuous speculation. He was William Bateson, and the word was 'genetics'. Across the Atlantic, Thomas Hunt Morgan was also leaving old-fashioned evolutionary biology for experimental science, abandoning various worms for the fruitfly, on whose narrow and bristly shoulders rests a century of work on comparative developmental biology. The work on the genome of *Ciona* can thus trace its lineage to Bateson and Morgan and the very foundations of genetics.

The origin of vertebrates may seem like a question straight out of some dusty zoology textbook. But we owe the existence of genetics, genomics and, arguably, a large part of molecular biology to the frustration of two young men unable to coax the secrets of evolution from creatures such as *Ciona intestinalis*. Welcome back, little squirt. ■

Henry Gee is a senior editor at Nature.

LAWSON WOOD/CORBIS

Figure 1 **Object of taxonomic affection — *Ciona intestinalis*, the tube-like organisms seen in this underwater display.**

1. Dehal, P. *et al. Science* **298,** 2157–2167 (2002).
2. Bateson, W. *Materials for the Study of Variation* (Macmillan, London, 1894).

Computational biology

Evolution plays dice

Philip Gerrish

In computer simulations, initially identical populations of organisms growing in identical environments follow very different evolutionary trajectories. Mutational interdependence is a key factor.

Because of the long timescales involved, evolution has principally been a historical or retrospective science, relying in large part on a sparse fossil record. This constraint makes drawing inferences about evolutionary processes very difficult, but it does not apply to systems where evolution happens fast. With such systems, evolution ceases to be a strictly historical science and becomes amenable to experimentation, allowing hypotheses about evolution to be tested directly. Notably, fast evolution has been studied in laboratory populations of microbes[1,2], and really fast evolution has been studied in *in silico* populations of virtual organisms[3]. Virtual organisms are self-replicating, mutable computer programs, and as Yedid and Bell show on page 810 of this issue[4], they provide a powerful way of addressing fundamental questions about evolution.

One such question, remarkably, continues to confound evolutionary biologists. Is evolution governed mainly by chance or is its outcome largely predetermined? With microbial or *in silico* systems, this question can be addressed experimentally as follows. Create several parallel populations with identical organisms; let them evolve independently in identical environments for a period of time; then see whether the evolved organisms are the same or not. If evolution is governed solely by chance, each population will follow a unique evolutionary trajectory and the evolved organisms will all be different. If evolution is completely deterministic, each population will follow the same trajectory and the evolved organisms will be the same.

This simple experiment has been performed with different microbes[5–7]. Yet even with today's sophisticated biotechnological

tools, the level of observable detail is limited, restricting the inferences that can be made about the underlying processes. To look at those processes in more detail, Yedid and Bell[4] performed the same experiment using populations of virtual organisms. Simply put, they conclude that chance rules.

This conclusion may seem obvious. After all, the mutations that drive evolutionary change occur at random, and any simulation that pretends to mimic them must do so using a random number generator. So, it seems, the accumulation of mutations over time must be a completely random process, whether *in vivo*, *in vitro* or *in silico*. The flaw in this logic is that it fails to acknowledge that a largely deterministic process — natural selection — is also at work in evolution.

Natural selection sorts out all the mutants in a population, sending the chaff to extinction and keeping only the best genotype — that is, the genetic constitution that confers the highest reproduction rate. In an unrealistically large (effectively infinite) population, all possible mutations are sampled, and so the best genotype will be predetermined by natural selection. In this extreme example, the outcome of evolution would be the same across replicate populations and one would expect the opposite of what Yedid and Bell found. But populations of realistic size are too small to sample all possible mutations (including double mutations, triple mutations, and so on). So the best genotype in one population will be different from the best genotype in another population. These two prizewinning genotypes may have similar reproduction rates, but their genetic configurations will probably be different. Indeed, Yedid and Bell found that their virtual bugs evolved similar biological properties (such as reproduction rate and genome size) across replicate populations, despite considerable divergence at the genetic level — a result that is corroborated by data from work with the bacterium *Escherichia coli*[6,7].

Evolutionary timescales are notoriously long, however. So it would seem that there ought to be plenty of time for each population to eventually acquire the best possible genetic configuration — which, again, should be the same across replicate populations. The reason this doesn't happen is that evolution occurs in 'steps', especially in asexual populations as studied by Yedid and Bell. Such steps happen when the best mutant available at the time becomes the predominant genotype in the population through the action of natural selection. At this point, the population is committed to carrying this particular mutation, even though it is probably not the best mutation the population could have acquired. That's because natural selection will not allow the population to go back to an inferior genotype. As if this commitment were not enough, the particular winning mutation largely determines which

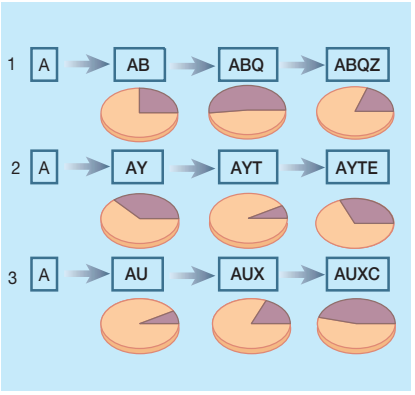

Figure 1 **The population-divergence ratchet. All populations start with an identical genotype, A. Population 1 acquires mutation B, creating genotype AB. B is not the best possible mutation, but rather the best mutation then available. The pie chart indicates the increment in reproduction rate conferred by mutation B as a fraction (in this case, 75%) of the rate conferred by the best possible mutation. In population 3, mutation U confers 90% of the maximum possible increment. Because of epistasis, future evolutionary trajectories are restricted by previous mutations. For example, the B mutation is advantageous on the A background, as evidenced by its acquisition in population 1. But B may not be advantageous on the AU background, the second genotype in population 3. Instead, a different set of advantageous mutations is available to genotype AU. One of them is X, which is acquired, giving genotype AUX. From a new set of available mutations, C is then acquired. The same principles apply to population 2 and, overall, the ratchet produces ever-increasing population divergence, as observed in computer simulations by Yedid and Bell[4].**

of the subsequent mutations will further improve replication rate, one of which will cause the next evolutionary step. By carrying this particular winning mutation, therefore, the population is further committed to a new and limited set of future trajectories.

Such contingency on previous evolutionary steps results from the fact that the effects of mutations are interdependent. This phenomenon, known as epistasis, is often politely acknowledged to be quite prevalent in biological systems but is then promptly swept under the rug for the theoretical complications it causes. The results of Yedid and Bell's *in silico* 'go-back' experiments, however, suggest that ignoring epistasis in long-term evolution is a mistake. They find that interdependence among mutational effects — both physiological and social — plays a key role in the dynamics of population divergence. This idea is actually quite old[8], but such strong experimental evidence supporting it is new and may be the most consequential contribution of Yedid and Bell's paper.

All in all, then, three factors together produce a temporal asymmetry and thus a ratchet-like process that drives genetic divergence: the interdependence among mutational effects; the randomness associated with the best genotype available in a finite population; and the unidirectional tendency of natural selection to increase replication rate (Fig. 1). Not only did Yedid and Bell find considerable genetic divergence in their *in silico* populations, supporting the view that chance was the governing factor, but also the detail afforded by their experimental system allowed them to identify why. Populations diverge genetically because the role of chance is amplified by epistatic contingencies.

How much weight can we put on these conclusions? True, they do not come from real biological populations. But nor do they come from armchair speculation, a practice that has bedevilled evolutionary biology. Such experiments with virtual organisms are a step away from the armchair and towards experimental biology. But there is some way to go before virtual populations can be comparable to real biological populations. One potential difference is in population size. The largest virtual population size achieved by Yedid and Bell was 10^4, whereas real populations of viruses, for example, can achieve sizes of the order of 10^{12}. Although the mutation supply rate can be adjusted in virtual populations to compensate for smaller population sizes (as Yedid and Bell did), the resulting population dynamics are not necessarily comparable to the dynamics of very large populations. As pointed out above, the degree of predictability in evolution should increase with increasing population size — a trend observed by Yedid and Bell and corroborated by data showing predictability in the evolution of HIV, where populations are typically very large[9].

A second difference between real populations and virtual populations is that there is some degree of genetic exchange in most real populations, whereas Yedid and Bell's virtual populations were strictly asexual. Introducing genetic exchange may increase or decrease the degree of predictability in evolution — a question that could nicely be addressed with populations of virtual organisms. ■

Philip Gerrish is in the Programa de Investigación en Matemáticas Aplicadas y Computación, Instituto Mexicano del Petróleo, Col. San Bartolo Atepehuacán, México DF 07730, Mexico.
e-mail: pgerrish@yahoo.com

1. Wahl, L. M. & Krakauer, D. C. *Genetics* **156,** 1437–1448 (2000).
2. Appenzeller, T. *Science* **284,** 2108–2110 (1999).
3. Zimmer, C. *Nat. Hist.* 42–45 (May 2001).
4. Yedid, G. & Bell, G. *Nature* **420,** 810–812 (2002).
5. Wichman, H. A., Badgett, M. R., Scott, L. A., Boulianne, C. M. & Bull, J. J. *Science* **285,** 422–424 (1999).
6. Papadopoulos, D. *et al. Proc. Natl Acad. Sci. USA* **96,** 3807–3812 (1999).
7. Lenski, R. E. & Travisano, M. *Proc. Natl Acad. Sci. USA* **91,** 6808–6814 (1994).
8. Wright, S. *Genetics* **16,** 97–159 (1931).
9. Boucher, C. A. *et al. J. Infect. Dis.* **165,** 105–110 (1992).