many constraints, and a local search algorithm proceeds by trying to 'fix' variable settings to reduce the number of violations in the search for a variable setting that satisfies all constraints.

In tackling the hitting-set problem, Mézard and Tarzia[3] follow a fundamentally different route. They take advantage of a significant advance that occurred in the early 1990s, when computer scientists banded together with physicists to study ensembles of randomly generated instances of various NP-complete problems[5–8]. An 'instance' here is simply a particular example of the generic problem, defined by a set of variables, and specific governing constraints; in our previous example, one specifies a set of students and the sports they play.

This work revealed that, at certain values of the ratio of constraints to variables, ensembles of random instances of the same generic problem underwent a sudden change, dubbed a phase transition. Below the phase-transition point, most instances have one or more solutions that satisfy all constraints; above the phase transition, most instances do not have any solution, because there are too many constraints to satisfy. The instances that were hardest to solve occurred with numbers of variables and constraints that lay exactly at these phase-transition boundaries. A natural conclusion was that tools from statistical physics developed to study physical phase transitions might help in developing more efficient algorithms for solving combinatorial problems.

An example is the 'survey-propagation method'[9] used by Mézard and Tarzia[3], which developed from the cavity method used in statistical physics to calculate ground-state properties of certain condensed-matter systems. Survey propagation solves random instances of the boolean satisfiability problem near phase transitions with large numbers of variables (more than 10^7), which are beyond the reach of backtrack and local-search methods. This archetypal NP-complete problem asks the question of whether, given a set of logical statements using boolean variables (variables that can be either 'true' or 'false'), there is any assignment of values to those variables that can satisfy all the statements.

Mézard and Tarzia use the survey-propagation method to compute statistical properties of the solutions of instances of the hitting-set problem. Such a strategy might seem doomed to fail because it is generally significantly harder to determine the properties of the set of solutions of a hard computational problem than it is to find a single solution. But survey propagation can efficiently approximate the requisite statistical information for instances of various combinatorial problems near phase boundaries. It does this by iteratively solving a large set of coupled equations, modelling the local interactions between variables probabilistically. This solution process can be performed in a parallel, distributed fashion using many different processors, and generally converges

to an answer extremely quickly — in seconds for equations with thousands of variables.

Survey propagation can be viewed as a generalization of the 'belief-propagation method'[10], which was discovered independently in several fields, including information theory and artificial intelligence. Belief propagation is a way of approximating the probability (the 'belief') that a variable takes on a particular value in a randomly sampled solution. This information can be used to set variables incrementally, thus simplifying a problem.

The method works well when solutions are nicely clustered together in the combinatorial space, which is the case reasonably far from a phase transition. Near phase boundaries, however, solutions break up into many smaller, unconnected clusters in the combinatorial space[11]. Conventional combinatorial search algorithms and standard belief-propagation techniques become trapped between these clusters, and cannot effectively search the solution space. The survey-propagation method, on the other hand, continues to provide reliable statistical information about the solution space[12–14].

It is this property that allows Mézard and Tarzia[3] to map out for the first time the space of hitting-set problems, identifying under what conditions belief- and survey-propagation methods can solve hard, random instances of the problem. They also identify regions where still more complex survey-propagation-style equations would be required.

Given the ever increasing role of computational methods in other disciplines, the fact that those other disciplines are, in turn, starting to contribute new concepts and ideas to the science of computation is an exciting development — one that, as the demands we make on computational methods continue to grow, we are sure to hear more of. ∎

Bart Selman is in the Department of Computer Science, Cornell University, 4148 Upson Hall, Ithaca, New York 14853, USA.
e-mail: selman@cs.cornell.edu

1. www.darpa.mil/grandchallenge/
2. Cook, S. *Proc. 3rd Annu. ACM Symp. Theor. Comput.* 151–158 (ACM, New York, 1971).
3. Mézard, M. & Tarzia, M. *Phys. Rev. E* **76,** 041124 (2007).
4. Gomes, C. P. & Selman, B. *Nature* **435,** 751–752 (2005).
5. Cheeseman, P., Kanefsky, B. & Taylor, W. *Proc. 12th Int. Joint Conf. Artif. Intell.* 331–337 (Morgan Kaufmann, San Francisco, 1991).
6. Mitchell, D., Selman, B. & Levesque, H. *Proc. 10th Nat. Conf. on Artif. Intell.* 459–465 (AAAI, Menlo Park, CA, 1992).
7. Kirkpatrick, S. & Selman, B. *Science* **264,** 1297–1301 (1994).
8. Monasson, R., Zecchina, R., Kirkpatrick, S., Selman, B. & Troyansky, L. *Nature* **400,** 133–137 (1999).
9. Mézard, M., Parisi, G. & Zecchina, R. *Science* **297,** 812–815 (2002).
10. Pearl, J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference* (Morgan Kaufmann, San Francisco, CA, 1988).
11. Mézard, M., Mora, T. & Zecchina, R. *Phys. Rev. Lett.* **94,** 197205 (2005).
12. Maneva, E., Mossel, E. & Wainright, M. J. *J. Assoc. Comput. Machin.* **54** (4), 2–41 (2007).
13. Braunstein, A. & Zecchina, R. *J. Stat. Mech.* P06007 (2004).
14. Kroc, L., Sabharwal, A. & Selman, B. *Proc. 23rd Conf. Uncert. Artif. Intell.* 217–226 (AUAI Press, Corvallis, OR, 2007).

EVOLUTIONARY GENETICS

Who shouldn't be your daddy

Patrick C. Phillips

Unusual reproductive incompatibility has been discovered between two strains of a nematode worm. This finding indicates that natural selection can generate long-term divergence within self-fertilizing populations.

Reproductive incompatibility is the stuff of speciation and lies at the heart of the world's tremendous biodiversity. It makes little sense for such incompatibility to be maintained within a species, however. After all, what is the advantage of having genes that kill your offspring? As they report in *Science*, Seidel and colleagues[1] attempt to answer this question by identifying two genes that mediate reproductive incompatibility between different populations of the nematode worm *Caenorhabditis elegans*.

The authors crossed two evolutionarily divergent worm lines, one from Bristol, UK, and the other from Hawaii, and analysed the genome of the resulting offspring for genetic markers — known DNA sequences that differed between the two strains. They noticed that, rather than the expected one-to-one ratio of Bristol

and Hawaii markers, nearly all of the markers on one specific region of chromosome I were of the Bristol type.

To obtain second-generation offspring from the Bristol–Hawaii hybrid, Seidel *et al.* crossed the hybrids with one another and found that a quarter of the resulting embryos died. Remarkably, nearly all of the dead embryos carried Hawaiian genetic markers in the same region of chromosome I.

To narrow this effect down, the authors again constructed hybrids between the two strains and then crossed them back to the Hawaiian strain (Fig. 1). The mating system of *C. elegans* is unusual in that this worm can occur as a male or as a self-fertilizing hermaphrodite. The authors found that when eggs of Hawaiian worms were fertilized by sperm from a Hawaii–Bristol hybrid (either a male
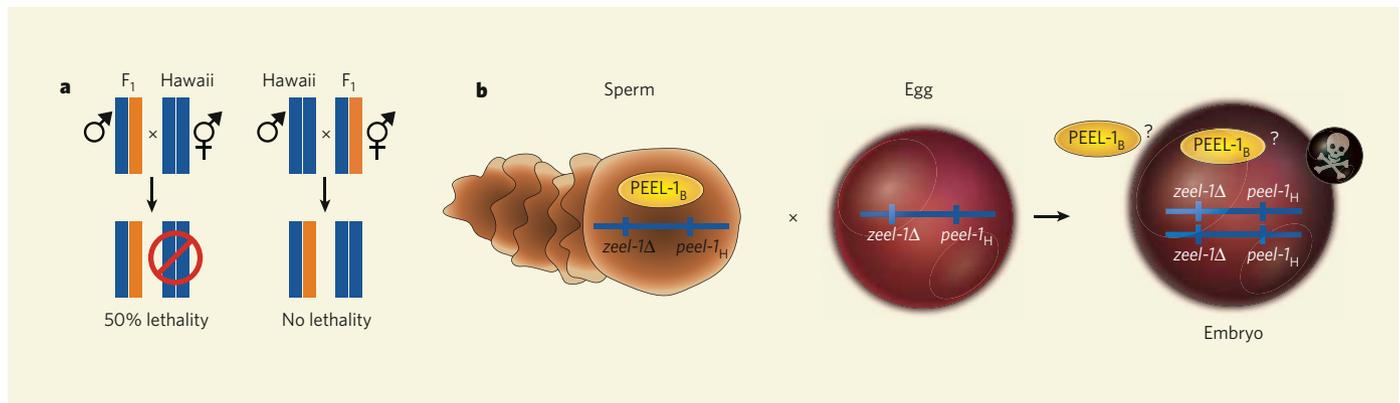
Figure 1 | Worms like to keep it in the family. a, Seidel *et al.*[1] crossed strains of *Caenorhabditis elegans* from Bristol and Hawaii. They then mated the hybrid male offspring of this cross with the Hawaiian strain. This led to a high proportion of embryonic lethality (50%) in the second-generation offspring. If, however, the sperm donor was a worm of the Hawaiian strain crossed with a Bristol–Hawaii hermaphrodite, no lethality was observed. **b**, Two genes are responsible for reproductive incompatibility between these divergent strains. The first, *peel-1,* is a paternally acting factor and the second is *zeel-1*, the expression of which is necessary in the embryo to counteract the lethal effects of Bristol *peel-1* product (PEEL-1_B). In the Hawaiian strain, *zeel-1* is missing entirely (*zeel-1Δ*). Here PEEL-1_B expressed by sperm precursor cells is transferred to the sperm containing the Hawaii copy of the incompatibility region (*zeel-1Δ*, *peel-1_H*). The lethal outcome is observed when an embryo with the *zeel-1Δ* locus is generated from a sperm containing PEEL-1_B. It is not known whether PEEL-1_B acts within or outside the embryo.

or a hermaphrodite), the embryos died. But if the Hawaii–Bristol offspring was the 'mother' (or egg donor) and a Hawaiian worm donated the sperm, the embryos survived (Fig. 1a). The authors therefore speculated that the first-generation Hawaii–Bristol offspring must carry a paternally acting gene in the Bristol-type genomic region of chromosome I — what they call the incompatibility region — that is responsible for embryonic lethality.

Post-fertilization effects of paternally expressed genes are quite rare because the sperm is thought to primarily transfer only its DNA to the developing embryo and few products originating from gene expression within the father (such as proteins or messenger RNA). Therefore, to test their hypothesis, Seidel *et al.* carried out further crossing of the inbred worm lines and narrowed down the incompatibility region to a 62-kilobase segment of nucleotides. Here they identified two new genes — *zeel-1*, which acts in the one-cell embryo, and *peel-1*, which is the paternal-effect factor from the sperm.

The locus (position) in which *zeel-1* occurs is missing in the Hawaiian strain. So the authors suggest that the incompatibility between Hawaii–Bristol hybrids and Hawaiian worms must be due to the action of the product of the paternal, Bristol *peel-1* in the absence of 'maternal' *zeel-1* (which would otherwise counteract the lethal effects of *peel-1*)(Fig. 1b).

Next, to gain insight into possible speciation in *C. elegans*, Seidel *et al.*[1] investigated the worldwide distribution of *zeel-1* and *peel-1* in natural isolates of this nematode. They found that both of these genes seem to be widely distributed, and can even occur in the same population. This is no way for an incompatibility system to behave! A possible explanation for this pattern could be recent, possibly human-assisted, broad-scale migration[2] of *C. elegans* populations. Consequently, variation in the geographical distribution of *zeel-1* alleles (copies) could be due to mixing between previously isolated strains. Seidel and colleagues, however, provide strong evidence against this view.

If this sort of mixing were to have occurred recently, we should see co-inheritance of large blocks of the genome in different strains. Instead, the authors find that, outside the incompatibility region, the Hawaiian and Bristol strains are genetically almost identical, and only within this region are there large differences in nucleotide sequence — differences that are more than 50-fold greater than usually found among *C. elegans* isolates. Seidel *et al.* take this high degree of polymorphism to mean that the *zeel-1* and *peel-1* loci are quite ancient and predate the divergence times of the surrounding genomic regions. This indicates that the incompatibility region has introgressed — infiltrated through repeated mating between different strains — into the genomes of many strains, and that this introgression has presumably been occurring on a worldwide scale for a long time.

What evolutionary forces might generate such a pattern? The authors offer two possibilities. First, because of the lethality induced by *peel-1* in worms not also carrying *zeel-1*, the Bristol *zeel-1* allele has a strong transmission advantage from the death of embryos with a Hawaiian background. The close proximity of *zeel-1* and *peel-1* loci is reminiscent of similar systems seen in the fruitfly *Drosophila* and other organisms[3] that allow, for instance, sperm of one type to prevent fertilization by sperm of another type. But alleles with this kind of advantage are thought to become quickly fixed within populations and not to result in the form of ancient polymorphism seen here.

Although it is possible that selection for the Bristol *zeel-1* allele is precisely counterbalanced by selection operating in the opposite direction on the Hawaiian allele, Seidel *et al.* dismiss this as highly unlikely. Instead, they favour the interpretation that polymorphism in these two genomic regions is maintained by balancing selection, which favours allele diversity, with the incompatibility that arises as a by-product of the long-term divergence generated by this selection. The most likely form of balancing selection in this case would be some kind of frequency-dependent selection in which rare variants are favoured, as is commonly observed in pathogen-resistance systems.

Why the incompatibility, then? This is where the unusual mating system of *C. elegans* (males and hermaphrodites) comes into play. Hermaphrodites can only self-fertilize, and so inter-crossing of animals of different strains occurs only through males, which are thought to be quite rare in many natural populations[4,5]. If most individuals are self-fertilizing, these divergent alleles can be maintained for a long time with little threat from induced lethality because the alleles would come into contact only infrequently.

Seidel and colleagues' work[1] is important because it is one of the few examples in which a specific incompatibility system has been identified at the molecular level. But it is perplexing because what one might have expected to be an elegant case of incipient speciation is something quite different. At the very least, this work highlights the emerging picture that genomic evolution in *C. elegans* is strongly dominated by its self-fertilizing mode of reproduction. ■

Patrick C. Phillips is in the Center for Ecology and Evolutionary Biology, University of Oregon, Eugene, Oregon 97403, USA.
e-mail: pphil@uoregon.edu

1. Seidel, H. S., Rockman, M. V. & Kruglyak, L. *Science* doi:10.1126/science.1151107 (2008).
2. Phillips, P. C. *Trends Genet.* **22,** 405–407 (2006).
3. Lyttle, T. W. *Annu. Rev. Genet.* **25,** 511–557 (1991).
4. Barrière, A. & Félix, M.-A. *Curr. Biol.* **15,** 1176–1184 (2005).
5. Sivasundar, A. & Hey, J. *Curr. Biol.* **15,** 1598–1602 (2005).