

growth of subcutaneous 'pancreatic' tumours by about half, but growth was slowed even more if the treatment was begun at the same time that the cells were injected under the skin.

Cancers of the digestive system are often incurable, so the prospect of a new treatment is exciting. But much work remains to be done before we know whether cyclopamine, or any other inhibitors of Hh signalling, such as Hh-blocking antibodies, are effective therapies. Although some of the cell lines tested by Thayer *et al.* were highly susceptible to cyclopamine treatment, others were insensitive to the drug. So some forms of pancreatic cancer might not be linked to defective Hh signalling. Alternatively, some cancers might be caused by mutations in components of the pathway that function downstream of Smo, in which case Smo inhibitors such as cyclopamine would have no effect.

Several other components of the Hh signalling pathway, in addition to Ptc and Smo,

have been identified from studies of fruitfly development. Each molecule in the pathway has the potential to cause cancer if its activity is disrupted or increased. But, as Berman *et al.* and Thayer *et al.* show, revealing the signalling defects that promote tumour growth creates opportunities to devise rational therapies. ■

Matthew P. Scott is in the Departments of *Developmental Biology and Genetics, Howard Hughes Medical Institute, Stanford University School of Medicine, California 94305, USA.* e-mail: [scott@cmgm.stanford.edu](mailto:scott@cmgm.stanford.edu)

1. McMahon, A. P., Ingham, P. W. & Tablin, C. J. *Curr. Top. Dev. Biol.* **53**, 1–114 (2003).
2. Taipale, J. & Beachy, P. A. *Nature* **411**, 349–354 (2001).
3. Berman, D. M. *et al.* *Nature* **425**, 846–851 (2003).
4. Thayer, S. P. *et al.* *Nature* **425**, 851–856 (2003).
5. Hahn, H. *et al.* *Cell* **85**, 841–851 (1996).
6. Johnson, R. L. *et al.* *Science* **272**, 1668–1671 (1996).
7. Goodrich, L. V. *et al.* *Science* **277**, 1109–1113 (1997).
8. Hahn, H. *et al.* *Nature Med.* **4**, 619–622 (1998).
9. Xie, J. *et al.* *Nature* **391**, 90–92 (1998).
10. Taipale, J. *et al.* *Nature* **406**, 1005–1009 (2000).

Evolution

## Ending incongruence

Henry Gee

Recovering the true evolutionary history of any group of organisms has seemed impossible. The availability of large amounts of genomic data promises an era in which the uncertainties are better constrained.

The careful reader of this issue will come across a picture that, at first sight, has a startling message. Those impatient to see it should turn to the report from Rokas *et al.* (A. Rokas, B. L. Williams, N. King & S. B. Carroll *Nature* **425**, 798–804; 2003). The authors' Fig. 4, on page 801, is the object of interest. What, you might ask yourself, is so remarkable about a phylogeny — an evolutionary tree — of seven species of yeast of the genus *Saccharomyces*? Closer inspection, however, will show that the authors are making an unprecedented claim: that this is a fully resolved phylogeny with five internal branches in the tree, each of which has unequivocal support from all the data. For years biologists have tried to find methods to tease evolutionary history from obtuse data. This looks like the best attempt yet.

In principle, it should be possible to describe the phylogeny of a group of organisms as a nested set of bifurcating branches with the appearance of a tree, in which the branching order is a graphic expression of the distribution of features among the organisms concerned. The question that always arises, however, is how one can ever justify the choice of characters that govern branching at each node. Which subset of characters, out of hundreds of possibilities, best reflects the 'true' phylogeny of a group? Even if we can possibly make such decisions,

how should the information from characters be 'weighted'? When technology allowed examination of the sequences of genes, rather than features of anatomy or physiology, the feeling was that truth would be simpler to reach. Genes are a direct expression of inheritance, and not signals refracted through the distorting lens of an organism's physical or biochemical characteristics.

But genes are not immune to external influences: like any feature of anatomy, they have histories that can confuse as well as enlighten. The result has been several decades over which phylogenies of various groups of organism have been based on sequences from one or a few genes, the assumption being that the genes of choice stand reliably for the many thousands of others that remain unsampled. The worry was that no justification could exist for this assumption. But this concern was not often articulated, because little could be done to address the problem.

For many years, single genes were all there were, and people had to make the best of them. The result was endless argument, because a phylogeny supported by analysis of one gene would often be very different from that created using data from another. Retreating into statistical thicket, researchers had to employ more or less sophisticated criteria to decide which of many millions of possible trees was most likely to represent the true history of the group in which they were interested.

The availability of genome-sized quantities of sequence information offers the possibility of changing things. Rokas *et al.* used a database of orthologous genes from seven species of *Saccharomyces*, with a more distant relative, *Candida albicans*, as an 'outgroup'. Orthologues are genes in different species that sequence analysis shows have an ancestral gene in common, while the inclusion of an outgroup is common practice to provide an external standard in phylogenetic analysis. With the database information, Rokas *et al.* could ask just how many genes are needed to produce a reliable phylogeny. By the same token, they could address the issue of why single-gene phylogenies are almost always unreliable.

Rokas *et al.* started with 106 genes represented by orthologues in all their candidate species, and then computed phylogenies using each gene in turn. As had been expected, each gene supported its own particular branching order. In other words, the trees were 'incongruent'. Incongruence can be explained in many ways, but the bottom line — only evident when it is possible to compare large numbers of genes simultaneously — is that there are no identifiable parameters that can predict the performance of genes in any systematic way. But when the researchers combined all the data from these disputatious genes, a *pax genetica* emerged — a single tree that was statistically robust at all points. This is the authors' Fig. 4.

They next considered how much genetic information was needed to recover this phylogeny reliably. The results varied according to the method, and the minimum amount of data required to achieve a single, fully resolved tree will vary according to the nature of the problem being analysed. In the case of the yeasts, however, the 'true' phylogeny could be recovered with remarkably little information. Given that nucleotide sequences in genes do not evolve independently, it was often possible to achieve a better result with small numbers of nucleotides sampled from many genes, rather than whole, single genes.

But hindsight is a wonderful thing, and it would have been interesting to know how confident Rokas *et al.* would have been in their results with minimal data had they not already created a perfectly robust tree with what turned out to be a superabundance of data. This, then, will always be a source of unease. There may be no way to know, without question, how many data are necessary to create a perfectly resolved tree, or indeed if this tree is necessarily the 'true' tree. But evolutionary biologists, like scientists generally, can only ever deal in provisional solutions. What is certain is that the work of Rokas *et al.* has raised the game of phylogenetic reconstruction to a new level. ■

Henry Gee is a senior editor at Nature.