

the beads were moved apart to produce an unzipping or unfolding transition, and then moved back together to allow refolding.

By integrating the force over the change in length of the molecule through many cycles, Collin and colleagues¹ obtained separate statistical distributions for the work done on the molecule during the forward and reverse transitions. They used these to verify that the CFT holds. By finding the energy at which the two distributions crossed for a range of speeds of bead movement, the authors were able to evaluate the free energy of transition between the folded and unfolded state of an individual RNA hairpin. They also demonstrated the robustness of the CFT by using their analysis to quantify the equilibrium free energy accurately even when the beads were moved apart too fast for the system to respond — in other words, when the system was taken far from equilibrium.

Collin and colleagues' work clearly verifies that the CFT can be applied to small biomolecular systems. But what are its wider implications for single-molecule research? The method provides a unique approach to quantifying the free energy in two-state systems — even in those systems far from equilibrium, where the transition between the two states requires very different amounts of work depending on its direction (a phenomenon known as hysteresis). This is important, as it is often the drifting baseline of a measuring instrument that limits how slowly, and thus how close to equilibrium, an experiment can be performed.

The CFT method is unlikely, however, to be suitable for all molecular systems and probe techniques. Its success in Collin and colleagues' experiments is due in part to certain features of the RNA hairpins and junctions used that made force a relatively smooth function of displacement between the well-defined initial (folded) and final (unfolded) states. Specifically, the free-energy landscape between these states resembles a 'staircase' of small energetic steps of short duration spread over an easily observable distance of many nanometres. The nonlinear, elastic behaviour of the polymer linkers and unfolded RNA chain also ensured that forward and reverse responses could be easily aligned in the presence of instrument drift. These attributes, together with the exquisite sensitivity of the optical tweezers — in regard to both the force applied and the displacement caused — allowed accurate integration for calculating the work during both unfolding and refolding. It also allowed for the subtraction of energetic contributions from the optical tweezers and chemical linkers, a necessary step in precisely determining the free-energy difference of the molecular transition. Extending Collin and colleagues' technique to other biomolecular transitions is thus sure to present fresh challenges.

Even so, this exciting experiment has provided insight into the way that recently developed fluctuation theories can be applied

to transitions in biomolecular systems. Not only does physics teach us new ways to understand single-molecule measurements, but such measurements also feed back into theory, extending our knowledge of thermodynamics to systems that are small and out of equilibrium. Stir a cup of coffee with a small enough spoon, and the coffee might just stir you. ■

Wesley P. Wong and Evan Evans are in the Department of Biomedical Engineering, Boston

University, 44 Cummington Street, Boston, Massachusetts 02215, USA. Wesley P. Wong is also in the Department of Physics, Harvard University, 17 Oxford Street, Cambridge, Massachusetts 02138, USA.
e-mails: wesley@physics.harvard.edu; evanse@bu.edu

1. Collin, D. *et al.* *Nature* **437**, 231–234 (2005).
2. Crooks, G. E. *Phys. Rev. E* **60**, 2721–2726 (1999).
3. Jarzynski, C. *Phys. Rev. Lett.* **78**, 2690–2693 (1997).

EVOLUTIONARY GENETICS

Microarrays and species origins

Roger Butlin and Cally Roper

Whole-genome arrays have been used to reveal small islands of genetic differentiation in *Anopheles* mosquitoes. Analysis of these regions will identify genes involved in the initial stages of speciation.

What are the genetic modifications that underlie the first steps in the origin of new species? The use of microarrays to make whole-genome comparisons between populations that can interbreed, but only to a limited degree, provides a fresh approach to the question. As reported in *PLoS Biology*, Turner *et al.*¹ have applied this methodology to the malaria-carrying mosquito, *Anopheles gambiae*. They have identified three small regions, less than 1.5% of the genome, that contain critical genes for the initiation of speciation.

In the classic model of speciation, two populations diverge during a period of geographical separation and cannot exchange genes when the physical barrier is removed. In this case the two genomes accumulate differences uniformly. There is, however, often gene flow during the period of divergence, or after secondary contact if reproductive isolation is incomplete. In these cases, genetic differentiation becomes unevenly distributed across the genome as gene flow erodes differences between species, leaving islands of differentiated sequence only around the genes responsible for reproductive isolation². This pattern has been nicely illustrated by the sequencing of multiple genetic loci in closely related species pairs of the fruitfly *Drosophila*^{3,4}, and by surveying differentiation at many anonymous markers, for example in the winkle, *Littorina saxatilis*⁵.

Whole-genome microarrays are used primarily for studies of gene expression, but they can also be used to detect differences in DNA sequences. Genomic DNA is hybridized to the array and a weak signal indicates a mismatch between the individual's DNA and the probe sequence. This principle has been successfully applied to the genetic mapping of quantitative traits in a plant, *Arabidopsis*⁶. Now, Turner and colleagues have used it to scan the whole mosquito genome for genetic

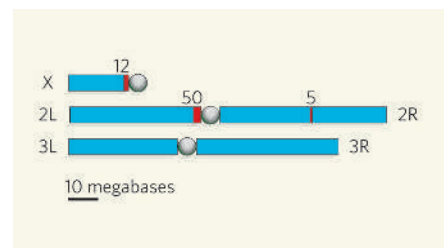


Figure 1 | Differentiated regions in the genomes of the M and S forms of *Anopheles gambiae*.

Anopheles gambiae has three chromosome pairs (X/Y, and 2 and 3, with the two arms of 2 and 3 being designated L and R). Numbers above the differentiated regions, shown in red, are counts of predicted genes within them: among these are genes critical for the initiation of speciation. The small differentiated region on 2R is least strongly supported statistically, perhaps simply because it is small. It may be significant that two of the differentiated regions are close to centromeres (circles).

variants that distinguish two forms of *Anopheles gambiae* known as M and S.

The M and S forms probably began to diverge very recently, perhaps within the past 10,000 years. The M form is restricted to West Africa, where it occurs alongside the S form, and the two interbreed at a low frequency (about 1.2%)⁷. The resulting hybrids are viable and fertile. Fixed differences unique to each species occur in the genes for ribosomal RNAs on the X chromosome⁸. But variable genetic markers elsewhere in the genome show little or no differentiation⁹.

Turner and colleagues' approach involved hybridizing DNA from seven individuals of each form, from Cameroon, to arrays of 142,065 probes each 25 base pairs long. Regions containing concentrations of differences between M and S were detected using a 300-probe sliding window, and just three small regions turned out to be significantly

differentiated (Fig. 1). On its own, this analysis may not have the power to distinguish regions protected from gene exchange from those that have large differences due to the chance effects of genetic drift. But sequencing of a sample of loci within the differentiated islands, close to these islands, or elsewhere in the genome, confirmed the result for a larger sample of individuals by detecting fixed nucleotide substitutions only in the candidate regions.

Among the small number of genes contained within these genomic islands of differentiation there must be critical genes for the first stages of speciation. The few currently known examples of such 'speciation genes', detected by other approaches¹⁰, are in much older species pairs; they may have evolved after the origin of reproductive isolation, rather than initiating it. So the genes identified in the M-S comparison will be especially interesting. They are likely to contribute to nonrandom mating, the major known barrier to gene flow, and possibly to differential environmental adaptation.

Even though this was a whole-genome survey, it would be premature to assume that all speciation genes fall within the identified islands. The small region detected on chromosome arm 2R (Fig. 1) was only marginally statistically significant. This is, at least partly, because it is small. The methods used would not be able to identify a single speciation gene in an area of high genetic recombination, such as *OdsH* in *Drosophila*⁴, because its signal in terms of the number of differentiated probes would be too weak to detect.

It may be significant that the two major differentiated regions are close to centromeres (Fig. 1), structures by which chromosomes are attached to the spindle during cell division. Centromeres are usually surrounded by areas of reduced recombination, and there is growing evidence that reduced recombination facilitates speciation by preventing the break-up of interacting combinations of variants at specific gene loci¹¹. Restricted recombination also aids the detection of differentiated regions. One of the exciting features of the mosquito system is that patterns such as this can be tested at different stages of speciation. There are seven recognized species in the Gambiae complex that are capable of breeding with each other to a greater or lesser degree, and the microarray approach can be extended to other pairwise comparisons within the group.

The methodology can also be applied to other such pairs of populations in species for which microarrays are available. These systems are rare at the moment, but they will rapidly become more numerous and will provide many insights into the origin of species. ■

Roger Butlin is in the Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK.

e-mail: r.k.butlin@sheffield.ac.uk

Cally Roper is in the Department of Infectious

and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK.

e-mail: cally.roper@lshtm.ac.uk

1. Turner, T. L., Hahn, M. W. & Nuzhdin, S. V. *PLoS Biol.* **3**, e285 (2005).
2. Wu, C.-I. *J. Evol. Biol.* **14**, 851-865 (2001).
3. Machado, C. A. & Hey, J. *Proc. R. Soc. Lond. B* **270**, 1193-1202 (2003).

4. Ting, C.-T., Tsaur, S.-C. & Wu, C.-I. *Proc. Natl Acad. Sci. USA* **97**, 5313-5316 (2000).
5. Wilding, C. S., Butlin, R. K. & Grahame, J. W. *J. Evol. Biol.* **14**, 611-619 (2001).
6. Borevitz, J. et al. *Genome Res.* **13**, 513-523 (2003).
7. Tripet, F., Touré, Y. T., Dolo, G. & Lanzaro, G. C. *Am. J. Trop. Med. Hyg.* **68**, 1-5 (2003).
8. Gentile, G. et al. *Insect Mol. Biol.* **10**, 25-32 (2001).
9. Lehmann, T. et al. *J. Hered.* **94**, 133-147 (2003).
10. Noor, M. A. F. *Nature* **423**, 699-700 (2003).
11. Butlin, R. K. *Mol. Ecol.* **14**, 2621-2635 (2005).

SYNTHETIC CHEMISTRY

Glycosylation with a twist

Sabine L. Flitsch

Nature has a whole battery of dedicated enzymes to make the complex links between sugar rings — how can synthetic chemists compete? An ingenious approach fills a big gap in the synthetic tool-kit.

Sugars can be linked together in a number of different ways to create chains. The resulting glycoside molecules have numerous biological functions, both by themselves and particularly when attached to proteins and lipids. The chemical reactions involved in making these molecules are formidable, and synthetic chemists have yet to find a set of robust chemical reactions that can generate the full complement of glycosidic linkages. In the *Journal of the American Chemical Society*, Boons and colleagues¹ describe a cunning method that could fill a major gap in the chemical tool-kit required for the efficient synthesis of many glycosides that occur in nature.

In glycosides, each sugar ring is linked through an oxygen atom to its neighbouring ring. The precise configuration of this glycosidic bond determines the molecule's physicochemical properties and biological function. The glycosidic bond occurs in two forms: the attached group can either be below (an α -bond) or above (a β -bond) the plane of the ring (Fig. 1). The chemical similarity of α - and β -bonds presents a problem for the synthesis of glycosides — nature solves this by using a dedicated enzyme (glycosyltransferase) to synthesize each different glycosidic bond. So it seems unlikely that one synthetic glycosylation method will be able to generate all glycosidic linkages. Some very robust reaction conditions have, however, been developed that can be used in different contexts to reliably produce broad classes of glycosides. The most successful of these uses a 'neighbouring group' method.

The classic version of this method uses a protective ester group attached to the carbon-2 position in the sugar ring to direct the formation of the new glycosidic bond to the neighbouring carbon-1 (Fig. 2a, overleaf). During the reaction the oxygen atom of the ester temporarily forms a bond to carbon-1 in the intermediate, termed an 'acetoxonium ion'.

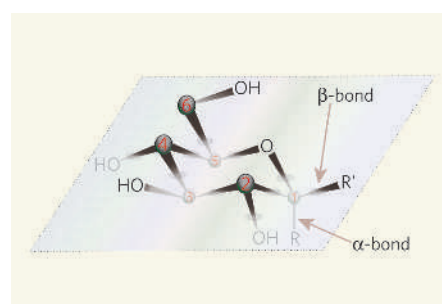


Figure 1 | Configurations of glycosidic bonds using glucose as an example. Carbon atoms are numbered around the ring as shown. R and R' represent different chemical constituents, such as another sugar ring or a hydrogen atom. The bond from carbon-1 to R is an α -bond as it forms below the plane of the ring; the bond from carbon-2 to R' is a β -bond as it forms above the plane of the ring. The R group is *cis* to the nearest -OH group as they are both on the same side of the ring (below the plane); the R' group is *trans* to the nearest -OH group as they are on opposite sides of the ring.

The neighbouring group blocks the lower face of the sugar ring so the glycosidic bond must form from above, making a β -bond. This method has been very useful for making glycosides in which the chemical groups attached to carbon-1 and carbon-2 are on opposite faces of the ring (that is, 1,2-*trans* glycosides). Recently, 1,2-*trans* glycosides have been synthesized by 'one-pot'^{2,3} and even automated routes⁴.

No such robust method is available to synthesize the many glycosides that contain 1,2-*cis* linkages (where the chemical groups attached to carbon-1 and carbon-2 are on the same face of the ring, such as the α -glucosides, α -galactosides or β -mannosides). To address this problem, Boons and colleagues¹ designed a different neighbouring group to create the opposite configuration to that of the classic method (Fig. 2b).