

alignments computed by other means. However, work remains before the method explored here can be considered totally reliable. For instance, inversions not associated with a duplication event currently cause problems for our approach to identifying orthologs. Another hurdle is presented by gene-conversion events¹⁸, which can present the appearance of an alignment with zones of differing percent identity. We anticipate that promising approaches will be developed by several groups, and that before long biologist will have free access to automatically generated alignments in gene clusters of higher quality than what is currently available.

References

1. O. Elemento, O. Gascuel and M. P. Lefranc, *Mol. Biol. Evol.* **19**, 278–288 (2002).
2. M. Lajoie, D. Bertrand, N. El-Mabrouk and O. Gascuel, *J. Comput. Biol.* **14**, 462–468 (2007).
3. Z. Jiang, H. Tang, M. Ventura, M. F. Cardone, T. Marques-Bonet, X. She, P. A. Pevzner, E. E. Eichler, *Nat. Genet.* **39**, 1361–1368 (2007).
4. J. Ma, A. Ratan, B. J. Raney, B. B. Suh, L. X. Zhang, W. Miller and D. Haussler. *J. Comput. Biol.* To appear (2008).
5. J. Ma, A. Ratan, B. J. Raney, B. B. Suh, W. Miller and D. Haussler. *Proc. Natl. Acad. Sci. USA*. To appear (2008).
6. Y. Zhang, G. Song, T. Vinar, E. D. Green, A. Siepel and W. Miller, *Proceedings of RECOMB* (2008).
7. S. Schwartz, W. J. Kent, A. Smit, Z. Zhang, R. Baertsch, R. C. Hardison, D. Haussler and W. Miller, *Genome Res.* **13**, 103–107 (2003).
8. J. H. Nadeau and B. A. Taylor, *Proc. Natl. Acad. Sci. USA* **81**, 814–818 (1984).
9. P. Pevzner and G. Tesler, *Proc. Natl. Acad. Sci. USA* **100**, 7672–7677 (2003).
10. J. S. Liu, *Monte Carlo Strategies in Scientific Computing*, Springer, New York (2001).
11. H. Wang, K. M. Donley, D. S. Keeney and S. M. G. Hoffman, *Environmental Health Perspectives* **111**, 1835–1842 (2002).
12. S. Ohno, *Evolution by Gene Duplication*, Springer, Berlin (1970).
13. A. Force, M. Lynch, F. B. Pickett, A. Amores, Y. Yan and J. Postlethwait. *Genetics* **151**, 1531–1545 (1999).
14. E. D. Green, *Nat. Rev. Genet.* **2**, 573–573 (2001).
15. E. Birney, M. Clamp and R. Durbin. GeneWise and Genomewise. *Genome Res.* **14**, 988–995 (2004).
16. Rhesus Macaque Genome Sequencing and Analysis Consortium. *Science* **316**, 222–234 (2007).
17. B. Hurler, W. Swanson, NISC Comparative Sequencing Program and E. D. Green. *Genome Res.* **17**, 276–286 (2007).
18. J.-M. Chen D. N. Cooper, N. Chuzhanova, C. Ferec and G. P. Patrinos. *Nature Reviews Genetics* **8**, 762–775 (2007).