as non-specific hybridization noise, which is sometimes observed in CGH experiments involving this region of chromosome 19. However, Hemminki and co-workers selected a polymporhic marker at 19p (*D19S886*) that corresponded to the deleted region, and demonstrated allelic loss in the polyps. Using a selective ultraviolet radiation fractionation technique, they were able to restrict their allelic loss studies to the epithelial cells of the polyps, the likely precursors of the malignancy. The lost allele seemed to originate from the unaffected parent's side: a finding that one would expect if a germ-line mutation of a tumour suppressor gene was present at this locus and segregating in the family (Fig. 1). To prove that this locus harbours the gene causing the Peutz-Jeghers syndrome, the researchers performed a linkage study in 12 Peutz-Jeghers families using *D19S886* as the genetic marker. Unambiguous linkage was obtained and by adding more markers, a multipoint lod score of 7.0 was achieved with no evidence of heterogeneity. This first marker that the authors tested was therefore critical in achieving locus assignment for the Peutz-Jeghers syndrome.

This is by no means the first example where chromosomal clues have proved critical for assigning a locus for cancer predisposition. Back in 1986, cytogenetic and molecular evidence of 13q14 deletions in tumours from patients with retinoblastoma provided the critical lead to the identification of the disease locus and the subsequent cloning of the *RB1* gene[7]. The availability of modern molecular cytogenetic techniques raises the question of whether Knudson's hypothesis and the search for somatic clues to cancer predisposition could also be helpful in defining loci for other inherited cancers, especially the hereditary forms of common solid tumours. Most of the cloned genes predisposing to cancer are tumour suppressor genes, whose wild-type allele is deleted as a result of a somatic loss[7]; however, finding such informative somatic deletions is not likely to be trivial. Both Peutz-Jeghers and retinoblastomas are rare diseases whose phenotypes are quite easy to distinguish, whereas most of the hereditary forms of common cancers have no phenotypic features that distinguish them from sporadic cases. Additionally, a large number of different predisposing genes may be involved. Distinction of early genetic changes is hampered by the genomic instability that typically arises at some point during tumour progression. For example, both *BRCA1* and *BRCA2* loci appear to undergo frequent allelic loss during tumour progression of non-familial cancers[9]. This would have made it quite difficult to find such predisposition loci based on CGH analyses of breast cancers from *BRCA1* and *BRCA2* carriers. Thus, even though it is technically straightforward to perform CGH analyses from familial cancers, secondary genetic changes may make it difficult to distinguish the predisposing loci. Careful statistical comparisons between matched sporadic and hereditary cases of the same histological type and clinical features could turn out to be useful. Alternatively, one could follow the Hemminki example and microdissect a preneoplastic early lesion, where such a critical deletion is more likely to appear as a sole genetic change. Overall, researchers should keep in mind that it is the second hit which makes cancers grow. □

1. Hemminki, A. *et al. Nature Genet.* **15**, 87–90 (1997).
2. Ginns E. *et al. Nature Genet.*, **12**, 431–435 (1996).
3. Spigelman A.D., Murday V., Phillips R.K. *Cancer*, **30**, 1588–1590 (1989).
4. Giardello F.M. *et al. N. Engl. J. Med.* **316**, 1511–1514 (1987).
5. Hall J.M., LeDuc C. A., Watson A. R., Roter A.H. *Genome Res.* **6**, 781–790 (1996).
6. Smith J.M. *et al. Science* **274**, 1371–1373 (1996).
7. Knudson A.G. *J. Cancer Res. Clin. Oncol.* **122**, 135–140 (1996).
8. Kallioniemi O-P. *Trends Genet.* **12**, 237–238 (1996).
9. Beckmann M.W. *et al. Br. J. Cance* **73**, 1220–1226 (1996).

Landmarks in the Rosetta Stone of mammalian comparative maps

Joseph H. Nadeau[1] & David Sankoff[2]

[1]*Genetics Department, Case Western Reserve University School of Medicine, Case Western Reserve University, 10900 Euclid Ave. Cleveland, Ohio 44016-4955, USA. e-mail jhn4@po.cwru.edu.* [2]*Centre de Recherches Mathematiques, Universite de Montreal, Montreal, Canada*

Since the divergence of the various mammalian lineages approximately 70 Myr ago, chromosome rearrangements have altered the original order of genes along many chromosomes and, in some cases, have exchanged segments among chromosomes[1]. The number of rearrangements, however, has been sufficiently modest that remnants of the ancestral chromosomes can be found in genetic maps by comparing the location of homologous genes from different species. Such analyses generate comparative maps which have many applications, including linkage predictions, disease gene identification, and studies of genome organization and evolution. Exploiting these comparative maps depends on the availability of detailed genetic maps for many mammalian species where homologous genes are numerous and readily identified. Homologous genes are crucial because they are the landmarks necessary to identify corresponding chromosomal locations and conserved segments in the genome of different species. Once complete genetic maps and complete genomic DNA sequences are available, map comparisons among mammalian species will be easily undertaken. In these early days of comparative genomics, however, the landmarks for interrelating maps are remarkably few.

While the principles of comparative mapping have been evident for many years, progress in defining landmarks in comparative maps has been unexpectedly slow. At present, at least 55 mammalian species are included in the comparative mapping portion of the Mouse Genome Database (http://www.informatics.jax.org)[2]. Tens to thousands of genes have been mapped in many species, but the number of genes that have been mapped in any pair of species is often very small[3]. A recent tabulation shows that 4,344 genes have been mapped in humans and 5,554 in mouse, but only 1,889 have been mapped in both (J. Ehrlich, D.S. and J.H.N., unpub. data). For other species pairs, the number of genes that have been mapped in both ranges from only one for deermice and dogs to 226 for humans and rats. Additionally, progress on comparative maps is too often the result of a fortuitous consequence of independent

gene mapping efforts in different species. For further progress it seems there are two options: wait for additional mapping and hope that new genes fill in the gaps, or address the problem systematically by developing reagents that can be used for comparative gene mapping. In this issue of *Nature Genetics*, Lyons *et al.*[4] have taken the proactive approach to this problem and report an important step towards the goal of developing reagents that define homologous landmarks in the genome of many mammalian species.

Essentially, the researchers identified 410 mapped genes for which DNA sequences were available in at least two species. Wherever possible primers were designed that spanned an intron to increase the chance of finding variation that can be used for gene mapping. The primers were also designed to include parts of exons so that homology could be verified by sequence analysis in different species. Of these 410 primer pairs, 318 (83%) produced a single PCR product with cat genomic DNA. Sequence analysis of PCR products in the cat verified amplification of the predicted gene in 59 (83%) of the 71 cases; however, gene families sometimes complicated the analysis. When these primer pairs were tested against genomic DNA from 20 other mammalian species representing a total of 11 orders, PCR products were obtained from 35% (voles) to 52% (deer) of the 318 primers. Finally, about 10% of the PCR products revealed size differences as assayed with agarose gel electrophoresis in mouse versus human or human versus Chinese hamster (the progenitor species of commonly used somatic cell hybrid panels). Obviously, PCR conditions could be modified for each species to improve detection of homology, and assay conditions could be improved to detect additional variation. In the spirit of STSs (sequence tagged sites) and ESTs (expressed sequence tags), the comparative mapping landmarks are called CATS (comparative anchor tagged sequences). As the authors point out, their work was successful both because of the large databases of genes and other conserved loci that have been mapped and sequenced in humans, mice and other mammalian species and because of the availability of DNA sequence analysis software.

By dramatically increasing the amount of comparative mapping information, CATS maps have important applications. The first is that they will enhance genetic analysis of disease and other traits in species that are not formally part of the Human Genome Project. Most genome structure reagents and information, such as genetic maps, gene and EST sequences and large insert DNA libraries, are much more complete for humans and mice than for other species, although considerable progress is being made in the other species[5]. Comparative maps greatly facilitate the transfer of these reagents and information from 'map rich' to 'map poor' species, thereby speeding progress towards disease gene discovery.

The other important application of CATS is for the analysis of genome organization and evolution. Many new analytical methods are being developed that depend on numerous high quality comparative maps[6–13]. Progress has been made on the limited data sets already available, but these analyses are based on limited information. Although many of these methods are remarkably robust with small sample sizes, concerns linger that sampling artefacts bias the results and interpretations. Additional comparative mapping information is essential. Examples of some of the applications of detailed comparative maps include reconstructing the ancestral mammalian genome ('original synteny')[14], estimating lineage-specific rates of chromosome rearrangements, comparing rates of chromosome rearrangements and nucleotide substitutions, and testing whether some segments remain intact or rearrange more often than expected.

Chromosome painting is an alternative method for identifying homologous chromosome segments. Here, chromosome- or segment-specific libraries from one species are fluorescently labelled and hybridized to chromosome preparations from another species[15–18]. The homologous segments are readily visualized. This is perhaps the fastest and easiest way to identify homologous chromosome segments and is the current method of choice for this purpose. However, painting, as presently implemented, usually lacks the resolution that is already available with comparative genetic maps, where homologies can be defined at the DNA sequence, the ultimate level of resolution. However, both FISH and CATS analysis techniques must be developed in parallel because their strengths are complementary.

Do we have enough CATS? Approximately 140–180 conserved linkages are estimated to exist in the human–mouse comparative map[3,5,6,19,20]. This corresponds to an average of about 2 CATS per conserved linkage. As measured by the proportion of the genome that is covered, the existing CATS identify the large segments that span most of the genome. However, the lengths of these segments follow a negative exponential distribution; there are many short segments and few long segments. Thus, many short segments are probably undetected and this represents an important reason to expand the number of CATS. Another reason is map precision: a high density of CATS will make the use of comparative maps more precise for linkage predictions and disease gene identification. Although more CATS and other similar kinds of markers[21] are needed, those described by Lyons *et al.* provide an important step towards a complete set of universal mapping reagents for translating mammalian genetic maps. □

1. Ohno, S., *Nature* **244**, 259–262 (1973).
2. Nadeau, J.H. *et al. Nature* **373**, 363–365 (1995).
3. O'Brien, S.J. *et al. Nature Genet.* **3**, 103–112 (1993).
4. Lyons, L.A. *et al. Nature Genet.* **15**, 47–56 (1997).
5. Comparative genome organization: first international workshop. *Mamm. Genome* **7**, 717–734 (1996).
6. Nadeau, J.H. & Taylor, B.A. *Proc. Natl. Acad. Sci. USA* **81**, 814–818 (1984).
7. Sankoff, D. *et al. Proc. Natl. Acad. Sci. USA* **89**, 6575–6579 (1992).
8. Zakharov, I.A., Nikiforov, V.I. & Stepanyuk, E.V. *Genetika* (trans. *Soviet Genet.*) **28**, 77–81 (1992).
9. Zakharov, I.A., Nikiforov, V.I. & Stepanyuk, E.V. *Genetika* (trans. *Soviet Genet.*) **31**, 1163–1167 (1995).
10. Bengtsson, B.O., Levan, K.K. & Levan, G. *Cytogenet. Cell Genet.* **64**, 198–200, (1993).
11. Hannenhalli, S., Chappey, C., Koonin, E.V. & Pevzner, P.A. *Genomics* **30**, 299–311 (1995).
12. Hannenhalli, S. & Pevzner, P.A. *Proc. 27th Ann. ACM–SIAM Symp. Found. Computer Sci.* 178–189 (1995).
13. Sankoff, D., & Nadeau, J.H. *Discrete Appl. Math.* (in the press).
14. Ferretti, V., Nadeau, J.H. & Sankoff, D. *Lecture Notes in Computer Science.* in *Combinatorial Pattern Matching* 1075 (eds Hirschberg, D. & Heyers, G.) 159–167 (Springer–Verlag, New York, 1996).
15. Jauch, A. *Proc. Natl. Acad. Sci. USA* **89**, 8611–8613 (1992).
16. Reid, T., Arnold, N., Ward, D.C. & Weinberg, J. *Genomics* **18**, 381–386 (1993).
17. Schertham, H. *et al. Nature Genet.* **6**, 342–347 (1994).
18. Rettenberger, G. *et al. Genomics* **26**, 372–378 (1995).
19. Copeland, N.G. *et al. Science* **262**, 57–66 (1993).
20. DeBry, R.W. & Seldin, M.F. *Genomics* **33**, 337–351 (1996).
21. Hino, O. *et al. Proc. Natn. Acad. Sci. USA* **90**, 730–734 (1993).