siblings; these families should be most genetically informative, and the statistical difficulties of treating such families should not be intractable.

Although the replication by Hamer's team is encouraging, the Xq28 hypothesis has also suffered a recent setback. In September, Ebers and Rice of the University of Western

> "Just because something is ethically complicated doesn't mean it shouldn't be investigated."
>
> — Dr. Bryan Sykes, Institute of Molecular Medicine, Oxford, July 1993 (ref. 12)

Ontario and collaborators presented a report at the twenty-first annual meeting of the International Academy of Sex Research that showed they found no hint of linkage between sexual orientation and Xq28 markers in 41 pairs of gay brothers. The Canadian study's design differed slightly from the earlier studies because the former did not exclude subjects from families with possible patrilineal inheritance. This lowers the Canadian study's effective sample size slightly, but it is unlikely that the different exclusion critieria can account for the different results. Even with plausible adjustments, its negative results differ reliably from combined results of the two studies from Hamer's lab.

The Canadian study did generate one finding consistent with an X-linkage hypothesis, however. In a replication of pedigree findings from Hamer *et al.'s* original report[1], gay men reported a significantly higher rate of gay uncles on their mother's side than on their father's side, 13.4% versus 6.9%, respectively. Contrary to previous concerns[7], this did not appear to result from probands having greater knowledge of their mothers' sides; homosexual aunts showed an opposite, patrilineal bias.

Reasons for the inconsistency of linkage results between the two labs are unclear, and in any case they are ultimately less important than the question of whether a gene within Xq28 affects male sexual orientation. That question could be answered definitively with a sufficiently large study, on the order of 100 pairs of gay brothers. Gay publications, potentially useful for recruiting families, are read widely, and many gay men are intensely interested in participating in research on biological origins. Thus, completing a large study would not be difficult. Indeed, both Elliot Gershon of the National Institute of Mental Health and Richard Pillard of Boston University have begun studies of this intended scope.

The potential payoff of a linkage finding is enormous, though not for reasons commonly believed. There are, for example, no clear social implications of a genetic explanation for sexual orientation[10]. If one is a determinist, then all behavior is caused, and in that fundamental sense, out of one's control. It is difficult to see why, for example, one is more responsible for behavior resulting from parental treatment than from one's genotype. In contrast, the discovery of a gene for sexual orientation, and subsequent elaboration of its function, would probably have far-reaching scientific

implications in the areas of sex differences, sexuality and neuroscience.

One final intriguing reason to entertain an X-linkage hypothesis for male homosexuality is evolutionary. Evidence for genetic contributions to sexual orientation has raised the vexing question of how genes for homosexuality could persist despite their obvious reproductive disadvantage. Biologist Robert Trivers, in his presentation at last year's annual meeting of the Human Behavior and Evolution Society, has offered an elegant solution— Assume that male homosexuality is caused by an uncommon X-linked gene that is evolutionarily advantageous in women. Because women have two X chromosomes, approximately twice as many women as men have the gene. Thus, the average increase in female fitness need only be half as great as it would if the gene were autosomal (in which case its frequency would be the same in men and women). A hypothesis this clever should be true. With luck, we may know soon. □

1. Hamer, D.H., Hu, S., Magnuson, V.L., Hu, N. & Pattatucci, A.M.L. *Science* **261**, 321–327 (1993).
2. Hamer, D.H. & Copeland, P. *The Science of Desire.* (Simon & Schuster, New York, 1994).
3. Horgan, J. *Sci. Am.* **268**, 122–130 (1993).
4. Hubbard, R. *The New York Times* p. A15, August 2 (1993).
5. Pillard, R.C. & Weinrich, J. D. *Archs. gen. Psychiat.* **43**, 808–812 (1986).
6. Bailey, J.M. & Pillard, R.C. *Archs. gen. Psychiat.* **50**, 217–223 (1991).
7. Risch, N., Squires-Wheeler, E. & Keats, B.J.B. *Science* **262**, 2063–2065 (1993).
8. Hu, S. *et al. Nature Genet.* **11**, 248–256 (1995).
9. Bailey, J.M., Pillard, R.C., Neale, M.C. & Agyei Y. *Archs. gen. Psychiat.* **43**, 808–812 (1986).
10. Greenberg, A.S. & Bailey, J.M. *J. Sex Res.* **30**, 245–251 (1993).
11. Henry, G. W. *Sex variants; a study of homosexual patterns.* Vol 2. 128 (New York, London, P.B. Hoeber, Inc. 1941).
12. Quotes taken from the *Independent*, November 1, 17 (1995).

How do you compute a lod score?

Jurg Ott

Department of Psychiatry, Columbia University, Unit 58 New York, New York 10032, USA

How do you compute a lod score by hand? Many people would be hard pressed to give an answer to this question. Yet up to about 40 years ago, there was no other way to compute lod scores or likelihoods than by hand. Nowadays, various computer programs for calculating lod scores exist, but, in typical linkage analysis applications, these computations are very computer intensive. In this issue of *Nature Genetics*, O'Connell and Weeks[1] describe a significant methodological advance in the area of lod score/likelihood calculation.

I would like to put their contribution into a more general perspective and discuss it on a rather basic and intuitive level.

In a nutshell, the likelihood for pedigree data is simply the probability of occurrence of the pedigree members' phenotypes, with respect

to two (or more) loci, viewed as a function of parameters, notably the recombination fraction, θ; this is written L(θ). The likelihood ratio, L(θ)/L(½), measures the evidence provided by the pedigree data that the recombination fraction is equal to θ (<½, linkage) as opposed to being equal to ½ (no linkage). The lod score, finally, is just the logarithm of that ratio of likelihoods.

For given pedigree data, the likelihood may be developed as a formula with pencil and paper, but this typically is extremely tedious and requires skills in probability calculus and a good understanding of mendelian genetics. It is this combination of scientific fields that makes statistical genetics particularly attractive, but also difficult. Newton Morton[2] in 1955 streamlined lod scores and published lod score tables for many types of two-generation families. An important extension to multigenerational pedigrees became possible with the Elston-Stewart[3] algorithm and the subsequent development of the LIPED computer program[4]. Some ten years later, another breakthrough occurred with the development of the LINKAGE programs[5], which allow likelihood calculations for multiple loci. As linkage analysts know only too well, such calculations can be very time consuming, even on fast computers. Now, after another ten years or so, the new algorithm presented by O'Connell and Weeks[1], also called VITESSE, presents another major step forward in that their methods greatly reduce the computational burden of likelihood calculations.

What problem have O'Connell and Weeks solved? Molecular genetics has provided us with markers that provide an ever increasing heterozygosity. On the one hand, this raises informativeness for linkage by increasing the chance that parents are heterozygous at a region of interest. On the other hand, an increased number of alleles at multiple loci leads to greatly prolonged execution times and/or much heavier demands on computer memory.

Linkage analysts currently tackle this problem in a variety of ways, as discussed by O'Connell and Weeks[1], but none of these

approaches provides satisfactory solutions. Downcoding is one of the techniques being used[6], but it is error prone, and its efficient application may be tedious. Alternatively, one may consider only those alleles actually occurring in a given pedigree while all other alleles in the marker system are lumped into one additional 'mega-allele'. In contrast to this global lumping of alleles, the authors apply this principle locally, that is, for each suitable individual. Specifically, they recognized that, for each individual, all those alleles not transmitted to any descendant need not be distinguished from each other, so they may be lumped into a single mega-allele. Consequently, the authors no longer work with identifiable single alleles but rather with allele sets and call the resulting transmission rules fuzzy inheritance. They further economize their approach by introducing a novel scheme they call set recoding, that is, for each untyped individual, sets of alleles with identical roles are lumped into a single representative allele so that the number of possible genotypes for an untyped individual tends to be greatly reduced. In addition, they apply genotype elimination, which is, in its most general sense, a technique to disregard for a given individual those genotypes that are incompatible with pedigree phenotypes[7].

This all may sound simple enough, but implementing such techniques is tricky and requires great care so that mendelian rules are not violated. Also, while developing new methods is important, they only become really useful when implemented in computer programs. O'Connell and Weeks did this too, and called the resulting computer program VITESSE; it currently works on pedigrees without consanguinity and marriage loops, that is, parents must not be related and, for example, two brothers must not be married to two sisters. As the tables in their paper show, for many pedigrees, the new methods greatly reduce analysis time and memory requirements. Consequently, these authors have greatly extended the limit of practical problems that can be done. This is important because it has been shown repeatedly[8] that

informativeness for linkage is increased with the use of additional flanking marker loci.

It should be stressed that the method developed by O'Connell and Weeks provides exact likelihoods. Alternative approaches to analysing problems larger than those that can currently be handled involve approximations in the calculation of likelihoods, typically carried out with computer simulation as outlined by O'Connell and Weeks[1]. Each of these approaches has its pros and cons. Approximate solutions certainly have appeal, too. Data always contain errors, typically on the order of 1% or higher[9]. In the presence of such errors, 'quick and dirty' results may be sufficient, at least in initial investigations. One may speculate whether accurate or approximate likelihood calculations will eventually dominate the field. In the future, global optimizations in a comprehensive sense may become the standard, where error rates, computer time, the salary of the statistical geneticist, and so on will all be evaluated to minimize the total cost of a certain goal.

In the past 20 years, linkage analysis methods have become very sophisticated. O'Connell and Weeks have just pushed the limits of what can be achieved to a higher level. It has been gratifying to witness this phenomenal pace of development; only a dozen years ago, in my capacity as associate director of the City Statistics Office in Zurich, I was running LIPED and early versions of LINKAGE on an original IBM PC with two floppy drives and no hard drive! □

1. O'Connell, J.R. & Weeks, D.E. *Nature Genet.* **11**, 402–408 (1995).
2. Morton, N.E. *Am. J. hum. Genet.* **7**, 277–318 (1955).
3. Elston, R.C. & Stewart, J. *J. hum. Hered.* **21**, 523–542 (1971).
4. Ott, J. *Am. J. hum. Genet.* **26**, 588–597 (1974).
5. Lathrop, G.M., Lalouel, J.M., Julier, C. & Ott, J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3443–3446 (1984).
6. Ott, J. *Ann. hum. Genet.* **42**, 255–257 (1978).
7. Goradia, T.M., Lange, K., Miller, P.L. & Nadkarni, P.M. *Hum. Hered.* **42**, 42–62 (1992).
8. Kruglyak, L. & Lander, E.S. *Am. J. hum. Genet.* **57**, 439–454 (1995).
9. Brzustowicz, L.M., Mérette, C., Xie, X., Townsend, L., Gilliam, T.C. & Ott, J. *Am. J. hum. Genet.* **53**, 1137–1145 (1993).