



In Search of Human Variation

Kenneth M. Weiss

Genome Res. 1998 8: 691-697

Access the most recent version at doi:[10.1101/gr.8.7.691](https://doi.org/10.1101/gr.8.7.691)

References This article cites 14 articles, 5 of which can be accessed free at:
<http://genome.cshlp.org/content/8/7/691.full.html#ref-list-1>

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).



To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

Cold Spring Harbor Laboratory Press

PERSPECTIVE

In Search of Human Variation

Kenneth M. Weiss¹Departments of Anthropology and Biology, Pennsylvania State University,
University Park, Pennsylvania 16802 USA

There is widespread interest in documenting the amount and geographic distribution of genetic variation in the human species. This information is desired by the biomedical community, who want a densely packed map of SNP (single nucleotide polymorphism) sites to be used to identify genes associated with disease by linkage disequilibrium between sets of adjacent markers and the occurrence of disease in populations, and to characterize disease-related variation among populations. Anthropologists use genetic variation to reconstruct our species' history, and to understand the role of culture and geography in the global distribution of human variation. The requirements for these two perspectives seem to be converging on a need for an accessible, representative DNA bank and statistical database of human variation. However, both fields have been using conceptual models that are oversimplified, and this may lead to unrealistic expectations of the questions that can be answered from genetic data.

“... subtle and difficult of detection.”

With these words recognizing the challenge presented by the biochemical uniqueness of the individual, Archibald Garrod founded the formal discipline of human genetics (Garrod 1902). Ninety years later, genetics is in ascendance. Two areas of investigation that began early in this century have expanded rapidly in recent years: One is the search for genetic variation associated with disease; the other is the use of genetic variation to infer human origins and prehistory.

Both areas are now slowed by the inadequate resolving power of current methods, and better resources are sought in the form of DNA samples more representative of natural variation in our species. Although anthropological and biomedical genetics have different objectives, there are similarities in the resources they seek and the problems to be faced in obtaining them. These problems arise in part from expectations about our ability to answer questions from genetic variation that may not be well matched to the resolution that the processes generating the variation can provide, so rethinking as well as retooling may be necessary.

The Biomedical Perspective

Garrod grappled with the source and amount of in-born variation (Bearn 1993), but human genetics

was then largely restricted to rare, relatively simple afflictions of childhood, like cystic fibrosis, for which phenotypes closely follow Mendel's classical rules of inheritance in families, and there are relatively direct genotype-to-phenotype ($G \rightarrow P$) relationships. The DNA era has been ushered in by a parade of successes in identifying genes responsible for such disorders.

Key to a systematic strategy for finding such genes is the ability to map phenotypic variation to known locations in the genome. Families are screened for co-occurrence of disease and variant alleles at genetic markers that have already been mapped. A major objective of the Human Genome Project (HGP) was to identify a large set of easily typed markers spanning the genome (Collins and Galas 1993; <http://www.nhgri.nih.gov/hgp>). This has been achieved with the mapping of thousands of highly polymorphic microsatellite markers.

Success with single-locus disorders has raised expectations of a similar march through the common, complex chronic diseases that are so important today. Chronic diseases aggregate in families, which suggests that there is a genetic component to be found. But chronic diseases usually do not segregate in neat Mendelian ways in families. Attempts to map such diseases have been plagued by inconsistent and inconclusive results, and weak signal-to-noise ratios compared with clear-cut Mendelian traits.

Part of the problem is that microsatellite mark-

¹E-MAIL kmw4@psu.edu; FAX (814) 863-1474.

WEISS

ers are generally spaced at ~1 cM or greater intervals across the genome (~1% recombination per meiosis between adjacent markers). One centimorgan is ~10⁶ bp, already enough to contain a large number of genes. To identify a narrow chromosomal region to facilitate the search for disease-associated genes, a large number of meioses must be observed. The amount of family data that can be collected may not, in practical terms, provide sufficient numbers. This particularly applies to late-onset multifactorial diseases, in which penetrance (the probability that a person with a given genotype will be affected) may be low, or multiple affected family members difficult to find, weakening the genotype–phenotype associations needed for linkage mapping.

Ideas for ways to enhance mapping resolution have emerged recently, driven in part by the promise of new tools, like oligonucleotide arrays (e.g., hybridization chips), for rapidly genotyping a much denser set of markers. With such data, more powerful mapping approaches become possible. Haplotypes, sequence variants linked together on a single chromosome, reflect the shared population history of contiguous DNA segments. In a small population in which a genetic disease has sufficiently homogeneous etiology, most affected individuals may be distant relatives, essentially comprising a clone of cases produced by alleles identical by descent (ibd) from a single mutation event in the population many generations ago. The deep genealogy that connects the haplotypes of these affected individuals includes the effects of all the meiotic events of the generations since their common ancestor (many more than the few meioses observable in family data). Only a narrow span of marker alleles flanking the causal site will not have been broken up by recombination during those generations. These alleles will have been together on the chromosome since the original mutation, and the length of the shared region will be proportional to the age in generations (number of meioses) since the causal mutation. These markers will be in linkage disequilibrium (LD) with the original causal mutation, that is, statistically associated with the disease. The nearby markers can be used to identify the causal site via this association.

Case–control studies can be used to find spans of markers statistically shared by the cases compared with controls, if the individuals can be typed for a sufficiently dense set of mapping markers. This is a population analog of linkage studies in families that takes advantage of the richer population history of meiotic events to narrow the candidate chromosomal region (Risch and Merikangas 1996; Weir

1996; Kruglyak 1997). LD mapping methods have been used on some occasions in the past to show disease associations with the *HLA* genes, to map cystic fibrosis, and to understand the Rh blood group system. New to LD mapping would be the application of the required high-density genotyping methodology to population studies.

Microsatellite markers have high variability, which made them ideal for disease mapping in family studies, but they have a high mutation rate and the same allele can arise recurrently over time. Even in the absence of recombination, these changes can obscure the trace of shared ancestral relationships between marker and casual disease alleles, reducing the association that is the object of LD mapping in populations. SNPs—single nucleotide polymorphisms—are less variable than microsatellites but are mutationally more stable (e.g., Kruglyak 1997) and may provide a better platform for LD mapping. The National Human Genome Research Institute (NHGRI) is in the process of creating a resource for the identification of a dense SNP map (Collins et al. 1997; Wadman 1998). This resource will include a public domain DNA bank, an online statistical database of SNPs mapped across the genome, and means (e.g., PCR primer sequences) for typing them efficiently. Subjects from any human population could then be typed for up to 100,000 SNP markers in the search for disease-associated genes. But to make this possible, the SNPs must, in some sense, be representative of human variation generally and not just that occurring in a single population. How this will be achieved will be discussed below.

The Anthropological Perspective

Anthropologists have a rather different view of human genetics. Variation is used as a marker, but of geographic rather than chromosomal map location, to understand human history and evolution. This dates back almost as far as Garrod (1902), to the discovery by the Hirzfelds in 1915 that ABO blood group frequencies differed among human populations. Anthropological genetics has progressed from the dull instrument of blood typing, through protein electrophoresis, to RFLPs, microsatellites, and sequences. Allan Wilson energized the field a decade ago by using mitochondrial DNA (mtDNA) as a haploid, nonrecombining molecular clock to argue that humans arose by expansion from Africa ~100,000 years ago (Cann et al. 1987), a date consistent with prior estimates made from classical markers (discussed in Weiss 1988; Cavalli-Sforza et al. 1994). This asserted the authority of genetics,

even before the court of paleontology, and continues to captivate the public media. However, it is important to note that the core questions—the time and place of human origins—are still unresolved, even after a decade of intense work on microsatellites, nuclear haplotypes and sequence variation, and the complementary Y-chromosome haploid system (Templeton 1997; Harpending et al. 1998; Jorde et al. 1998).

At least one reason for these uncertainties is that we have never had a really systematic human geographic marker map with which to work. The patchy, heterogeneous data that have been available have impaired aggregate analysis; different laboratories have typed their samples for different markers, and the DNA is not generally available to any interested laboratory. The resulting noncomparable genotype data are difficult to analyze statistically.

A major decision when the HGP launched the genomics era was to exclude natural variation (except for identifying mapping markers) to avoid having their baseline mapping and sequencing effort confounded by having to deal with variation. Therefore, to obtain a good geographic marker map, Cavalli-Sforza and others proposed a Human Genome Diversity Project (HGDP) (e.g., Cavalli-Sforza et al. 1991; Kidd et al. 1993) to build a bank of DNA samples systematically collected from hundreds of populations around the world, accessible like the CEPH DNAs, with an on-line statistical database like GenBank.

The proposal raised several technical issues involving the sampling design and societal issues concerning informed consent, confidentiality, patenting, and the risk of commercial exploitation of sampled populations. Some of these are particularly emotional, especially among indigenous ethnic groups, who have experienced a history of exploitation by the outside world (see Vol. 20, Summer 1996, *Cultural Survival*) and who may fear inclusion in such a project more than they do exclusion from biomedical science. Much of the developing world has heard denials of exploitive intent from (sometimes) well-meaning, but ultimately self-serving scientists before, and are understandably wary. Little in the hubris of current human genetics would instill confidence to the contrary; one must assume the necessity for any such project to be operated only under external regulation.

Because this was a major proposal and raised so much controversy, a National Research Council (NRC) Committee was formed to evaluate the idea, much as occurred prior to the HGP. Their report

(<http://www.nap.edu/readingroom/books/genetic/>) notes the potential value of a global assessment of human genetic variation, accepts many of the premises of the original HGDP design (like accessible DNA samples), reiterates the ethical concerns, and asks for a more precisely specified sample design. The NRC urges that such a project be started, but that to avoid the problems of achieving and monitoring international consensus, sampling should initially be restricted to US-based populations or investigators where institutional review board ethical standards can be imposed.

Geneticists should try to understand the visceral nature of the most sensitive ethical concern—that study of global variation might reinforce racism or ethnic determinism by virtue of its sample design alone. If there are objections to the human genome because it omits variation, there are objections to dividing the world into population units stereotyped by race, language, ethnicity, or nation. Such categories are to some extent a product of our own culture, and may impose a culturally imperious imprint that may not accurately reflect either the distribution of variation or the world view of other peoples. For example, many times we (people from the developed world) give a group an ethnic name that we take as equivalent to nation, when to them (those in other populations) it only means people. In fact, genetic variation is basically continuously distributed over geographic space, no matter how we label or package our necessarily discrete samples of that variation.

An illustrative example are studies of the correspondence between genes and populations hierarchically categorized by language. Such a treatment might be erroneously interpreted as suggesting internal homogeneity and species-like phylogeny of human populations according to ethnic traits like language, or that indigenous peoples are frozen replicas of the human past, a kind of Victorian taxonomy of people denied their own history (e.g., Wolf 1982; Moore 1995). People in large numbers continue to pay with their lives for strife rationalized by equating inborn value with geography, ethnicity, or nation. English-speaking North American readers will include people from all over the globe, who speak a language that was brought here from England, and should understand that language, culture, genes, and geography do not necessarily mark the same thing, much less do they support a bifurcational history of our forebears.

The human species inhabits a geographic range that is vast relative to the traditional migration distance (from birth to death) of individuals. As a re-

WEISS

sult, genetic differences among populations, averaged over many loci, tend to correspond to geographic distance and are not unexpectedly correlated with contemporary aspects of ethnicity for similar reasons. But ethnicity is not phylogeny. Individual genes have patterns of geographic variation determined by gene-specific histories of drift, mutation, and selection.

A Convergence of Interests

Anthropological and biomedical geneticists alike realize the value of a systematic, representative resource for characterizing human genetic variation, respectively, to understand and to use the results of population history. In designing both the proposed SNP identification resource and the HGDP, similar problems must be faced. These include the ethical issues just discussed. There is also a shared need to represent genetic variation that arose on all continents, which in practice means a stratified sample of some kind. Both projects face the fundamental question: Which populations should be sampled?

The HGDP has objectives that require detailed geographic knowledge of human variation and proposes a base collection of at least 25–100 individuals, from at least 100 populations from each continent, a total sample size of >10,000. Choosing the populations immediately raises the ethnicity issues just described. An acceptable, systematic design for sampling the world would best be developed through open discussion of the issues. The HGDP organizers began this in an exploratory way, but it seems that in the absence of a single funding or organizing source, practical politics including national interest, nervousness about exploitation by the west, and variable ethical criteria around the world, may restrict what can actually be achieved. What seems most likely to develop is a loose international consortium, with locally determined sample choice, that is a mixture of samples of convenience and locally identified ethnic groups. Activities are under way in many countries, but whether this will lead to a global, well-understood, accessible DNA resource, without exploitation, is unclear.

The SNP map project faces practical constraints and has more general objectives. Its organizers seek a modest but immediate goal of ~100 individuals from single samples representing each continent, a total of ~500 individuals. For practical reasons, this must be U.S. based, and the obvious source is the major ethnic groups here. By understanding the nature of ethnicity in the Americas, the SNP mapping

project can take advantage of the correlation between geographic origin and ethnic identity that is the result of reasonably understood patterns of immigration and subsequent admixture here (in contrast, the HGDP faces an essentially undocumented global prehistory of population dynamics).

This admixture history can be used to apportion genomes collected in samples from U.S. ethnic groups to their geographic origins. For example, the African American population is, genetically, ~85% African and 15% European in geographic origin (Chakraborty et al. 1995). A sample of 118 African Americans and 82 European Americans would thus yield a net sample of ~100 genomes from each continent of origin. Other populations can be treated similarly.

There is considerable debate about the nature and size of samples required to identify SNPs. Some advocate using only a small sample of 10 or less individuals, identifying variants at high frequency likely to be found globally. Others believe that more population restricted, or recent, variants need to be found. In fact, we do not know how efficient each strategy will be (see below). Different investigators are likely to use different-size subsets of the basic DNA resource.

Although it is interesting to compare the relative roles of culture and simple isolation by distance, in the dispersion of human variation, the Americas are unusual in terms of the very recent, large-scale, long-distance nature of the amalgamation that has taken place here. More generally, however, ethnicity may be ephemeral relative to deeper human population history. Rather than struggle to identify an appropriate set of ethnically defined sampling units, some investigators have suggested that individuals instead simply be sampled from a geographic grid. Whether or not this is practicable on a global scale, thinking in terms of geography rather than ethnicity might help liberate ethnicity from the biological determinism that has plagued the first century of human genetics.

The overall sample design for the SNP resource will be made public, but the ethnic identity of the individual DNA samples will not be provided. This decision has been controversial. Laboratories can find SNPs without knowing the ethnic origin of the DNAs, but users of the SNPs for disease-mapping purposes might benefit from knowing where the various alleles were found in the world. The ethnic identity of individual samples is being withheld to prevent any misuse that might promote ethnic or racial determinism. However, the decision to with-

hold that information might be misunderstood as a tacit admission that unpleasant truths about ethnicity exist but are being hidden; a less defensive and more proactive position might be preferable from our public institutions.

Some serious issues remain in the design of representative population samples. For example, alleles found in 100 individuals (200 genomes) from a single sample of a given continental origin will have frequency $\geq 0.5\%$ in the sample, but if there is population substructure, a single local sample may not accurately reflect variation elsewhere on the continent (e.g., Weir 1996). Sampling from agglomerated sources, such as urban areas that include people from many parts of the continent, should ameliorate this problem because alleles that are common over a large part of the region are likely to be included. Similar issues pertain to the number and nature of HGDP-related samples, though on a different scale. Basically, allele frequencies and frequency-dependent measures like linkage disequilibrium can only be estimated accurately from properly identified and sampled populations.

Successful LD mapping depends on the association between marker and disease-related alleles. It is already clear that some human variants are present or common only in restricted regions. Association is a frequency-based phenomenon that depends on several factors, including the presence, relative ages, and frequencies of disease mutations and the marker alleles used to search for them; whether recombination has occurred between them; the degree of etiologic heterogeneity; and the mutation rate and heterozygosity (amount of variation) at the marker loci. These factors all depend on the population history of the group in which a study is carried out. Each disease may have to be approached differently.

For this reason the sample design is important, and a general-purpose SNP resource may prove to be less powerful than is currently hoped. It is likely often to be necessary to study a disease in a specific local population in which it is particularly common or homogeneous in its manifestation. The most celebrated successes of LD mapping to date have been in religious isolates or populations like parts of Finland, where the founder effect has produced homogeneity for relatively recent mutations for otherwise rare disorders. In some circumstances, microsatellites rather than SNPs, or even just the amount of disequilibrium rather than association between specific alleles and disease (Terwilliger et al. 1998), may be more powerful if disease-associated mutations are recent. In any case, mapping in local popula-

tions requires detailed knowledge of local variation, and it is foreseeable that if LD mapping shows promise, a second-generation resource to include detailed SNPs specific to many local populations may be needed—a direct convergence with the more detailed sampling interests of a HGDP.

None of these strategies will work well if the relationship between causal alleles at a locus and the flanking sites that happen to be selected for the marker map is not sufficiently strong. The pattern of LD among sites within a gene or between a gene and nearby markers may be more complex than has been thought and may differ greatly depending on which marker sites are chosen (Clark et al. 1998). Generally, we do not yet know how powerful LD mapping will turn out to be.

In addition to these problems there is a serious issue related to statistical power. An increase of the number of markers to 100,000 SNPs would seem an obvious improvement in detection power. But more is not always better. Even for single-site association tests some adjustment in the statistical significance level will be needed to avoid being swamped by thousands of false positives. The usual approach is to use a very stringent significance level for a single-site test before it is considered useful evidence of association. But with this many markers, the sample sizes of cases and controls required to do this adequately may be unachievable, in practice. This is true even for simple, fully penetrant homogeneous etiology; the situation will be worse for less clear, heterogeneous etiology as is found in complex chronic diseases.

Digital Sampling of an Analog World. Is Retooling Enough?

If chronic diseases are causally heterogeneous and difficult to map, and the most basic questions about human origins are not clearly resolved, it is fair to ask whether better mapping tools are the answer. Or do we need to rethink the questions as well?

In both instances, problems arise from what can be described as digital sampling to understand an analog world of variation. The elusiveness of complex diseases arises in part from trying to use discrete genetic variation to map what are essentially quantitative phenotypes like cholesterol level or risk of cancer. Refined mapping efficiency can only work if what is finely (finally) mapped is related materially to disease outcomes. In many ways we have been forcing classical Mendelian two-allele (e.g., normal–abnormal) concepts onto a more complex reality (e.g., Weiss 1996). Many biomedical ge-

WEISS

neticians still speak as if there were a wild-type normal allele and one or only a few disease alleles. However, we know that human genes have hundreds of alleles (and thousands of possible genotypes), varying among populations, and that each allele has its own particular effects on the trait. In the end, even a single locus generates an essentially quantitative spectrum of effects such as disease severity.

There has also been a kind of theory creep associated with the widespread contemporary belief in the primacy of genes and genes as the object of evolution (e.g., Strohman 1997). In important ways, this inverts the direction of evolutionary causation. Only the phenotype is the direct object of natural selection. It seems not to be widely appreciated that selection does not specify a single good sequence for a gene but, instead, is a tolerant process that allows as much variation as can survive to survive—survival of the fit (and lucky) rather than just the fittest. Aspects of genes vital to survival or embryogenesis, or embedded in pleiotropic interactions, may be less variable, but for the same reason cannot be responsible for much late-onset disease. Mutation is an inexorable source of new variation, most mutations are unique at the DNA level, and most mutations have little fitness effect (especially for postreproductive chronic disease). Because it screens phenotypes rather than genotypes, evolution generates heterogeneity in the $G \rightarrow P$ relationship among and within populations.

This is what has been discovered in biomedical genetics as well. Things are most tractable and mappable where the $G \rightarrow P$ relationships are simplest—in genes vital to survival, in families or isolated populations with high identity by descent, and for rare, idiosyncratic phenotypes. Darwin, who knew no genes, recognized this, and it was the basis of Garrod's work. Better mapping tools may not answer questions about disease in the general population when based on a misplaced expectations of the biological reality.

In a somewhat analogous way, the lack of resolution of questions in anthropological genetics also has to do with inadequate assumptions about population history. We treat digital (spot) samples of populations as meaningful units relative to the essentially analog (quantitative) processes that distribute variation across time and space. Population history is tolerant to breaches of ethnic and geographic boundaries and generates a distribution of variation that is often more quantitative than qualitative. We should not expect genotypes to capture population identity precisely. Local populations are

nearly as variable as our whole species (Cavalli-Sforza et al. 1994; Barbujani et al. 1997). Humans appear to have had a more episodic than stable population history (Templeton 1997; Harpending et al. 1998; Jorde et al. 1998), so that theoretical models of neutral equilibrium fail to fit many aspects of human data. One result is the ambiguity in the inferences we can make about our history. For example, rapid expansion by small subpopulations drawn from a large source population, with no subsequent migration among them, can yield patterns of variation similar to those of a large stable population with migration between local areas. Methods for using genetic data to infer history more definitively are under active investigation.

Conclusion

Evolution generates heterogeneity because it is a tolerant process, rather than one that prescribes a single normal allele and purges the population of other variation. Genetic variation is tied less tightly to phenotypes and is similarly a more ambiguous reflection of population history than is often thought. Except when caused by rare mutations with dramatic effect, complex disease traits typically aggregate but do not segregate in families, and human variation aggregates but does not segregate discretely among populations. Finding those rare dramatic mutations can be an important tool of discovery but may fail for similar reasons to provide the closed explanations of the same traits in the broader population. In his context, Garrod knew this 90 years ago, and he was circumspect about the ability of genes to explain the biochemical uniqueness of each human being.

It is important that we use what we already know about genetic variation in a sophisticated way to develop durable resources for understanding what we do not yet know. To be considered acceptable, such resources should adequately represent the diverse populations that comprise our species. We should embrace rather than fear that diversity, for both its biomedical and historical value. Perhaps we also should temper our expectations to be appropriate for the questions that we can answer with genetic data.

REFERENCES

- Barbujani, G., A. Magagni, E. Minch, and L.L. Cavalli-Sforza. 1997. An apportionment of human DNA diversity. *Proc. Natl. Acad. Sci.* 94: 4516–4519.
- Bearn, A.G. 1993. *Archibald Garrod and the individuality of man*. Clarendon Press, Oxford, UK.

IN SEARCH OF HUMAN VARIATION

- Cann, R., M. Stoneking, and A. Wilson. 1987. Mitochondrial DNA and human evolution. *Nature* 325: 31–36.
- Cavalli-Sforza, L.L., A. Wilson, C. Cantor, R. Cook-Degan, and M.-C. King. 1991. Call for a worldwide survey of human genetic diversity: A vanishing opportunity for the Human Genome Project. *Genomics* 11: 490–491.
- Cavalli-Sforza, L.L., P. Menozzi, and A. Piazza. 1994. *History and geography of human genes*. Princeton University Press, Princeton, NJ.
- Chakraborty, R., M. Kamboh, M. Nwankwo, and R. Ferrell. 1992. Caucasian genes in American Blacks: New data. *Am. J. Hum. Genet.* 50: 145–155.
- Clark, A. K. Weiss, D. Nickerson, S. Taylor, A. Buchanan, J. Stengard, V. Salomaa, E. Vartiainen, E. Boerwinkle, and C. Sing. 1998. Haplotype structure and population genetic inferences from nucleotide sequence variation in human lipoprotein lipase. *Am. J. Hum. Genet.* (in press).
- Collins, F. and D. Galas. 1993. A new five-year plan for the US Human Genome Project. *Science* 262: 43–46.
- Collins, F., M. Guyer, and A. Chakravarti. 1997. Variations on a theme: Cataloging human DNA sequence variation. *Science* 278: 1580–1581.
- Garrod, A. 1909. *Inborn errors of metabolism. The Croonian lectures delivered before the Royal College of Physicians of London, in June, 1908*. Oxford University Press, London, UK.
- Harpending, H., M. Batzer, M. Gurven, L. Jorde, A. Roges, and S. Sherry. 1998. Genetic traces of ancient demography. *Proc. Natl. Acad. Sci.* 95: 1961–1967.
- Jorde, L., M. Bamshad, and A. Rogers. 1998. Using mitochondrial and nuclear DNA markers to reconstruct human evolution. *BioEssays* 20: 126–136.
- Kidd, J.R., K.K. Kidd, and K.M. Weiss. The human genome diversity initiative. *Hum. Biol.* 65: 1–6.
- Kruglyak, L. 1997. The use of a genetic map of biallelic markers in linkage studies. *Nature Genet.* 17: 21–24.
- Moore, J.H. 1995. Putting anthropology back together again: The ethnogenetic critique of cladistic theory. *Am. Anthropol.* 96: 925–948.
- Risch, N. and K. Merikangas. 1996. The future of genetic studies of complex human diseases. *Science* 273: 1516–1517.
- Strohman, R. 1997. The coming Kuhnian revolution in biology. *BioTechnology* 15: 194–200.
- Templeton, A.R. 1997. Out of Africa? What do genes tell us? *Curr. Opin. Genet. Dev.* 7: 841–847.
- Terwilliger, J., S. Zollner, M. Laan, and S. Paabo. 1998. Mapping genes through the use of linkage disequilibrium generated by genetic drift. *Hum. Hered.* (in press).
- Wadman, M. 1998. Gene marker database stirs up debate. *Nature* 391: 726.
- Weir, B. 1996. *Genetic Data Analysis. II*. Sinauer, Sunderland, MA.
- Weiss, K. 1988. In search of times past: Gene flow and invasion in the generation of human diversity. In *Biological aspects of human migration* (ed. N. Mascie-Taylor and G. Laskev), pp. 130–166. Cambridge University Press, Cambridge, UK.
- Weiss, K.M. 1996. Is there a “paradigm” shift in genetics? Lessons from the study of human diseases. *Mol. Phylogenet. Evol.* 5: 259–265.
- Wolf, E. 1982. *Europe and the people without history*. University of California Press, Berkeley, CA.