Chapter 1

Foundations of Anthropological Genetics

Michael H. Crawford

Department of Anthropology University of Kansas, Lawrence, KS 66045

What is anthropological genetics?

Anthropological genetics is a synthetic discipline that applies the methods and theories of genetics to evolutionary questions posed by anthropologists. These anthropological questions concern the processes of human evolution, the human diaspora out of Africa, the resulting patterns of human variation, and bio-cultural involvement in complex diseases. How does anthropological genetics differ from its kin discipline, human genetics? Both fields examine various aspects of human genetics but from different perspectives. With the synthetic volume of 1973 (Methods and Theories of Anthropological Genetics), it became evident that the questions posed by the practitioners of anthropological genetics and human genetics tended to be somewhat different. I compared and contrasted these two fields in the introduction to the special issue of *Human Biology* (2000) on Anthropological Genetics in the twenty-first century (see Table 1.1). What distinguishes anthropological genetics from human genetics is its emphasis on smaller, reproductively isolated, non-Western populations, plus a broader, biocultural perspective on evolution and on complex disease etiology and transmission. Judging from the contents of the *American Journal of Human Genetics* (premiere journal in the field of human genetics) there is a greater emphasis on the causes and processes associated with disease, and the examination of these processes in affected phenotypes (probands) and their families. Anthropological geneticists tend to focus more on normal variation in non-Western reproductively isolated human populations (Crawford, 2000). Anthropological geneticists also attempt to measure environmental influences through co-variates of quantitative phenotypes, while human geneticists less often attempt to quantify the environment in order to assess the impact of environmental-genetic interactions.

Table 1.1. | Differences between human genetics and anthropological genetics (Crawford, 2000).

Anthropological genetics	Human genetics
1. Broader biocultural perspective on genetic/environmental interactions	1. Mechanisms and processes — particularly in disease
2. Population focus, pedigrees utilized to measure familial resemblance	2. Families of proband, twins and twin families
3. Small, reproductively isolated populations — often, non-Western	3. Larger, urban, and clinical samples
4. Culturally homogeneous populations	4. Populations may be heterogeneous by race, socio-economic factors, occupation, and lifestyle
5. Sampling representative of normal variation in population	5. Sampling based on clinical ascertainment
6. Attempts made to characterize and measure the environment	6. Environmental variation rarely assessed. It is assumed that $e^2 = 1 - h^2$
7. Study of normal variation in complex traits	7. Dichotomy of disease vs. normality — usually observed

History of anthropological genetics

The ancestral roots of the field of anthropological genetics are intimately intertwined with the developments in evolutionary biology, population genetics, and biological anthropology. O'Rourke (2003) correctly noted that this modern amalgamated discipline was further cross-fertilized by molecular biology and bioinformatics. Through cross-fertilization this hybrid field has acquired the analytical and laboratory tools to dissect the molecular and genetic bases of human variation, a traditional focus within biological anthropology. The addition of genome scanning and linkage analyses have contributed to the fluorescence of genetic epidemiology and the mapping of genes involved in complex phenotypes, particularly those associated with chronic diseases.

Anthropological genetics of the late 1960s and early 1970s was preceded by almost a century of discovery and development in evolutionary theory and genetics. Many of the ideas associated with natural selection can be traced to the publication of Charles Darwin's *Origin of Species* in 1859 (see Table 1.2). Because Darwin was unaware of Gregor Mendel's experiments on the particulate nature of genes (using characteristics of pea plants) Darwin lacked specific mechanisms for generating new variation and had to settle for a blending form of inheritance. Darwin also used Lamarck's concept of the inheritance of acquired characteristics, a concept that persisted well into the twentieth century.

Table 1.2. Time-line of significant developments in genetics and anthropological genetics.*

1859	Publication of Darwin's *Origin of Species*
1860	Meischer first isolated DNA
1880	Weismann demonstrated the separation of the germ plasm from somatic cells
1900	Rediscovery of Mendel's laws of inheritance
1901	Documentation of first polymorphism in humans, ABO blood group system, by K. Landsteiner
1902	Garrod demonstrated that the mode of inheritance of inborn errors of metabolism were Mendelian in nature
1908	Formulation of the principle of genetic equilibrium, generally attributed to Hardy and Weinberg, but preceded by Castle in 1903
1919	Population variability in the frequency of blood group genes demonstrated in World War I by Hirschfeld and Hirschfeld
	Fisher integrated Darwin's theory of natural selection with Mendel's formulations
1930-2	Fisher, Haldane and Wright publish the mathematical basis of modern population genetic theory
1937	Dobzhansky published *Genetics and the Origin of Species* and further fleshed-out the modern synthesis by reconciling the evidence of the naturalists with the geneticists
1944	DNA is shown to be heritable material
1949	Molecular basis for sickle cell disease demonstrated by J. V. Neel (1957), Pauling *et al.* (1949), and H. A. Itano and L. Pauling (1961)
1953	Watson and Crick break the genetic code
1954	Allison reveals relationship between sickle-cell trait and malaria
1956	Human chromosomal numbers correctly characterized by J. HinTjio and A. Levan (1956)
1955	Smithies (1955, 1959) develops starch-gel electrophoresis, a method for separating protein variation based on charge and size of molecules
	Y-chromosome shown to determine the sex of organisms
1972	Lewontin apportioned human genetic diversity and demonstrated the 85% is within populations
1973	Publication of the first major synthesis of anthropological genetics, Crawford and Workman
1977	DNA sequencing methods described
1978	Restriction Fragment Length Polymorphism (RFLP) first described
1981	Human mtDNA genome sequenced
1984	Methods of DNA fingerprinting first described by Jeffries
1985	Development of Polymerase Chain Reaction (PCR) methods
1987	Development of laser based fluorescent detection of DNA
1988	Beginning of the Human Genome Project
1991	Human Genome Diversity Project Proposed
1997	First Neandertal mtDNA sequence
1998	Completion of sequencing of the first human chromosome (Ch. 22)
2001	Draft of human genome sequence

*Time-line was based in part on Jobling, Hurles and Tyler-Smith (2004).

Despite the brilliant research of August Weismann, who demonstrated the separation of germ plasm from the soma, Lamarckian concepts were adopted in Stalinist Soviet Union because they better fitted the ideology. Taken to extremes, there was a belief that changes in the phenotype affect the genotype, which is then transmitted to the next generation. However, later geneticists demonstrated that the alternation of the phenotype does not get inherited by subsequent generations because of the separation of the sex cells from other somatic cells. The concept of mutation initially arose from Hugo DeVries' research on primroses. He concluded that most mutations had drastic effects and that speciation was driven by mutations. However, the creative research of Thomas Hunt Morgan on the fruit fly demonstrated that mutations introduced variation in populations at incremental levels but rarely resulted in the formation of new species.

Measures of human variation

Blood groups

At the turn of the twentieth century, Karl Landsteiner's (1901) immunological characterization of the ABO blood group system and its mode of inheritance provided a genetic marker for the measure of human variation. Ludwik Hirschfeld and Hanka Hirschfeld (1919), during World War I, demonstrated that military personnel of various so-called 'racial groups' or ethnicities differed in the frequencies of the ABO blood groups. In the few decades that followed, additional blood group systems, such as the Rhesus, MNS, Duffy and Diego, were shown to vary in human populations. These blood group systems were polymorphic and differed in allelic frequencies in human regional populations. Yet, the function of these complex gycolipid (sugar/fat) or glycoprotein (sugar/protein) molecules expressed on the surface of human erythrocytes were unknown until relatively recently. For example, the Rh (Rhesus blood group system), discovered by Landsteiner and Weiner in 1941, came to medical attention because of its importance in pregnancy and the risk of maternal/fetal incompatibility. It only became evident in the year 1997 that the evolutionary history of the Rhesus blood group system was of great antiquity and that the function of the RH glycoproteins is the transportation of ammonium ions (NH4+) across the cell membrane (Marini and Urrestarazu, 1997; Marini *et al.*, 2000). My chapter, The Use of Genetic Markers of the Blood in the Study of the Evolution of Human Populations, contains a summary of the known genetic variation in human blood group systems (Crawford, 1973).

Protein variation

In the mid 1950s, Orville Smithies developed a method (starch gel electrophoresis) for separating protein mixtures on the basis of both the electric charge and the size of the molecules (Smithies, 1955).

Thus, the degree of genetic variation in the serum proteins in populations could be ascertained electrophoretically (Smithies and Connell, 1959; Smithies, 1959). This methodological innovation provided a glimpse into the genetic variation contained within the human genome, using primary gene products such as serum and red blood cell proteins and enzymes (extracted from the red blood cells into hemolysates). Refinements in electrophoretic methods, from filter paper electrophoresis (which separated proteins on the basis of molecular charge) to starch gel electrophoresis (separation based on both the charge and the size of the molecule), to isoelectric focusing (IEF, a method of electrophoresis performed on a gel containing a pH gradient), were suggestive of a human genome consisting of approximately 100,000 genes. With the sequencing of the human genome, this estimate was later down-sized to approximately 30,000 genes.

Modern evolutionary synthesis

Most fields of inquiry are fortunate to have one, or maximally two, highly innovative 'founders', such as a Charles Darwin or an Albert Einstein. However, in addition to Charles Darwin, evolutionary theory was developed by three contemporaneous major figures, namely: Sewall Wright, J.B.S. Haldane, and R.A. Fisher (Table 1.2 contains a time-line of the significant genetic breakthroughs). They set the mathematical foundations for the modern synthetic evolutionary theory and provided the formal underpinnings for the measurement of natural selection and statistical methods for estimating the effects of stochastic processes. Other scientists, such as Thomas Hunt Morgan and Ernest Muller, using animal (fruit fly) models provided an understanding of mutation, the source of new genetic variation – which had eluded Charles Darwin. In an essay celebrating his 100th birthdate, the last survivor from that period of the development of the evolutionary synthesis, the eminent German evolutionary biologist, Ernst Mayr, recently reminisced about that era of evolutionary theory development (Mayr, 2004).

The next generation of population geneticists included the distinguished Russian émigré geneticist, Theodosius Dobzhansky, the great Chinese agronomist, C. C. Li, and US-born human geneticist, James Crow. They collectively added further refinements and detail to the synthetic theory of evolution. Although Dobzhansky's 'animals of choice' were the beetle and the lowly fruit fly (*Drosophila melanogaster*), he applied the principles of evolution learned from these models to humans and synthesized the available information on human evolution in a readable form. Similarly, C. C. Li synthesized much of the theory of population genetics in his concisely written primers, which assisted in the training of subsequent generations of evolutionists. James Crow coalesced demographic characteristics with population genetics by developing a method for assessing the opportunity of natural selection in human populations, based on

fertility and mortality components derived from church records and civil documents. Together with his former student, Arthur Mange, Crow also developed methods for estimating levels of inbreeding in human populations using marital records and the likelihood of individuals with the same surname marrying (isonymy). These methods were applied and further elaborated by anthropological geneticists, such as Gabriel W. Lasker (a Harvard Ph.D., trained, in part, by Ernst Hooton). James Spuhler, another student of Hooton's, was greatly influenced by Sewall Wright and applied some of the path methods for the computation of inbreeding coefficients to Ramah Navajo populations (Spuhler and Kluckhohn, 1953). Derek F. Roberts, an Oxford-trained biological anthropologist, applied some of Wright's formulations to an island population in the south Atlantic, Tristan da Cunha, and demonstrated the importance of unique historical events and founder effects on the population of this small, remote island (Roberts, 1969). He also described the high incidence of forms of congenital deafness and mental retardation in the Tristan population (1969) and more recently showed the reduction in genetic variation as assessed by mtDNA (Soodyall *et al*., 1997).

In the late 1950s and early 1960s, with the publication of Sewall Wright's insights into the actions of stochastic processes, physicians and medical geneticists discovered the usefulness of small, genetically isolated populations for the understanding of rare genetic diseases and anomalies. Recessive mutations (normally of low incidence in large populations) may appear at high frequencies in some of these small populations because of the founder effect and chance segregation. Victor McKusick, of Johns Hopkins University, spearheaded the study of rare genetic anomalies in Pennsylvania Amish populations. The value of this approach was further demonstrated by the discovery of rare, familial genetic conditions, such as Christmas hemophilia, forms of dwarfism, and adenylate kinase deficiency in Amish kindred. Physicians/geneticists such as Victor McKusick (1964), Arno G. Motulsky (1965) and James V. Neel (1957) integrated biochemical genetic methodologies with evolutionary theory to elucidate human adaptation to diseases such as malaria. L. L. Cavalli-Sforza, another medically trained geneticist, examined allelic frequency fluctuations due to stochastic processes in small villages of Parma, northern Italy. Recently, together with colleagues Moroni and Zei, Cavalli-Sforza expanded this research into a tome on consanguinity, inbreeding, and genetic drift in Italy (Cavalli-Sforza *et al*., 2004). During the 1960s, Motulsky followed up his biochemical genetic interests in metabolic diseases to conduct fieldwork in populations of the Congo (Motulsky 1960; Motulsky *et al*., 1966). Similarly, J. V. Neel, together with Brazilian geneticist, Francisco Salzano, mounted a highly successful research programme on the genetics of tribal populations of South America (1964). Thus, in the 1950s and 1960s, the margins between the fields of anthropological genetics and human genetics were somewhat blurred, with geneticists and physicians conducting anthropological research

and anthropological geneticists working in the realm of human genetics.

Early anthropological genetics

Anthropologists with training in genetics were useful to the medical profession in studies of small, highly isolated, non-Western populations. Unfortunately, until the 1950s, there were few anthropologists with adequate training in human genetics. The reason behind this paucity was that most physical anthropologists were traditionally trained in morphology and racial classification based on typology. However, several of Albert Hooton's last group of doctoral students at Harvard, namely Gabriel Lasker, Frederick Hulse and James Spuhler, had some training and interest in genetics. Lasker was influenced by the writings of Sewall Wright and applied these ideas to his field investigations with Mexican and Peruvian populations (Lasker, 1954, 1960). Hulse examined linguistic barriers to gene flow and blood group variation in Native American populations of northwestern United States (Hulse, 1955, 1957). In addition, he measured the effects of heterosis and exogamy in small-sized, Alpine Swiss communities (Hulse, 1957). James Spuhler collaborated with the cultural anthropologist, Clyde Kluckhohn, in applying Sewall Wright's pathway methods and measured the level of inbreeding among the Ramah Navajo (Spuhler and Kluckhohn, 1953).

Frank Livingstone, a former student of Neel and Spuhler at Michigan, conducted a study on the effects of culture (i.e. the introduction of slash-and-burn agriculture into sub-Saharan Africa) on the distribution of falciparum malaria. He demonstrated in his classic dissertation and subsequent publications that the destruction of the tropical rain forest resulted in the creation of standing bodies of water, a prerequisite for the successful breeding conditions of the Anopheles mosquito (Livingstone, 1958). The increased parasitization caused a shift from epidemic to endemic malarial infection and the action of natural selection against various phases of the life cycle of *Plasmodium falciparum*. Livingstone and Neel also trained a number of anthropological geneticists at Michigan, e.g. Kenneth Weiss, Alan Fix and the late Richard Ward – all went on to distinguished careers in anthropological genetics.

Several graduate anthropology students of W. W. Howells and Albert Damon at Harvard applied population genetic principles to anthropological populations. Eugene Giles tested theory of genetic drift on field populations of New Guinea. He sought to document that gene frequency fluctuations were due to genetic drift in small, isolated villages (Giles *et al*., 1966). Jonathan Friedlander, a graduate student of Damon's, conducted anthropological genetic investigations in the Solomon Islands (Friedlander, 1971).

Richard Lewontin (a population geneticist) statistically partitioned genetic variation within populations and between populations on the basis of 15 protein loci (Lewontin, 1967). He demonstrated that

85% of human genetic diversity is within populations. Thus, a much smaller percentage, 15%, is between populations. This research has been used to discourage genetic comparisons between so-called geographical 'races' because most of the variation is contained within the populations. Barbujani (1997) retested Lewontin's findings based on DNA markers and confirmed that 84.4% of the variation was within populations, 4.7% among samples, within groups, and 10.8% among groups (see Chapter 2, Madrigal and Barbujani). However, a controversial analysis of single nucleotide repeat (SNP) diversity (Seielstad *et al*., 1998) indicated that while autosomal and mtDNA SNPs provide a pattern similar to that observed by Lewontin and Barbujani (within populations 85.5% and 81.4% of the variation is subsumed), Y-chromosomal SNPs apportion almost 53% of the variation between continental populations.

Foundations of anthropological genetics

In 1988, when I assumed the editorship of the journal *Human Biology*, I inherited few manuscripts of publishable quality. Kenneth Weiss (a member of the editorial board) suggested that in celebration of the 60th anniversary of the journal I should consider publishing an issue of the journal containing the 'best' anthropological genetics articles that had graced the pages of *Human Biology* during its history. This special issue would, on one hand, provide the needed manuscripts for publication plus, on the other hand, connect the past with the new focus of the journal. I titled this special issue 'Foundations of Anthropological Genetics'. Gabriel Lasker and I selected the 'top-ten' most significant articles and had most of the original authors update their thoughts on the topic (Crawford and Lasker, 1989). Two of these classic articles were written almost 50 years ago, thus necessitating the preparation of the updates by willing specialists, namely David Hay and Robert Sokal, rather than the original authors. This special issue does establish connections between the early research in human genetics and the developments in anthropological genetics. The ten articles selected included distinguished authors such as J. B. S. Haldane, James Spuhler, D. F. Roberts, James Crow, J. V. Neel, Frank B. Livingstone, A. G. Motulsky, Morris Goodman, and P. T. Wilson. Their research and publications established a solid base, or foundation, for the field of anthropological genetics. While only Spuhler, Livingstone and Roberts were considered biological anthropologists, the field of anthropological genetics was built on the research and formulations of many disciplines and theoretical approaches to human evolution.

The synthesis

In 1970, due in part to prompting by my colleague at the University of Pittsburgh, the late C. C. Li, I organized a symposium on methods and theories of anthropological genetics at the School of American Research in Santa Fe, New Mexico. This symposium had a blend

Fig I.1 The participants in the School of American Research symposium on Anthropological Genetics, held in 1970, in Santa Fe, New Mexico. Back row (standing, from the left): Steven Vandenberg, Jean Benoist, Frank Livingstone, Gabriel W. Lasker, Peter L. Workman, Eugene Giles, Christy Turner III, Francis Johnston, and James Spuhler. Front row (seated): Michael H. Crawford, Derek F. Roberts, and William W. Howells.

of senior, established scholars, such as James Spuhler, William W. Howells, Gabriel W. Lasker, Frank Livingstone, Steven G. Vandenberg, and Derek F. Roberts. In addition, a number of younger researchers were invited: Eugene Giles, Peter L. Workman, Jean Benoist, Christy G. Turner, III, and Francis Johnston. Figure 1.1 identifies the original participants in the 1970 symposium in Santa Fe, New Mexico. During this symposium, the participants instructed Peter Workman and me to serve as the editors of the volume that was to be compiled and published in the series established by Douglas Schwartz with the University of New Mexico Press. Because we were limited to a small number of participants at the symposium, Workman and I decided to solicit a few additional chapters to fill the obvious lacunae. We added two contributions by human geneticists and six by anthropologists of a 'genetic persuasion'. The genetic additions included Newton E. Morton (who had elaborated on Malecott's bioassay approaches to the study of population structure and applied them to populations of Micronesia) and Jean W. McCluer (who at that time was applying computer simulation methods to the demographic structures of South American Native populations). The anthropological geneticists added to the mix included Henry Harpending (R-matrix analysis of South African populations), Russell Reid (synthesis of theory on inbreeding), and Solomon Katz (fearless prognosticator of the evolutionary future of humans), the late Richard H. Ward (genetic structure of Amazonian populations), Nancy Howell (the feasibility of characterizing the demographic structures of small populations), and Kenneth Morgan (historical demography of a Navajo community). Initially, Workman and I debated whether to

name the volume *Methods and Theories of Anthropological Genetics* or *Methods and Theories of Genetic Anthropology*. We eventually agreed on the use of Anthropological Genetics because it connoted a commingling of anthropology and genetics, yet this title suggested a unique anthropological approach to the field of genetics. This volume was published in 1973 and it comprised the first attempt by multiple authors of synthesizing this field. I later discovered that D. F. Roberts had preceded us in referring to anthropological genetics in a lecture that he had given to the Royal Anthropological Society (Roberts, 1965).

From 1973 to the 1980s, there was considerable research activity in anthropological genetics and related fields. The most significant developments were in the applications of quantitative genetic methodologies to complex phenotypes, particularly in chronic diseases. Developments in computer technology and programming facilitated the use of linkage methods, path analytical approaches of Sewall Wright, and segregation analyses to complex phenotypes. These methodologies provided information as to the mode of transmission of a complex phenotype and the chromosomal mapping (through linkage analysis) of Mendelian traits. These new developments, punctuated by a pronouncement from Newton Morton that all of the major questions in population structure have been answered and we should instead refocus on genetic epidemiology, prompted me to consider an update of the 1973 volume. After I was awarded a National Institutes of Health Career Development Award in 1976 a portion of my university salary was released by the administration of the University of Kansas. This award freed funds for a lecture series by distinguished speakers, each coming to Lawrence for one week, providing a public lecture, interacting with faculty and graduate students, and presenting one seminar to the graduate students and faculty. It was at this time that James H. Mielke (a former student of Peter Workman) was added to the faculty at Kansas and he joined me in developing this lecture programme. This collaboration resulted in the first volume of a three-volume series, *Current Developments in Anthropological Genetics: Theory and Methods*, published in 1980 by Plenum Press. Volume 2 focused on the effects of ecology on population structure and was released in 1982 (Crawford and Mielke, 1982). That volume contained a number of innovative approaches to population structure, including Robert Sokal's initial application of spatial autocorrelation to human populations of the Solomon Islands. In 1984, the final volume in the series was published. It was based on my research in Belize, Guatemala and St Vincent Island and was used as a case study applying the theories of population genetics to a series of historically related populations of the Caribbean and Central America (Crawford, 1984). This volume documented an evolutionary 'success story' of the Garifuna (Black Caribs). Although no unadmixed Carib or Arawak Native Americans now remain on St Vincent Island, their genes have been dispersed over a wide geographic expanse on the coast