Session 12.

3

An Application of DNA Sequencing to a Human Rights Problem

Mary-Claire King
School of Public Health
University of California at Berkeley
Berkeley, California

I. INTRODUCTION: THE HISTORICAL CONTEXT

Between March 1976 and December 1983, the Republic of Argentina was controlled by a military junta responsible for the abduction, torture, and murder of thousands of citizens (Amnesty International, 1980; Interamerican Commission on Human Rights, 1980). The total number of victims of extrajudicial execution will probably never be known, but in December 1983, the newly elected democratic government established the National Commission on the Disappearance of Persons (CONADEP), with the request that this commission document, insofar as possible, the history of disappearances of citizens during the period of military rule. By September 1984, CONADEP had prepared evidence on 8800 victims; however CONADEP and human rights organizations believe the actual number of "disappeared" persons to be much higher (National Commission on the Disappearance of Persons, 1984).

Approaches from human genetics have been used to help identify a special subset of these victims: the 210 children who were kidnapped at birth or as infants by military and police who murdered their parents and retained or sold the children. Unlike the older victims of the "Dirty War," many of these very young children remained alive, but were made to "disappear" (Abuelas de Plaza de Mayo, 1985). In 1977, during the period of military rule, the surviving relatives of kidnapped children formed the Grandmothers of the Plaza de Mayo, a human rights group devoted to finding kidnapped children and reuniting families. Through persistent collection and follow-up of circumstantial evidence, the

Molecular Genetic Medicine, Vol. 1

Grandmothers began to locate their kidnapped grandchildren, primarily in the households of military and police officials and their collaborators (Nosiglia, 1985).

It soon became apparent to the Grandmothers that it was necessary, but not sufficient, to establish that a specific child was a kidnap victim. It was also necessary to establish each child's true identity by objective means. After the fall of the military and the election of a democratic government, it became possible to bring charges in Argentinian courts against the kidnappers. The Grandmothers therefore asked geneticists for help in establishing objectively the identities of these grandchildren. In June 1984, at the request of the Grandmothers of the Plaza de Mayo and CONADEP, a commission from the American Association for the Advancement of Science first traveled to Buenos Aires to help develop procedures for genetic identification of kidnapped children.

In 1984, genetic relationships were most effectively established using serological testing of human lymphocyte antigens (HLAs) at the A, B, C, and DR loci (Albert *et al.*, 1984). The immunogenetics laboratory of Ana Maria DiLonardo at the Durand Hospital in Buenos Aires, which was experienced in HLA serologic typing, agreed to undertake the identification project. That laboratory has been successful in establishing the identities of a large number of children and in presenting evidence to the Argentinian courts that has led to the reunification of families (DiLonardo *et al.*, 1984; Diamond, 1987).

In the intervening 7 years, the genetic approaches applied to establishing the identity of these kidnapped children have been extended. Two developments motivated this work. First, at the request of the Grandmothers and other human rights organizations, the government of Argentina established in 1985 the National Genetic Data Bank. This voluntary service offers grandparents, aunts, uncles, cousins, and other surviving relatives of disappeared children the opportunity to have their blood sampled, HLA serology determined, and white cells stored for extraction of DNA. Pedigrees documenting the relationships of surviving relatives to the murdered parents and missing children are also constructed. The intent of the legislation is that, as kidnap victims are discovered, their identities can be determined by matching the child to surviving relatives. Several hundred surviving relatives have contributed to the Data Bank in the past 7 years. Consequently, as kidnap victims are found for whom little circumstantial evidence of identity exists, each must be matched against multiple families. To establish identity with a sufficiently high level of statistical certainty under these circumstances usually requires matching to more than one genetic linkage group.

The second development during the project has been the appearance of families with few surviving relatives. Deaths from natural causes and by murder during the military period and fear have led to incomplete information for some families. It is not uncommon for a family searching for a missing child to include only the maternal lineage, sometimes only the mother or a sister or brother of the murdered mother of a kidnap victim. Therefore, it has been necessary to employ

grandparent and his or her grandchild have only a 50% chance of sharing an allele by descent at a single nuclear locus. This limitation can be overcome by typing multiple independent loci, gradually accumulating odds in favor of, or against, relationship. A further complication is that differences between fragment lengths defining different VNTR alleles may be very small, so that it may not be clear whether two bands are identical by descent or not. This limitation can be overcome by mixing aliquots of DNA from two samples whose identity is to be tested, to determine whether they migrate identically. In addition, one can relax the effort to resolve so many alleles at these loci, although this reduces statistical power.

PCR amplification followed by hybridization to allele-specific oligonucleotide probes which reveal dots on filter paper is a relatively recent technique (Saiki *et al.*, 1986; Saiki *et al.*, 1988). The principal advantage of this technique is that PCR amplification permits identification based on extremely small amounts of material. However, for the determination of relatedness of living persons, the quantity of DNA available is not a limiting factor. At least three limitations of the "dot-blot" approach greatly reduce its effectiveness for identification of families. First, as of now, very few loci can be evaluated using allele-specific oligonucleotides, and these loci are not highly informative. Of course, sets of oligonucleotides for other loci are currently under development. Second, a genotype is defined by dots on filter paper made by probing the amplified sample DNA with the oligonucleotide. These dots generally vary in intensity, so that determining whether an allele is present or not may be subjective and the result ambiguous. Third, contamination of the PCR reaction by external material can lead to amplifying the target sequence from the wrong individual. Present efforts to circumvent contamination include keeping samples to be amplified in closed and controlled environments and carrying out the amplification step at least twice, testing the results independently.

Nevertheless, the future is bright both for DNA fingerprinting and for allele-specific oligonucleotides as tools for identification. The problems are well-known to everyone in the field, and a number of groups are actively addressing their solution. In addition, several professional and advisory organizations are cooperating to establish standards for the application of these approaches.

Against the background of increasingly widespread adoption of genetic methods of testing identity and of development of DNA sequencing technology, it is highly probable that direct determination of DNA sequences will soon be introduced to test identity and relationship. There are major advantages to DNA sequencing as a tool. Enzymatic amplification by PCR of target sequences enables tiny quantities of human material to be used. Furthermore, the genetic sequence provides the ultimate resolution of identity: properly chosen sequences would uniquely identify each individual. Finally, sequencing will circumvent many of the technical problems associated with DNA fingerprinting or with allele-specific oligonucleotides and dot blots.

polymerase after each heat denaturation step (Gelfand and White, 1990; Innis *et al.*, 1988; Saiki *et al.*, 1988). This modification not only has greatly simplified amplification and allowed it to be automated, but also has increased the yield, specificity, and length of the DNA fragments that can be amplified (Innis and Gelfand, 1990). Direct sequencing of enzymatically amplified mtDNA has been applied to the study of the evolution and diversity of human populations (Engelke *et al.*, 1988; Kocher *et al.*, 1989; Paabo, 1989; Paabo *et al.*, 1989; Vigilant *et al.*, 1988; Wrischnik *et al.*, 1987).

The control region of mtDNA, near the origin of replication, is particularly diverse among individuals. The mtDNA control region (also called the displacement or D-loop) is a region of approximately 1200 base pairs flanked by tRNA(Pro) and tRNA(Phe) genes. The control region does not code for any genes, which perhaps has released it from strict neocleotide conservation. In the control region, blocks of sequence homology among human, bovine, and rat sequences are separated by regions of high sequence diversity (Anderson *et al.*, 1981).

The method for amplifying and sequencing the control region of mtDNA in use for this project follows a two-step protocol that minimizes unintended amplification products (Gyllensten, 1989; Saiki *et al.*, 1988). The first stage is amplification of the entire mtDNA control region using primers that are complementary to the highly conserved flanking tRNA(Thr) and tRNA(Phe) genes (Kocher *et al.*, 1989; Orrego and King, 1990). The first step of the PCR reaction is carried out with equal concentration of primers. The second amplification step is carried out with nested primers using the unbalanced priming method (Mullis *et al.*, 1986; Gyllensten and Erlich, 1988). The single strand product is suitable for sequencing by the dideoxynucleotide chain termination technique (Sanger *et al.*, 1977). The region targeted in the second stage of the amplification is the most variable portion of the mtDNA control region, the first 400 basepairs. The asymmetric PCR amplification is carried out as in the initial amplification reaction, except for a 50-fold reduction of one of the primers. After the asymmetric PCR, the concentrated reaction mixture is used for sequencing with the Sequenase kit (U.S. Biochemical).

Several technical issues deserve further attention. Most important is the elimination of any sequencing artifacts due to replication errors of *Taq* polymerase (Scharf *et al.*, 1986; Tindall and Kunkel, 1988). Reaction conditions can be optimized to reduce errors during amplification. Parameters such as time and temperature of denaturation, annealing, and extension steps; the number of PCR cycles; the concentrations of salts, deoxyribonucleotide triphosphates, primers, and *Taq* polymerase all affect the fidelity of the reaction (Gelfand and White, 1990; Keohavong and Thilly, 1989; Saikai *et al.*, 1988). Recent experiments indicate that minimizing the concentrations of magnesium ions and of *Taq* polym-

sites compared, 32 were variable. When the sequences were compared between individuals in pairs, the number of differences ranged from 1 nucleotide to 13 nucleotides; the average pairwise difference was 5.9 nucleotides.

Obviously, mtDNA sequences will serve to identify the maternal lineage of an individual only to the extent that maternal lineages differ in sequences in the target region. Therefore it was crucial to be able to determine the likelihood that unrelated individuals would be identical by chance. For the 91 comparisons of the 14 sequences selected two at a time, the observed and expected number of sequence differences between pairs of individuals closely approximated a Poisson distribution (Orrego and King, 1990). Therefore, based only on this small sample of sequences, the probability of zero differences between two unrelated individuals is $p(0) = e^{-x}$, where x is the average number of differences between pairs of sequences. For $x = 5.9$, $p(0) = 0.0027$; in other words, for this 347 basepair sequence, the probability that two unrelated individuals will be identical by chance is approximately 1 in 370.

Two other sets of samples from Caucasian populations have been sequenced for the same region of mtDNA (DiRienzo and Wilson, 1991). Sixty-nine persons were sampled from five linguistically-defined regions of Sardinia. The linguistic subgroups of Sardinia appear genetically homogeneous based on nuclear genes and marriage patterns. This homogeneity was reflected in mtDNA sequences. For the 400-basepair region screened, nine sequences appeared more than once, including one sequence shared by 15 of the 69 subjects! Because samples were collected by requesting placentas from women giving birth in maternity hospitals, it is possible that the same extended family might have been sampled more than once. If so, this degree of concordance may be higher than for nominally unrelated Sardinians. In any case, the Sardinian sample illustrates how individual mtDNA sequences may appear multiple times in an endogamous population, and hence the importance of understanding relationships among maternal lineages in a population in which mtDNA sequencing will be employed.

A sample of 42 persons from the Middle East, including 29 Bedouins from Saudi Arabia, 8 Arabs from Israel, and 5 Jews from Yemen, were sequenced for the same 400-basepair region (DiRienzo and Wilson, 1991). In this sample, three sequences appeared more than once, each in a pair of individuals. Again, maternal relatives were not explicitly excluded. For the Middle Eastern group as a whole, the average number of nucleotide differences per pair of sequences was 7.3, and the distribution of pairwise differences was approximately Poisson.

The population genetics of mtDNA sequences involves determining the size of an mtDNA deme; that is, the historic and evolutionary relationships among maternal lineages. Furthermore, the appearance of individual sequences held in common by multiple individuals indicates that discriminating power will depend on the specific sequence, as well as on the genetic substructure of a population. For the genetically heterogeneous, European population of Argentina, with a finite

Teeth of relatively recent decedents are an excellent source of mtDNA. The nonliving exterior matrix of the tooth surrounds and protects a soft interior pulp consisting of living cells. The relative isolation of the dental pulp from the external environment protects its DNA from biological degradation. In addition, hydroxyapatite, a major component of the tooth, strongly binds and hence stabilizes DNA. In the unique environment of the tooth pulp chamber, DNA may be extremely long-lived relative to DNA in other cellular components.

As a preliminary test of sequencing DNA from teeth, we isolated DNA from the twenty-year-old baby teeth of an adult. Adequate DNA was obtained by amplification, and a portion of the mtDNA control region was sequenced. The sequence obtained matched that derived from a fresh blood sample of the adult, his sister, and his mother (Gintherc, unpublished results).

VI. TESTING RELATIONSHIPS IN ARGENTINA

The first application of mtDNA sequencing to the human rights work in Argentina was the analysis of the Lopez Guerra, Belaustegui Herrera, and Weisberg family, whose pedigree is shown as Figure 3.2. This extended family includes three Argentinian families related by marriage. In 1977, five of the young adults in this kindred were kidnapped and subsequently murdered by military forces in Argentina. Both Maria Cristina Lopez Guerra (C) and Valeria Belaustegui Herrera (V) were in the first trimester of pregnancies when they were abducted. Their parents later heard that their daughters had been kept alive until they gave birth, their newborn infants taken, and the young women then murdered.

In 1988, ten years after the births would have occurred, a 10-year-old boy (A,9) was brought to the attention of the Grandmothers. The woman who introduced A to the Grandmothers had cared for him since he was given to her as a "present" when a newborn in 1978 by her then male companion, who had close ties to the military. She reported that she had been terrified to tell anyone about the incident until her former companion was arrested in 1988 by the civilian

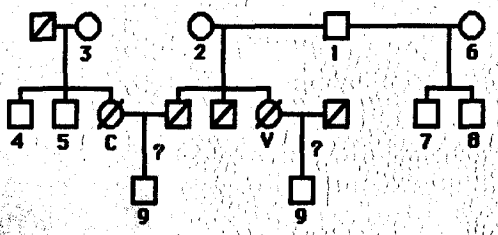

Figure 3.2. The Lopez Guerra, Belaustegui Herrera, and Weisberg families from Argentina.

The major concern underlying this work is to do what is best for the children (Abuelas de Plaza de Mayo, 1984). Clearly, circumstances vary enormously and the custody of each child must be decided individually. A few of the children born in captivity or kidnapped as infants were adopted in good faith by families with no ties to the military. The resolutions of these exceptional cases have generally been amicable, with the children being told the truth about their biological parents and spending time with both their biological and adoptive families. However, the cases of children living with military or police officers involved in the torture and murder of their parents are far more difficult. These have comprised the vast majority of children discovered so far. Certainly under normal circumstances, a child would not be left with kidnappers or their accomplices regardless of his or her age at abduction. The notion of assessing whether persons involved in kidnapping, torture, or murder are suitable parents for the children of their victims appears highly unlikely. Kidnapping has universally been considered a crime. Is the situation different in Argentina because kidnapping occurred on a large scale? The human rights groups with whom we work suggest that to abandon the search for the kidnapped children of Argentina is to abandon a group of children who will not grow up in carefree innocence. As these children become adults, what would their attitudes be toward relatives who knew they had disappeared but did nothing? What would be the effect on a young person to learn he or she has lived with people involved in the murders of his/her parents and that his/her surviving relatives did nothing to find him/her? Would failing to attempt to identify the kidnapped children implicitly grant immunity to kidnappers? Would this increase the sense of invulnerability of abusers of human rights in other countries?

The historical situation that led to this application of genetics to human rights is unprecedented. Thus, answers to these questions of ethics, law, and mental health are developing with our current experience. Meanwhile, as of this time (early 1991), the political situation in Argentina is much more hostile to the Grandmothers' efforts than during the Alfonsin presidency of 1984–1989. In particular, it is increasingly difficult to work within the Argentinian judicial system. However, the Grandmothers remain undaunted. They point out that the average age of the kidnapped children is now 15 years. Very soon, these children will have the legal right to determine for themselves their identities. For this purpose mtDNA sequences will be available. Even though the grandparents of a kidnapping victim may die before the grandchild is found, the young adult's maternal lineage will be identifiable using the genetic information the Grandmothers have left behind. A young person can thereby be put in touch with his/her family—surviving aunts, uncles, and cousins—and his/her history. For the past 15 years, the Grandmothers have been searching for their kidnapped grandchildren. Very soon, these grandchildren will come looking for them.

Giles, R. E., Blanc, H., Cann, H. M., and Wallace, D. C. (1980). Maternal inheritance of human mitochondrial DNA. *Proc. Natl. Acad. Sci. U.S.A.* **77**:6715–6719.

Greenberg, B. D., Newbold, J. E., and Sugino, A. (1983). Intraspecific nucleotide sequence variability surrounding the origin of replication in human mitochondrial DNA *Gene* **21**:33–49.

Gyllensten, U. B. (1989). PCR and DNA sequencing. *BioTechniques* **7**:700–708.

Gyllensten, U. B., and Erlich, H. A. (1988). Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. *Proc. Natl. Acad. Sci. U.S.A.* **85**:7652–7656.

Gyllensten, U., Wharton, D., and Wilson, A. (1985). Maternal inheritance of mitochondrial DNA during backcrossing of two species of mice. *J. Hered.* **76**:321–324.

Hagelberg, E., Sykes, B., and Hedges, R. (1989). Ancient bone amplified. *Nature* **342**:485.

Hecht, N. B., Liem, H., Kleene, K. C., Distel, R. J., and Ho, S.-M. (1984). Maternal inheritance of the mouse mitochondrial genome is not mediated by a loss or gross alteration of the paternal mitochondrial DNA or by methylation of the oocyte mitochondrial DNA. *Dev. Biol.* **102**:452–461.

Higuchi, R., vonBeroldingen, C. H., Sensabaugh, G. F., and Erlich, H. A. (1988). DNA typing from single hairs. *Nature* **332**:543–546.

Innis, M. A., and Gelfand, D. H. (1990). Optimization of PCRs. In "PCR Protocols: A Guide to Methods and Applications" (M. Innis, D. Gelfand, J. Sninsky, and T. White, eds.), pp. 3–12. Academic Press, San Diego.

Innis, M. A., Myambo, K. G., Gelfand, D. H., and Brow, M. A. D. (1988). DNA sequencing with *Thermus aquaticus* DNA polymerase: Direct sequencing of PCR-amplified DNA. *Proc. Natl. Acad. Sci.* **85**:9436–9440.

Interamerican Commission on Human Rights. (1980). "Report on the Situation of Human Rights in Argentina." General Secretariat, Organization of American States, Washington, C.C.

Jeffreys, A. J., Brookfield, J. F. Y., Semeonoff, R. (1985a). Positive identification of an immigration test-case using human DNA fingerprints. *Nature* **317**:818–819.

Jeffreys, A. J., Wilson, V., and Thein, S. L. (1985b). Individual-specific "fingerprints" of human DNA. *Nature* **316**:76–79.

Keohavong, P., and Thilly, W. G. (1989). Fidelity of DNA polymerases in DNA amplification. *Proc. Natl. Acad. Sci. U.S.A.* **86**:9253-9257.

Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Paabo, S., Villablanca, F. X., and Wilson, A. C. (1989). Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. *Proc. Natl. Acad. Sci. U.S.A.* **86**:6196–6200.

Mullis, K. B., and Faloona, F. A. (1987). Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Methods Enzymol.* **155**:335–350.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., and Erlich, H. (1986). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb. Symp. Quant. Biol.* **51**:263–273.

Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E., White, R. (1987). Variable number of tandem repeat (VNTR) markers for human gene mapping. *Science* **235**:1616–1622.

National Commission on the Disappearance of Persons. (1984) "Nunca Mas." Editorial Universitaria de Buenos Aires. English translation: 1986, Faber and Faber Ltd. London.

Nosiglia, J. E. (1985). "Botin de Guerra." Cooperativa Tierra Fertil, Buenos Aires.

Orrego, C., and King, M.-C. (1990). Determination of familial relationships. In "PCR Protocols: A Guide to Methods and Applications." (M. Innis, D. Gelfand, J. Sninsky, and T. White, eds), pp. 416–426. Academic Press, San Diego.

Paabo, S. (1985). Molecular cloning of ancient Egyptian mummy DNA. *Nature* **314**:644–645.

Paabo, S. (1986). Molecular genetic investigations of ancient human remains. *Cold Spring Harbor Symp. Quant. Biol.* **51**:441–446.