DNA FINGERPRINTING FOR THE DIAGNOSIS OF DISEASES AND CRIMES

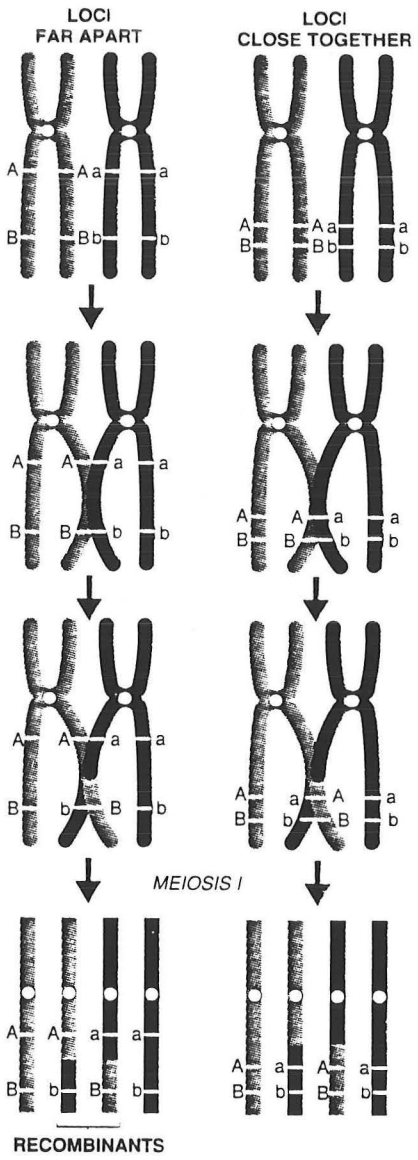
Helen H. Hobbs, M.D.
Internal Medicine Grand Rounds
November 8, 1990

In 1985 a new class of highly informative genetic markers was identified (Jeffreys et al., 1985a). These DNA genetic markers called variable number of tandem repeats (VNTRs), minisatellites, or DNA fingerprints are used to diagnose genetic disorders, map genes responsible for disease, determine parenthood, and solve crimes. To appreciate their contribution to the study of genetics requires a brief review of the history of genetic markers.

The development of recombinant DNA technology has dramatically improved the capability to diagnose genetic diseases. Gene mutations can now be detected directly using various well established methodologies. For diseases such as sickle cell anemia where every affected individual has the same mutation in the same gene, a diagnosis can be made in all cases by detecting the mutation directly. However, most genetic diseases can be caused by a myriad of different gene defects. Diseases like phenylketonuria, osteogenesis imperfecta, familial hypercholesterolemia, cystic fibrosis, hemophilia and neurofibromatosis can all be caused by a variety of different DNA mutations. It is impossible to know a priori which gene mutation is causative in any given affected individual. To perform prenatal diagnosis for a family with one of these disorders would require that the gene be analyzed in detail to identify the particular DNA mutation responsible for the disease. With new technological advancements, it has become much less cumbersome to find mutations in genes, but it is still a heroic amount of work to identify the gene mutation in each new patient.

An alternative strategy is to analyze the genotype indirectly using genetic markers linked to the mutant gene. Genetic markers are readily recognizable, simply inherited, genetic traits that allow one to indirectly examine the genotype. Genetic markers can be used to follow the segregation of closely linked genes during meiosis. Prior to metaphase I of meiosis, the maternal and paternal chromosomes replicate, and the two homologous chromosomes undergo recombination [Figure 1]. The frequency of recombination events between two loci is proportional to the physical distance that separate them. Thus, if two genes are located close together on a chromosome, it is likely they will remain linked through the two divisions of meiosis. Alternatively, if two genes are separated by a large distance, it is unlikely that they will remain together on the same chromosome through meiosis [Figure 1]. It has been estimated that there is a 1% chance of recombination between two loci if they are separated by 1 centimorgan (1000 kilobases). Two loci that are at opposite ends of a chromosome have only a 50% chance of remaining together after meiosis.

Figure 1 Recombination Between Homologous Chromosomes



Principles of Medical Genetics, T.D. Gelehrter & F.S. Collins, 1989

For two closely linked genes, information gleaned about the inheritance of one gene can be used to predict the inheritance of the other gene. In medical genetics, one of the genes acts as a genetic marker for the disease-causing gene. For a genetic marker to be informative for the mutant gene, the individual must be heterozygous for the genetic marker. Luckily, for the geneticist, there are polymorphisms in the sequences of proteins and genes that can be employed as genetic markers.

A polymorphism is defined as the occurrence of two or more genetically determined alternative alleles in a population such that the frequency of the rarest allele could not be maintained by the mutation rate alone. An operational definition is that the rare allele must have a frequency in the population of at least 1%. The types of polymorphisms that have been used as genetic markers include cytogenetic polymorphisms, protein polymorphisms, and DNA

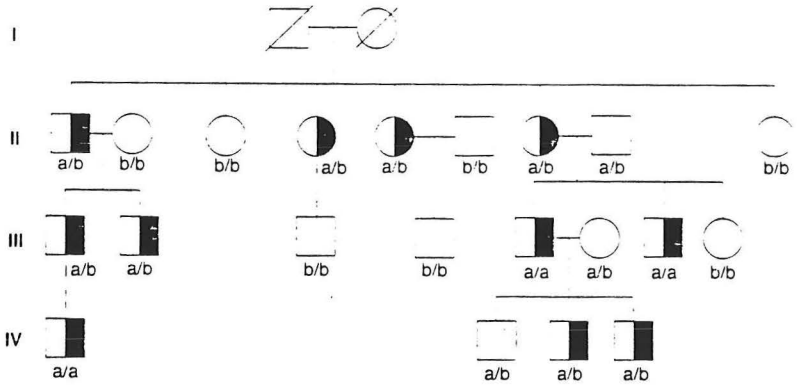
polymorphisms.

The first gene was mapped to an autosome using a cytogenetic polymorphism. Donahue was examining his own chromosomes by karyotyping and noted he was heterozygous for a region of expanded heterochromatin (condensed DNA that is not functional) in chromosome 1 [Figure 2, Panel A]. This is a cytogenetic polymorphism which has no known clinical sequelae. He also typed his blood and determined that he was heterozygous for the Duffy blood group (the Duffy blood groups are denoted by a and b in Figure 2, Panel B). He performed karyotype analysis and blood typing on his family. He found that the Duffy blood group co-segregated invariably with the "polymorphic" chromosome 1 in the family [Figure 2, Panel B]. That is, everyone who inherited the chromosome 1 with the expanded heterochromatin also had the same Duffy blood group. He concluded that this was not a coincidence but occurred because the gene encoding the Duffy blood group was on chromosome 1. The expanded heterochromatin acted as a genetic marker for Duffy blood group. The Duffy blood group is <u>not</u> encoded by the expanded heterochromatin, but the gene is closely linked to it.

<u>Figure 2</u> <u>Gene Mapping of Duffy Blood Group to Chromosome 1</u>

A. Karyotype

B. Pedigree of Donahue Family



Cytogenetic polymorphisms are only of limited usefulness as genetic markers. For many years protein polymorphisms were the most commonly employed genetic markers. Amino acid substitutions that result in an alteration of charge can be detected by gel electrophoresis. Alternatively, immunological assays have been used to detect protein polymorphisms.

Occasionally protein polymorphisms have been used as genetic markers to diagnose disease. An example is the ABO blood group which was used as a genetic marker for the autosomal dominant disorder, nail-patella syndrome. Individuals with nail-patella syndrome have dysplastic nails, absent, or hypoplastic patellae, and, in some cases, an associated nephropathy. The gene causing

this disorder is not known, but by analyzing numerous pedigrees it was shown that the disease cosegregated with the ABO blood group (Renwick and Lawler, 1985). A pedigree with nail-patella syndrome is shown in Figure 3. Individual II-1 has blood type AB and he has had seven children with a woman of blood type O. Since blood type O is recessive, we know that both of her alleles encode blood type O. The blood type of the children is determined by which allele each child inherited from the father. Every child with blood type A was found to have nail-patella syndrome. All offspring with blood type B were unaffected. Based on these observations, a schematic map of the involved chromosomes can be made and is shown in Figure 4. Individual II 1. is heterozygous at both the ABO blood group and nail-patella gene locus, and for this reason analysis of the ABO blood group can be successfully used to predict the inheritance of the normal or mutant gene. In this family blood group A is <u>linked</u> to the mutant gene whereas in another family the mutant gene may be linked to blood group B. Each pedigree must be analyzed individually to determine the blood group allele that co-segregates with the mutant gene at the nail-patella locus.

One problem using ABO blood groups as genetic markers for nail-patella syndrome is that the two loci are separated by about 10 centimorgans. This was determined by analyzing numerous families and finding that in approximately 10% of the offspring there was a recombinational event. In the family shown in Figure 3 and 4, a recombinational event between the loci would result in an offspring with blood group B and nail-patella syndrome or blood group A and no disease. Hence, the utility of such a marker is related to the distance between the marker and the gene of interest.

<u>Figure 3</u> <u>Segregation Analysis of ABO Blood Group in a Family with Nail-Patella Syndrome</u>

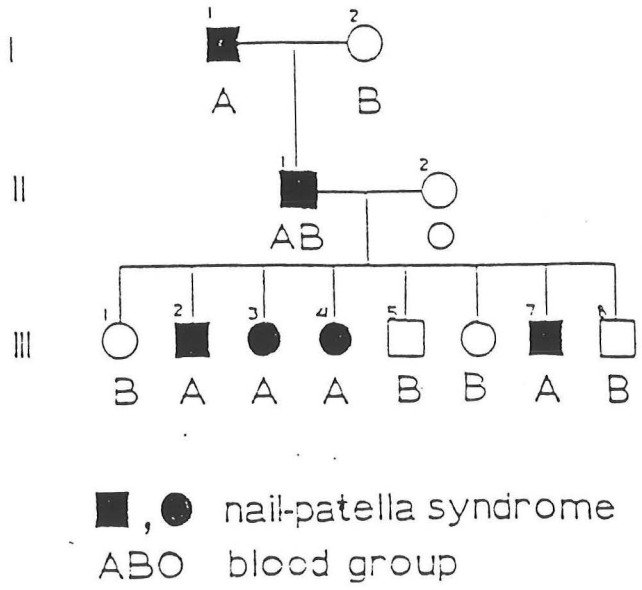
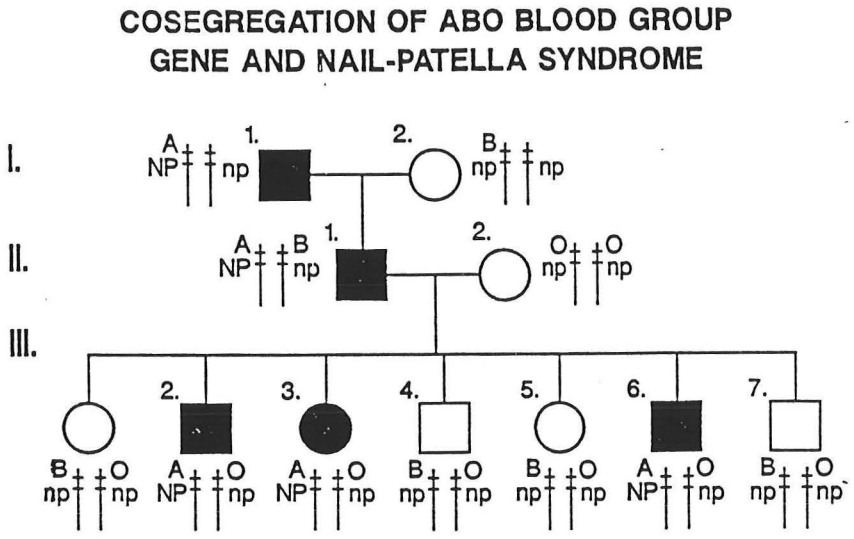


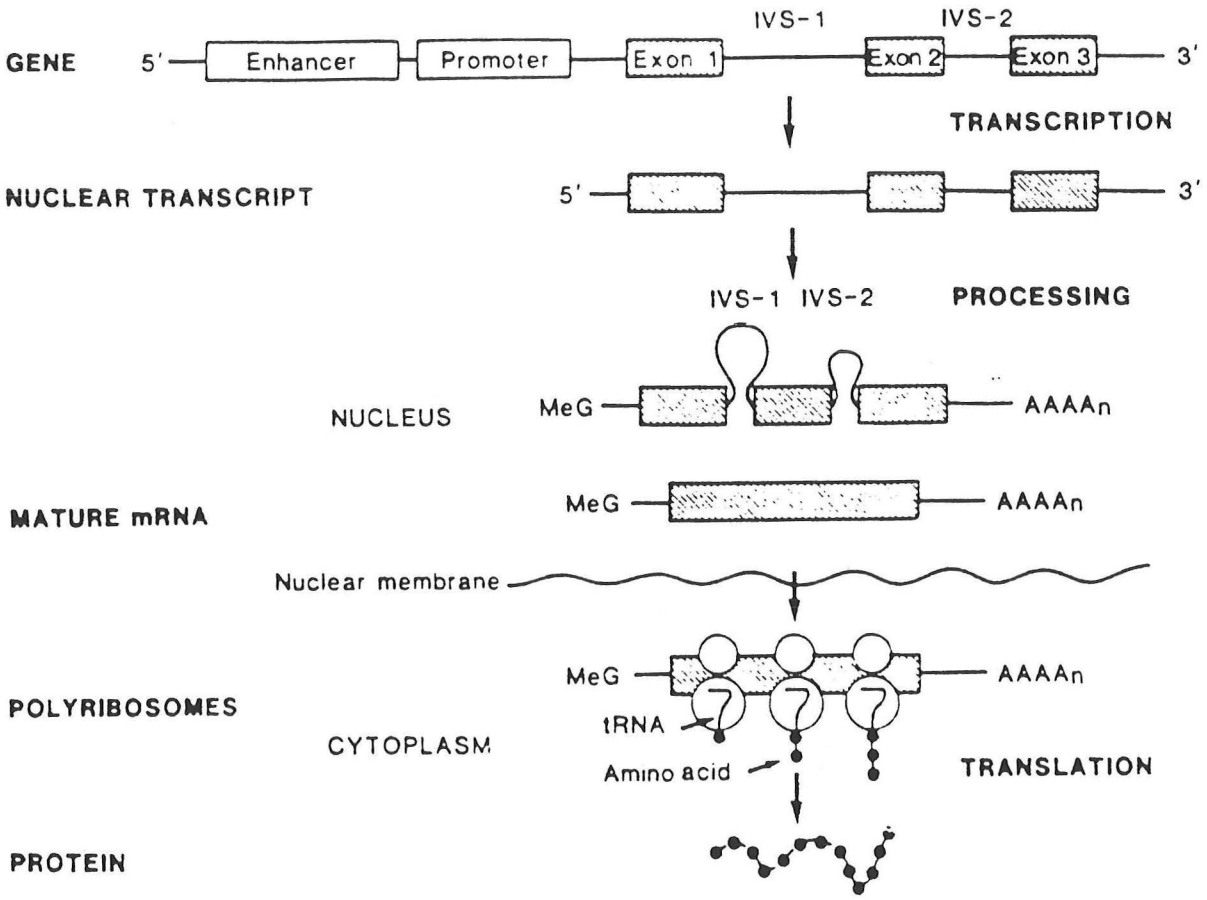
Figure 4



Protein polymorphisms have rarely been successfully used as genetic markers to diagnose genetic disease. The major reasons are that 1) proteins are not polymorphic enough (since protein sequences are under selective pressure) and 2) there are not a sufficient number of protein polymorphisms to find one encoded by a gene closely linked to the gene of interest. The analysis of the ABO blood group in families with nail-patella syndrome exemplifies this problem. The ABO blood group gene is too far away from the gene causing nail-patella syndrome. The chance of recombination between the two loci is 10% which is too great to use for prenatal diagnosis.

An ideal genetic marker would be easy to detect, highly polymorphic, and located close to the gene causing the disease. Only 5% of the human genome (3×10^9 base pairs) encodes protein. The coding regions of DNA (exons) are separated by introns and flanking regions [Figure 5]). Some sequences within the introns and flanking regions of genes, such as those involved in gene transcription, gene regulation, and mRNA splicing are under selective pressure. However, the vast majority of noncoding sequences are under no (known) selective pressure since changes in these sequences are phenotypically neutral. In the noncoding regions of genes, it has been estimated that there is a base pair polymorphism every 100-500 base pairs (Jeffreys, 1979).

Figure 5 Gene Structure



In 1978, it was recognized that these DNA polymorphisms could be used to follow the segregation of closely linked genes (Kan and Dozy, 1978). It was noted that some base pair substitutions destroy or create restriction sites, and therefore the polymorphism could be assayed by determining the integrity of the restriction site (so-called restriction fragment length polymorphisms). The methodology used to detect these silent polymorphisms and use them as genetic markers will be briefly reviewed.

First, total genomic DNA is isolated from a tissue source (most commonly white blood cells) and cut with a restriction enzyme. A restriction enzyme is an enzyme produced by bacteria that cuts specifically at a particular sequence, the consensus sequence, which is usually 4 or 6 base-pairs in length. Figure 6 gives a schematic representation of genomic DNA that has been cut with SmaI.

Figure 6

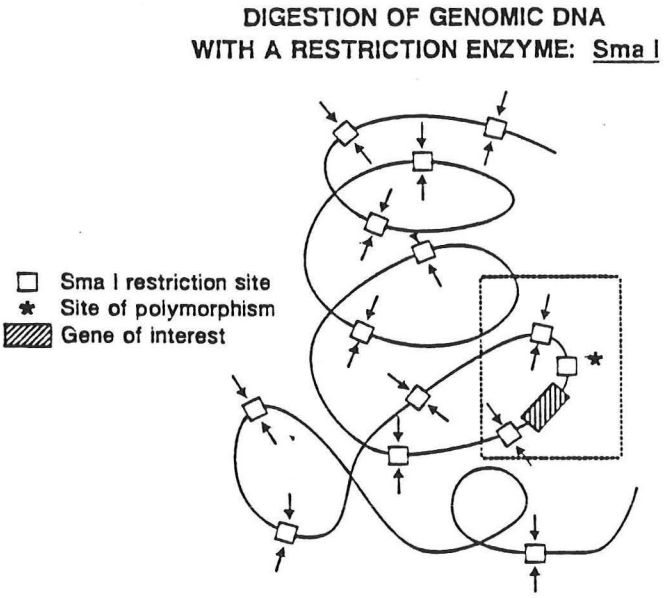
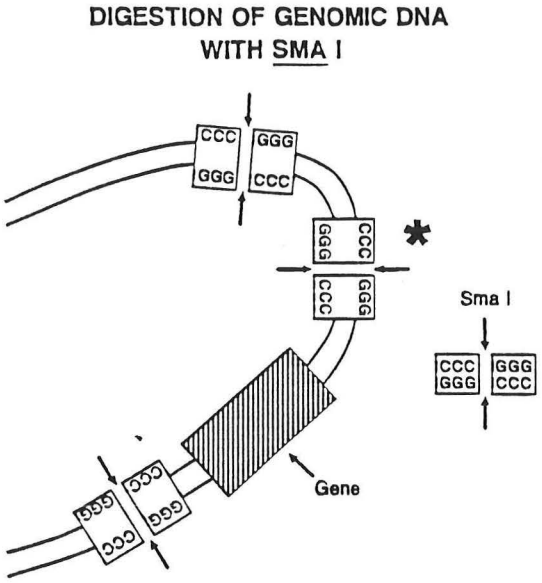
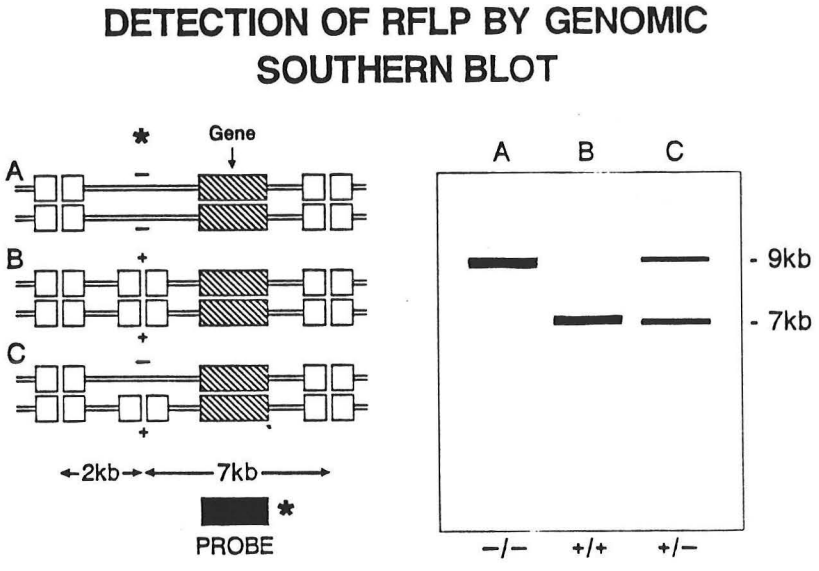


Figure 7



Two of the three SmaI restriction sites that surround the gene of interest [Figure 7] are not polymorphic; that is, every individual has both of these sites. However, one site, demarcated by the asterisk, is polymorphic; some alleles have a base pair substitution within the consensus sequence so that SmaI will not cut. Within the population there are three possible genotypes and they are shown on the left side in Figure 8. Panel A represents an individual homozygous for the basepair substitution that destroys the SmaI site. SmaI will not cut at this site on either chromosome. Panel B diagrams the two homologous alleles from an individual homozygous for the presence of the SmaI site. And finally, shown in panel C is a heterozygote; in one allele the SmaI site is present and in the other it is absent.

Figure 8

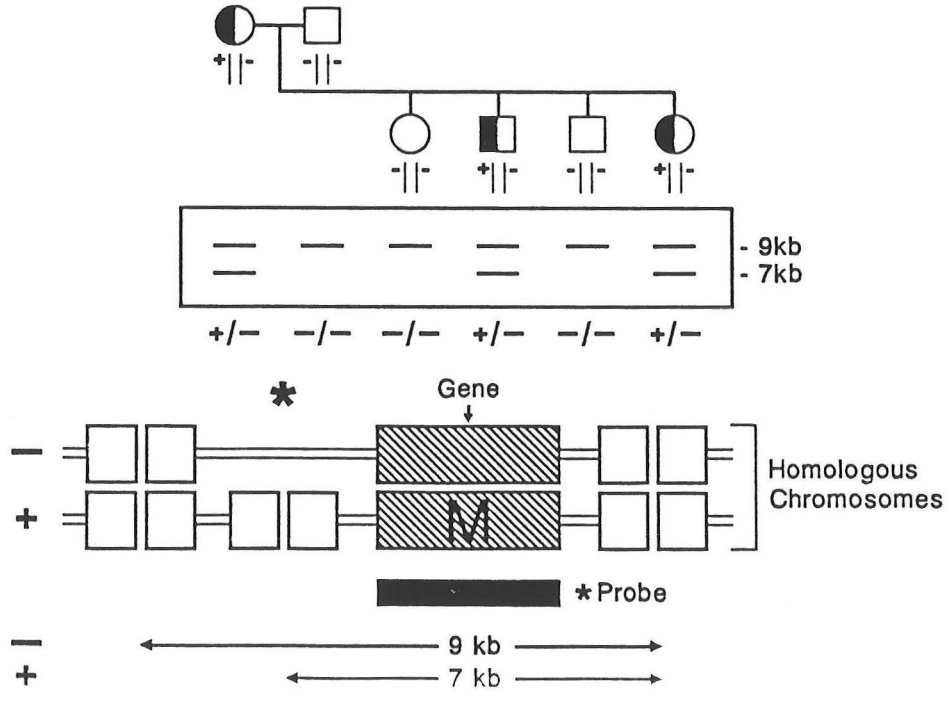


In order to detect this base pair polymorphism, genomic DNA from individuals A, B and C is cut with SmaI and the restriction fragments are size-fractionated by electrophoresis on an agarose gel. The DNA is transferred to a nylon membrane and then hybridized specifically with a radiolabeled probe derived from the gene (which has been cloned). The results are shown on the right side of Figure 8. In lane 1 there is only a single 9 kb restriction fragment since individual A does not have the polymorphic SmaI site in either allele. In lane 2, a smaller band of 7 kb is seen due to the presence of the SmaI site in both alleles. And finally, in lane 3, there is both a 9 kb and 7 kb band since individual C is heterozygous for the base-pair polymorphism at the SmaI site.

SmaI reveals a restriction fragment length polymorphism (RFLP). A RFLP is an inherited variation in the lengths of DNA fragments cut by a restriction enzyme due to the presence or absence of a particular consensus sequence. RFLPs also can be due to the presence or absence of a deletion or insertion between two conserved restriction sites, and this will be discussed later.

The RFLP revealed by SmaI can be used to analyze the segregation of a closely linked gene. Shown in Figure 9 are the results of an analysis of this restriction site in a family with an autosomal dominant disorder due to a mutation in a nearby gene. Genomic DNA from a father, mother, and four children has been digested with SmaI and analyzed by Southern blotting. The mother is heterozygous for the RFLP and she is also heterozygous for the mutation in the closely linked gene. She had children with a man who is homozygous for the absence of the restriction site. By analyzing the children for the presence or absence of the restriction site, it can be determined which gene each child got from their mother. The fact that the mother is heterozygous for the restriction site (and the father is homozygous) make the SmaI restriction site an informative marker for the linked gene.

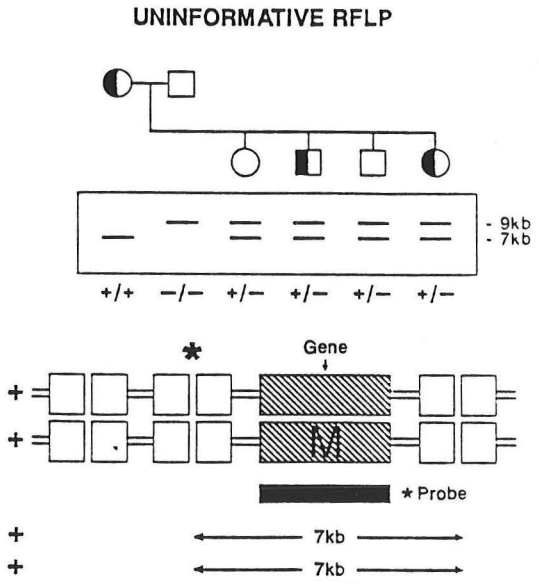
Figure 9



A schematic representation of the mother's two alleles is shown below the Southern blot in Figure 9; the mutation in the gene is on the allele that has the SmaI restriction site. This could not be known a priori but was determined by the analysis of her offspring for the presence, or absence, of the restriction site and the disease. The two offspring who inherited the allele with the restriction site (i.e., the 7 kb band) both have the disease, whereas the two offspring who inherited the 9 kb band are unaffected. Thus, in this pedigree, one can follow the segregation of the gene responsible for the disease by following the segregation of bands revealed by a SmaI restriction digest. The RFLP does not detect the mutation directly, but rather detects it indirectly because the polymorphic restriction site is linked to the gene with the mutation.

A polymorphic marker can be informative under two circumstances: 1) the affected parent must be heterozygous for the marker, and 2) the polymorphic marker must be located close to the gene causing the disease. An example of a closely linked RFLP not being informative in a pedigree is shown in Figure 10. In this family, the mother has an autosomal dominant genetic disorder. She is homozygous for the presence of the SmaI restriction site and has had children with a man homozygous for its absence. Therefore, the genotypes of the offspring are all identical (+/-) making it impossible to determine which children inherited the mutant allele. Both of the maternal alleles (the one with the mutation and one without the mutation) have the SmaI site present and thus, the segregation of each allele can not be ascertained.

Figure 10



Therefore, the best genetic markers are those which are 1) highly polymorphic and 2) closely linked to the gene of interest. The likelihood that a genetic marker will be informative is

proportional to its heterozygosity index (i.e. the percentage of individuals in the population who are heterozygous for the site). RFLPs can only have a heterozygosity index of 50% since there are just two possibilities - either the restriction site is present, or the restriction site is absent. If both alleles have a frequency of 50% in the population, it can be expected that the number of individuals who are heterozygous for the RFLP would be 50%. Multiple, linked, RFLPs can be used to construct haplotypes which can improve the heterozygosity index. However, to determine haplotypes requires the analysis of multiple different restriction sites in numerous family members.

Variable numbers of tandem repeats (VNTRs), minisatellites, or hypervariable regions are polymorphisms due to differences in length between conserved restriction sites, rather than the presence or absence of a restriction site. The length differences are due to the presence of allelic differences in the number of oligonucleotide repeats. These oligonucleotide repeats consist of sequences ranging in size between 9 and 64 base pairs that are tandemly (head-to-tail) arrayed [Figure 11 & 12]. The number of repeats varies greatly between alleles making these polymorphisms dramatically more informative than the previously described cytogenetic, protein, or DNA polymorphisms. It has been estimated that there are more than 1500 VNTR sequences in the human genome (Nakamura et al., 1987). The repeated sequences tend to be GC-rich and range in size from 100 base pairs to 20 kilobases.

Figure 11

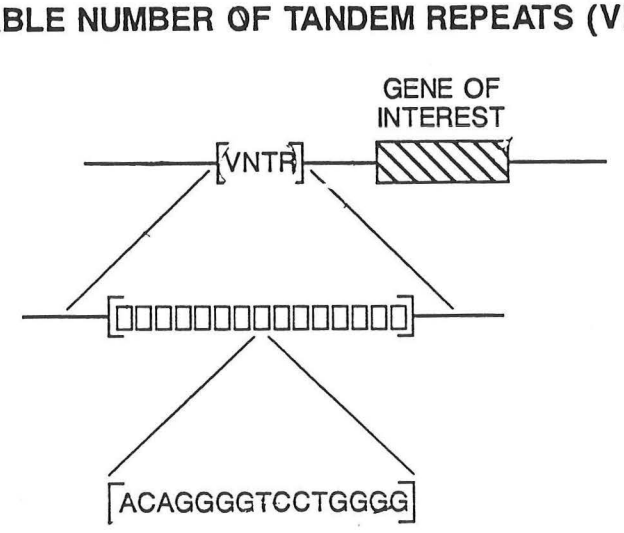
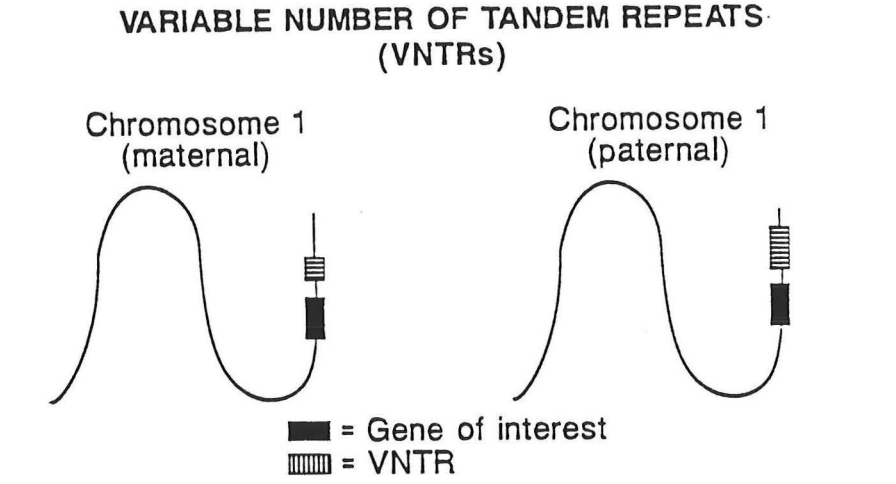


Figure 12



There are two methods used to detect VNTR alleles. One is by Southern blot analysis. Genomic DNA is restricted with an enzyme that does not cut within the repeat and the blot is hybridized with a unique sequence probe which is derived from sequences adjacent to the tandem repeat (a so-called locus-specific probe). The size of the resultant fragments reflects the number of tandem repeats between the flanking restriction sites [Figure 13]. The other method utilizes the polymerase chain reaction (PCR) to selectively amplify the repeated sequence (Jeffreys et al., 1988; Boerwinkle et al., 1989; Horn, 1989). Oligonucleotides homologous to sequences that flank the repeat are used in the amplification reaction [Figure 14]. To determine the length of the fragments, the amplified DNA fragments are size-fractionated by gel electrophoresis. The sizes of the bands are proportional to the number of tandem repeats.

Figure 13

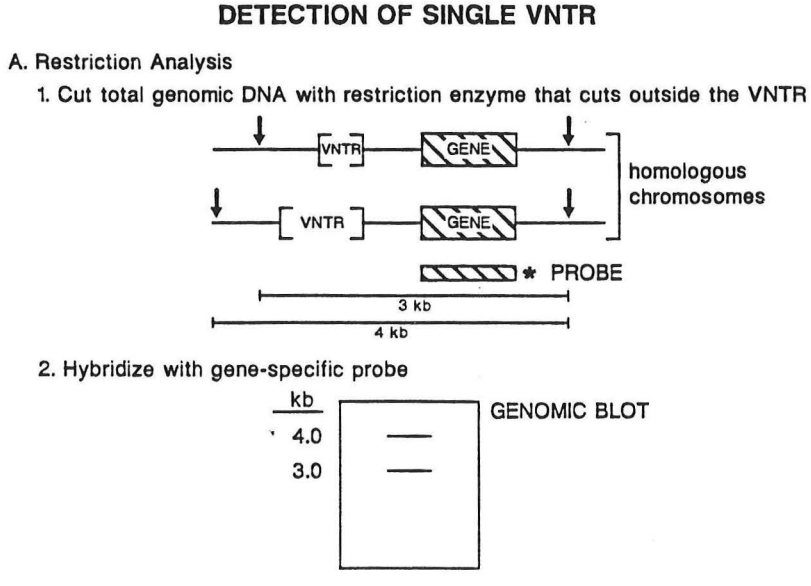
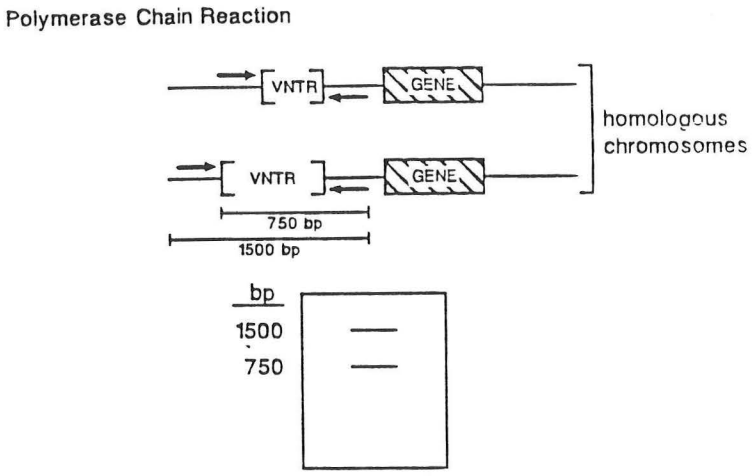
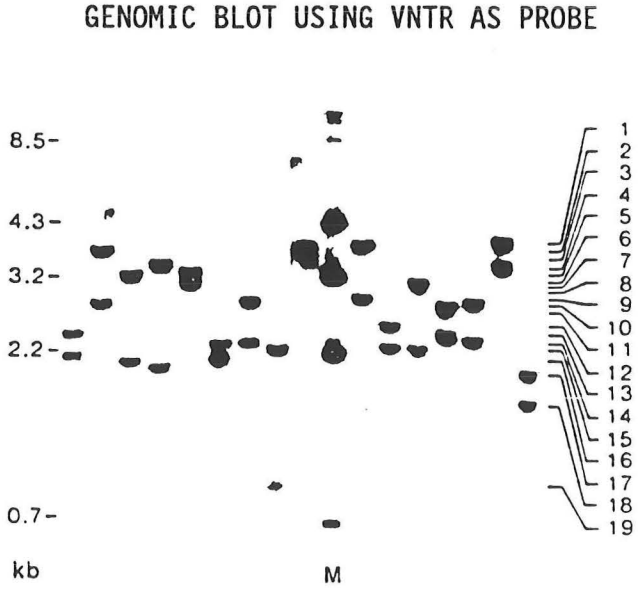


Figure 14



Since the number of repeats is highly polymorphic, and there can be many different alleles at one locus, the likelihood that an individual has a different number of repeats in each allele is much greater than that seen with a single RFLP. Figure 15 is a genomic blot of DNA from 16 unrelated individuals probed with a locus-specific VNTR locus probe (Nakamura et al., 1987). Nineteen different bands are seen and each individual is "heterozygotes" which reflects the high degree of polymorphism. Some other VNTR loci have been described that have as many as 70 alleles of different lengths (Wong et al., 1986). The heterozygosity indices for most VNTR loci range between 85% to over 99%.

Figure 15



An example of using the Southern blot technique to assay a VNTR locus in a pedigree is shown schematically in Figure 16A and an actual Southern blot analysis of a pedigree is shown in Figure 16B. Two bands are present for each family member and since all four parental alleles can be differentiated, it is possible to determine the parental origin of each allele in the offspring.

Figure 16A

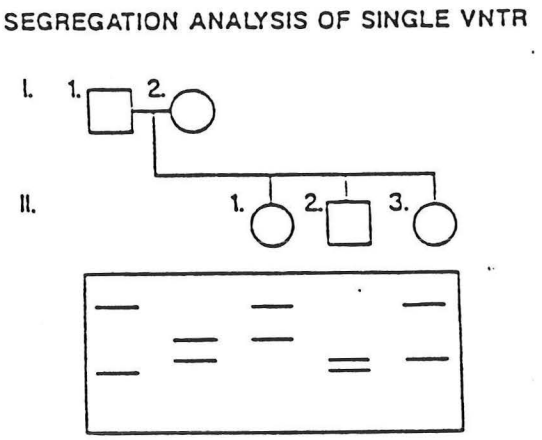
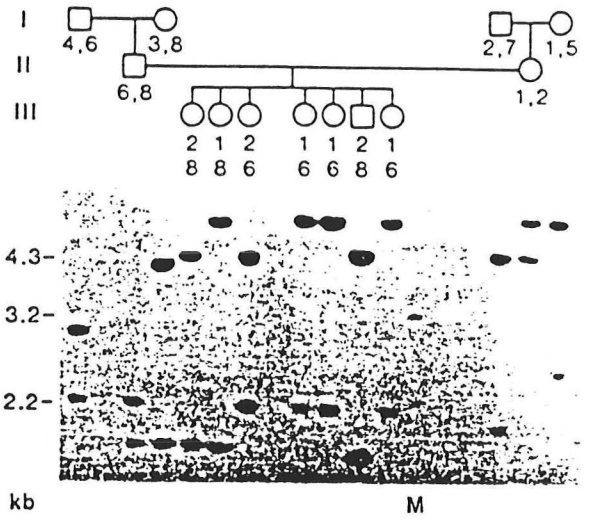
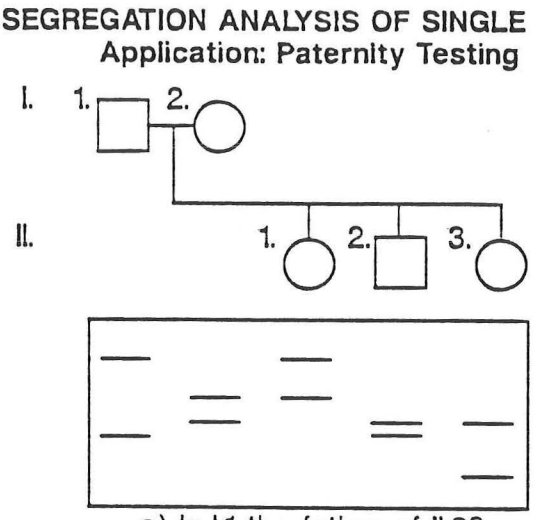


Figure 16B



Segregation analysis of individual VNTR loci has been used for paternity (and rarely, maternity) testing (Helminen, et al., 1988). A schematic representation of the segregation of alleles at a single VNTR locus in a family is shown in Figure 17. All four parental alleles generate different sized bands so their segregation can be followed. Individual II.3 has a band the same size as one of his mother's, but also has a band not found in the mother or father. There are two possibilities to account for the presence of this aberrant band. First, the father may not be the father. Second, perhaps a mutation has occurred in one of the father's gametes resulting in the formation of a new allele of different size (corresponding to a change in the number of repeats). Given the marked degree of length polymorphism seen at some of these VNTR loci, it would not be surprising to find that the spontaneous mutation rate can be high. In fact, the mutation rate has been measured in a variety of VNTRs and averages about 0.4%/VNTR locus/gamete (Jeffreys et al., 1988) though the rate varies significantly depending on the locus. Therefore, the spontaneous mutation rate of a locus must be determined prior to its use in paternity testing. And, since there is a real, though small incidence of new mutations, parental exclusions can not be made based on the analysis of a single VNTR locus.

Figure 17

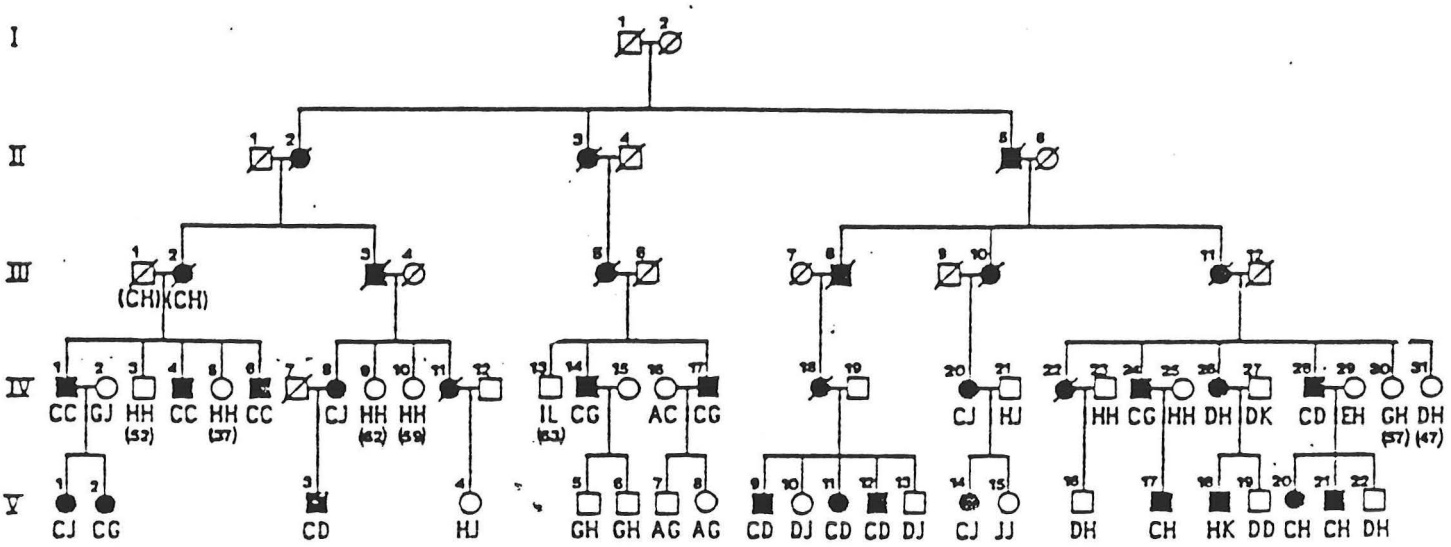


When DNA fingerprinting was systematically compared to HLA antigen analysis in 26 cases of paternity testing, in 24 cases it was determined to have a power of exclusion comparable to testing 63 HLA antigens. In two of the cases DNA fingerprinting was more informative than the HLA testing (Helminen, et al., 1988).

DNA fingerprinting has aided in the identification of genes causing disease. The most striking example of using a VNTR to map a gene causing a disease with no known mutant protein product is polycystic kidney disease (PCKD). Reeders and his colleagues found that a VNTR associated with the α-globin locus (Higgs, et al., 1983; Fowler et al., 1988) co-segregated with PCKD in a number of families (Reeders et al., 1985). One family is shown in Figure 18. The different VNTR alleles have been assigned different letters based on the length of the fragments identified on Southern blotting. In this family, every individual who inherits the "C" allele also has PCKD, suggesting that these two loci are linked. He analyzed nine families and there was only a single instance where the inheritance of a VNTR allele did not invariably co-segregate with the disease, and that was individual II.26 [Figure 18]. He has PCKD but inherited the "H" allele and not the "C" allele. He also transmitted the "H" allele (and the disease) to his son. Presumably this exception is due to a recombinational event between the VNTR locus and the gene responsible for PCKD in one of his parents' gametes. The frequency of such recombinations (which in this case is approximately 5%) is an indirect gauge of the distance between the genetic marker, i.e, the VNTR locus, and the gene encoding PCKD (which has been called PKD1). Efforts are now being directed to find a DNA marker closer to the gene, so that there is no recombination between the marker and the gene.

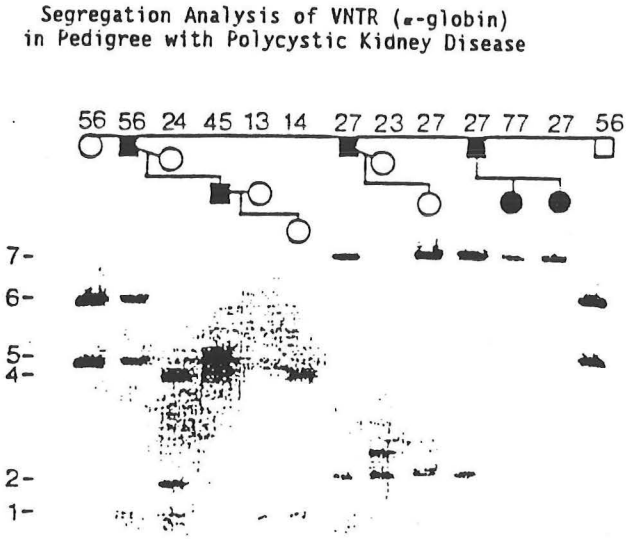
Figure 18

Segregation Analysis of VNTR (α-globin)
in Pedigree with Polycystic Kidney Disease.



Soon after this report, Kimberling, et. al. found a large Sicilian pedigree (over 250 members) with PCKD where the α-globin VNTR did not co-segregate with the disease (Kimberling et al., 1988). A partial pedigree of this family is shown in Figure 19. Above each pedigree member is a number that reflects the size of the two fragments on Southern blotting. 56 means the this individual has a band of size 5 kb and a band of 6 kb. Among the sibship of 5, an unaffected sister and an affected brother have inherited the same alleles from their parents, and these alleles are totally different from the two other affected brothers. The disease does not co-segregate with the α-globin VNTR in this family. Since these reports, numerous other families have been studied and three families from Italy, Sicily and Denmark have been found not to be linked with the α-globin gene. Approximately 4% of the families studied are not linked to PKD1 (Pieke et al., 1990). The lack of linkage between the VNTR locus and PCKD in these families is consistent with there being at least one other gene that can cause PCKD (Parkrey et al., 1990). The fact that the same disease can be caused by mutations in two different genes is not unprecedented, but complicates the use of these DNA markers to perform prenatal diagnosis.

Figure 19



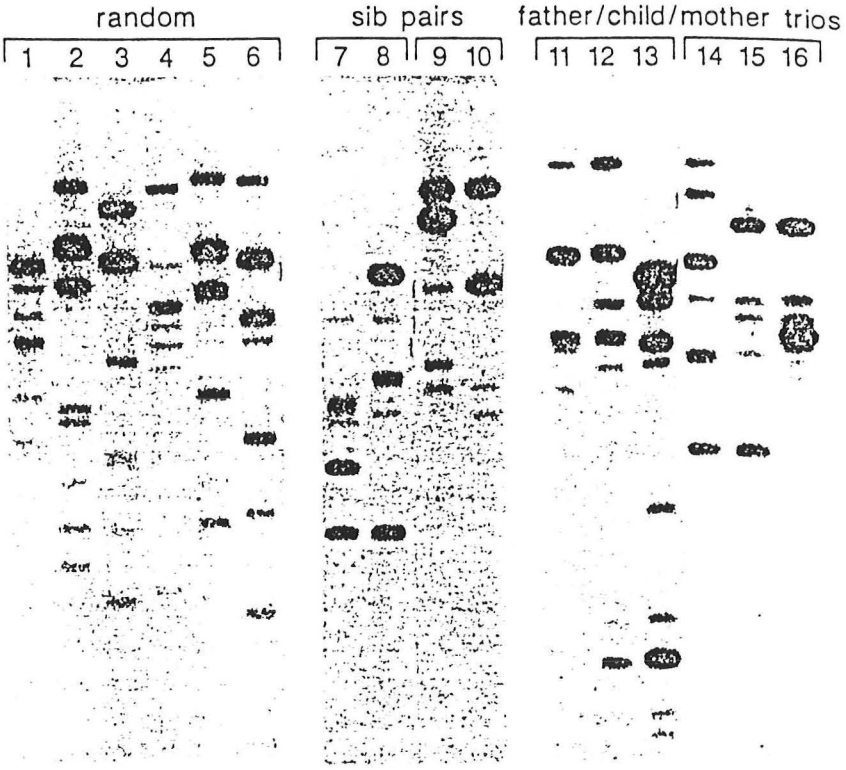
NEJM 319: 913, 1989

Numerous human genes have been found to have closely linked VNTRs which can be used for segregation analysis. These include the genes encoding insulin (Bell et al., 1982), α-globin (Higgs et al., 1986), the c-Ha-ras (Capon et al., 1983), myoglobin (Jeffreys, 1985), Type II collagen (Stoker et al., 1985), apolipoprotein B (Huang and Breslow, 1988, Knott et al., 1988), factor VII (O'Hara and Grant, 1988), the zeta-globin pseudogene (Goodbourn et al., 1983), and the gene responsible for Huntington's chorea (McDonald et al., 1990). Unfortunately, the distribution of VNTRs is not uniform in the genome. They are preferentially located in the telomeric regions of chromosomes (Royle et al., 1988, Simmler et al., 1987). Therefore, though VNTRs are highly informative genetic markers for closely linked genes, most genes do not have a VNTR nearby. However, the unequal distribution of VNTRs in the human genome does not limit their usefulness in determining individual identity. The high degree of polymorphism associated with VNTRs makes them excellent genetic markers to use for personal identification.

Two approaches have been taken to use VNTRs for personal identification. First, numerous VNTR loci in the genome can be analyzed at one time by performing Southern blot analysis and using multiple different locus-specific probes simultaneously. A genomic blot which was hybridized with 5 locus-specific probes is shown in Figure 20. Each of the VNTRs analyzed has a heterozygosity index of over 90%. Panel 1 shows the analysis of six unrelated individuals and no two individuals have the same pattern of bands (Wong, 1987). It is estimated that the chance of a second unrelated individual having an identical banding pattern using these 5 probes is 6×10^{-7}. In panel 2, sibling pairs are shown and,

as expected, they share about 50% of their bands. In the last panel, a father, child, and mother have been analyzed. Each band found in the child is seen in the father or the mother's sample.

Figure 20 Genomic Blot Using 5 VNTR Probes



The other method commonly used to analyze multiple VNTR loci at one time has been called DNA fingerprinting (Jeffreys, 1985). Jeffreys identified up to 40 different bands on Southern blot analysis using a probe derived from a VNTR locus associated with the myoglobin gene. This is because there are families of VNTRs which share a "core" sequence in each repeat. For example, Figure 21A shows the analysis of two different VNTRs that are revealed by Probe 1 and Probe 2. However, when Probe 3 is used, which contains the VNTR sequence itself, both sets of bands light up. A schematic representation of how a single VNTR probe can light up multiple restriction fragments is shown in Figure 21B. Chromosomes 1, 10, and 13 contain a single VNTR locus. Each locus is highly polymorphic in length but shares partial sequence homology. If a Southern blot is made using genomic DNA restricted with an enzyme that cuts outside the minisatellites, and is hybridized with a probe derived from the tandem repeat, multiple bands will be seen since the probe will hybridize to all the different VNTRs. If an individual is heterozygous for each site, the total number of bands seen will be twice the number of VNTR loci that hybridize with the probe [Figure 21B].

Figure 21A

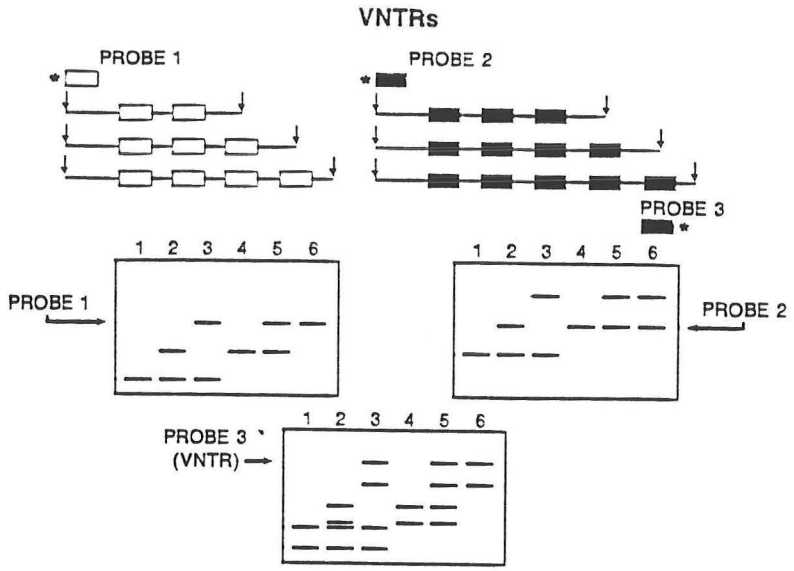
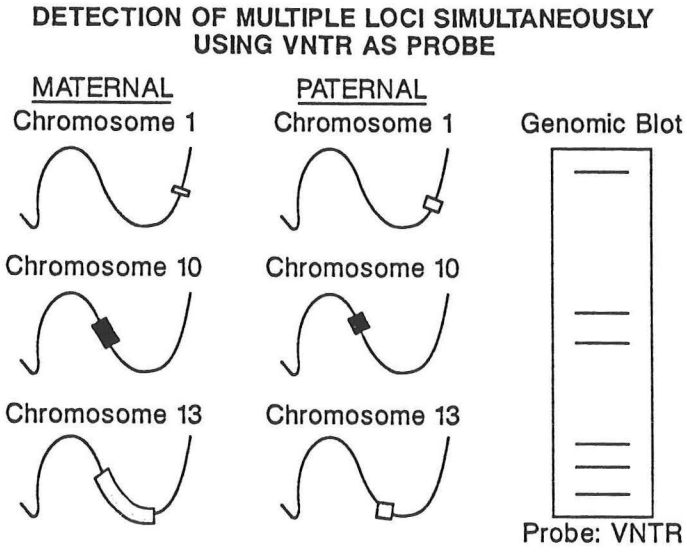
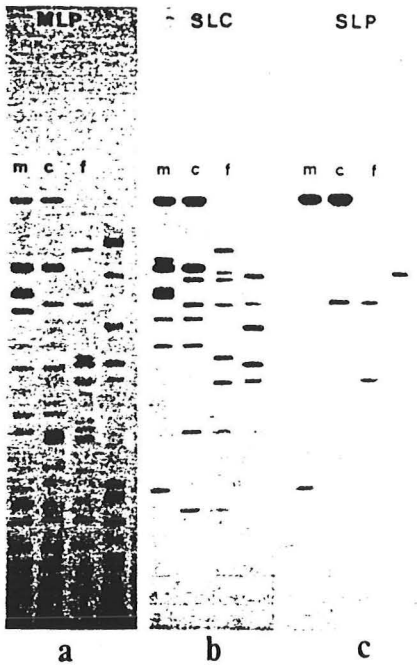


Figure 21B



Shown in Figure 22 is a comparison of the analysis of the same three individuals (m=mother, c=child and f=father) with a) multilocus probe, b) locus-specific probe "cocktail" and a c) single-locus probe.

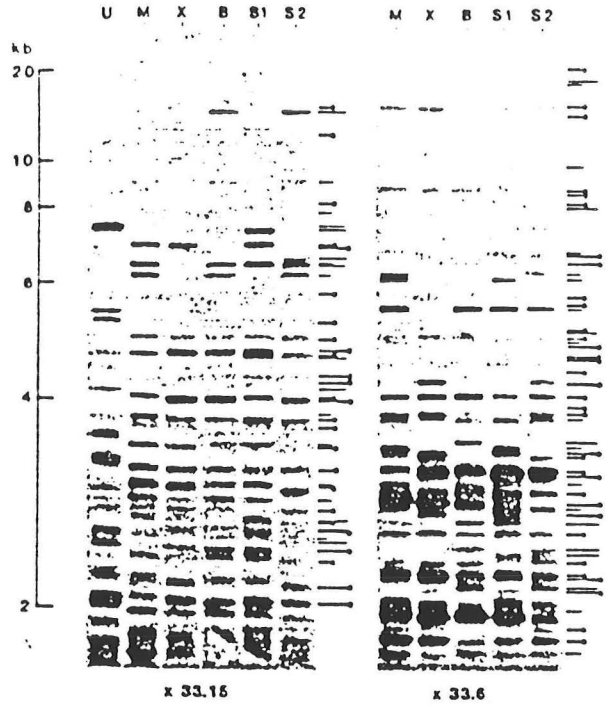
Figure 22



One of the first reports of using DNA fingerprinting to determine personal identity was in an immigration case (Jeffreys, 1985b). A Ghanaian boy who's mother lived in London went to visit his father in Ghana. When he returned, the immigration agency thought a substitution had been made with a maternal first cousin. Conventional protein marker analysis was done on the mother, her three children, and the purported son. The probability that the mother and the child were related was calculated as ~ 99%, but it could not be determined if she was the mother or an aunt. Unfortunately, the father and the mother's sister were not available for analysis. Therefore, Southern blot analysis was done using genomic DNA obtained from the mother, boy, brother, and two sisters. The blots were hybridized with a VNTR probe (from the myoglobin gene) and the results are shown in Figure 24 (U = unrelated, M = mother, X = index case, B = brother, S1 and S2 = sisters). A search was made in the mother's and sibs' samples for bands the same size as the child's. 50% of the proband's fragments were present in at least one sibling and not in the mother. These bands were presumed to be from the father. All the remaining fragments matched bands found in the mother. The probability of two unrelated individuals sharing a single band was estimated to be 26% (this is empiric). Since the proband shared 25 fragments with his mother, the probability of this occurring by chance alone was 0.26^{25} or 2.4×10^{-15}. (And the probability of the sister being the mother was calculated to be 6×10^{-6}). Therefore, it was determined that the child was indeed the son of the claimed mother and he was allowed back into the country. DNA fingerprinting has also been used successfully in Argentina to reunite relatives. Between 1975 and 1983, 9,000 Argentinean citizens disappeared. Some of the children born to captive women were adopted surreptitiously by other individuals. By performing DNA fingerprinting on children

and their relatives, 50 children have been reunited with their relatives.

Figure 23



This methodology has also been applied to the diagnosis of crimes, particularly those of rape and murder (Marx, 1988). Prior to the development of DNA fingerprinting, a suspect could be identified with only 90-95% certainty by analyzing blood group antigens, red cell enzymes, and HLA typing of semen, blood and hair left at the scene of the crime. The certainty of a correct identification is dramatically improved using DNA fingerprinting. There remains some controversy about the use of DNA fingerprinting in the courtroom, and though the methodology and standards are still evolving, it is clear that it is here to stay. The strategy employed is to match the "DNA fingerprints" of the suspect with the DNA fingerprints of any human material left at the scene of the crime.

Figure 24 summarizes the quantity of material needed to obtain a reliable "fingerprint" either by Southern blotting or by using the PCR technique. The material - whether it be blood, semen or hair - should be stored dried or frozen. Humidity accelerates the degradation of DNA. DNA fingerprinting has been successfully performed on 4 year old blood and semen stains (Gill, et al., 1985). To perform Southern blot analysis to look at multiple VNTR loci requires between 0.1 to 1 milliliters (ml) of blood or 5 microliters (μl) of semen. To perform a multilocus DNA fingerprint requires 0.1 to 1 micrograms (μg) of genomic DNA (Jeffreys, 1985). If a locus-specific probe is used, only 50 nanograms (ng) of DNA

are required. The PCR technique can be used on a single cell or a single sperm (Li, et al., 1988). The most common evidence at the scene of a crime is hair, and a freshly plucked hair (a hair that includes the root) has about 100 ng of DNA. It requires between 1 to 5 hairs to perform a Southern blot analysis. However, with PCR, one freshly plucked hair is sufficient to perform DNA fingerprinting (Higuchi, 1988). Finally, shed hairs have much less DNA and Southern blot analysis cannot be done on a single (or many) shed hair. However, PCR analysis can be done on a single shed hair which contains less than 10 ng of DNA (Higuchi, 1988).

Figure 24

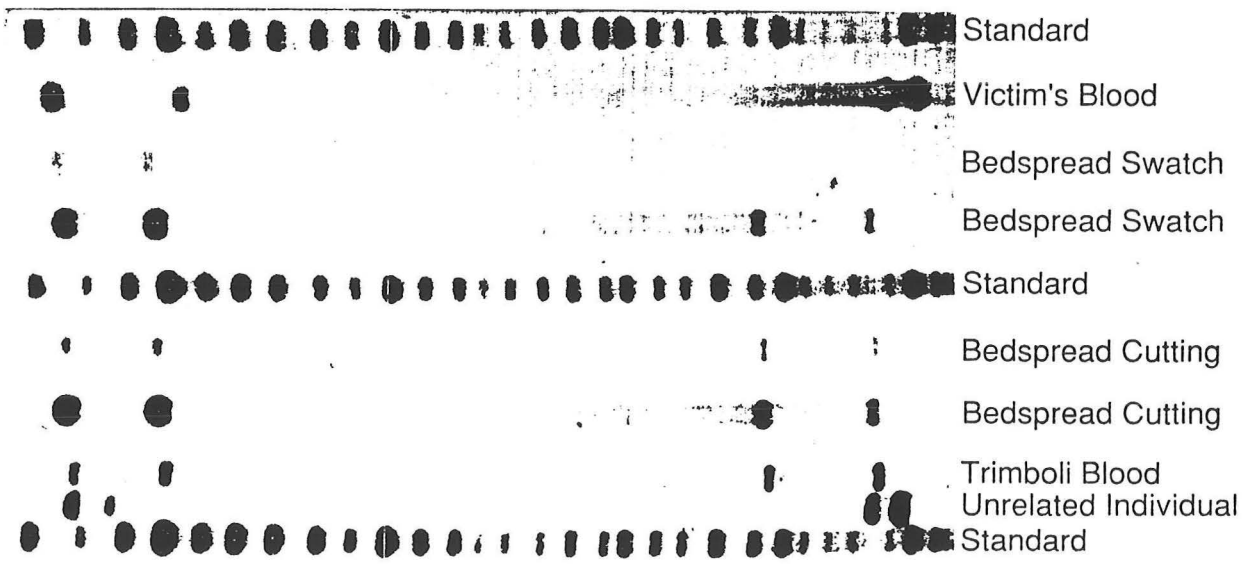
		SOUTHERN BLOTTING		PCR	
	Tissue	Amount	DNA	Amount	DNA
1.	Blood	100-1000 ul	0.1-1.0 ug	1 cell	< 1 ng
2.	Semen	5 ul	"	1 sperm	< 1 ng
3.	Hair-plucked	1-5	$\leq$ 0.5 ug	1 hair	~200 ng
4.	Hair-shed	--	--	1 hair	<10 ng

The first time this methodology was used to solve a crime was in 1987. In Narborough Village, England, there were two rape/murders of teenage girls - one in 1983 and another in 1986. Semen had been collected and saved. The DNA fingerprints of the two semen samples were identical so it was concluded that the same man had raped both women. There was a 17 year old suspect who was charged with the murder based on circumstantial evidence. He actually confessed to the crime. However, when his DNA was examined it had a completely different banding pattern so he was released. The police collected blood from ~ 3,500 males in the region between the ages of 18 and 30. 60% could be excluded as based on blood typing. The other 40% were excluded by DNA fingerprinting. However, someone overheard a man in a pub who said he had given 2 blood samples - one in the name of a person with whom he worked. That person was contacted and, when his DNA was analyzed it gave an identical pattern to that of the semen samples. Subsequently, the man confessed to both crimes and went to prison.

Closer to home there was a case solved using DNA fingerprinting: the case of Mr. Trimboli. Mr. Trimboli was accused of killing a 14 year old neighbor, her boyfriend, and her sister. The young teenage girl was raped before (or after) she was killed, and semen was left on the bedspread. Protein studies were performed on the sperm and disclosed a match with Mr. Trimboli. However, the test results also matched 22% of the population of Arlington. The first trial ended in a mistrial and the second

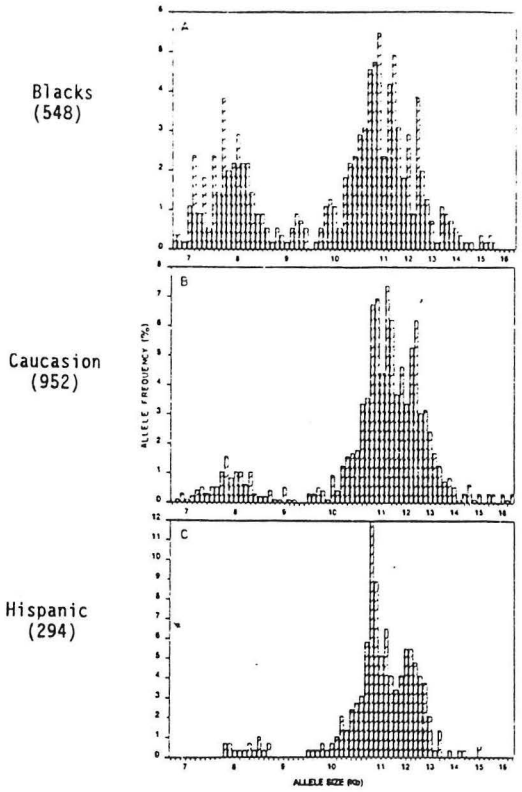
trial resulted in a hung jury. Then, during the preparations for the 3rd trial, the defense team learned about DNA fingerprinting. They presented the option of having DNA fingerprinting performed to Mr. Trimboli to prove his innocence. Mr. Trimboli agreed to be tested. DNA was both eluted and extracted from the semen stains left on the bedspread. A blood sample was obtained from Trimboli and had been obtained from the rape victim. All of the DNA samples were digested individually with a restriction enzyme, PstI, fractionated on an agarose gel, transferred to nitrocellulose, and probed with four different probes that each reveal single VNTR loci. Two bands were seen with each probe. A reproduction of one of the Southern blots is shown in Figure 25. Restricted DNA from the victim is shown in lane 2. The pattern of bands are clearly different from those in lanes 3,4,6 and 7 which contain the sperm samples taken from the bedspread. The pattern is also different from the DNA sample from an unrelated individual (lane 9). However, the banding pattern of the sperm DNA from the bedspread (lanes 3,4,6 and 7) and from Trimboli (lane 8) were identical. Actually, they were not exactly identical. The bands seen in Trimboli's sample migrated a little slower. This slower migration was seen consistently in the analysis of all eight bands and was thought to be an artifact due to differences in the methodology used to isolate the DNA. It was concluded from these studies that all eight bands analyzed from the sperm sample taken at the scene of the crime were the same size as those found in the DNA sample from Trimboli. In order to calculate the likelihood of a match, the frequency of each band seen in Trimboli's sample was determined in the general population.

Figure 25



When computing the probability of another individual having bands of the exact same size, it is essential that the correct population be used as a reference. This is important since the likelihood of the banding pattern matching by chance depends on the frequency of the bands in the population. If a particular band is very frequent in the population then the fact that there is a match is much less significant. The frequencies of bands differ (sometimes quite dramatically) depending on the population being studied (Baird, 1986; Baird, 1989; Balazs, 1989). This is demonstrated in Figure 26 where the frequency of each band of a VNTR was determined in unrelated individuals from 3 different populations.

Figure 26



Am. J. Hum. Genet. 44: 185, 1989

Once the frequency of each band is determined in the population, the probability of an unrelated individual having the same band is calculated. Then the probabilities (for each band) are multiplied together. The final calculation is based on the assumption that all the different VNTR loci are in linkage equilibrium. In the case of Trimboli, it was determined that the likelihood of another individual having the exact same constellation of bands with the same four probes was 1 in 54,900,000,000 (see Figure 27). This information was admitted as evidence in Mr. Trimboli's third trial, and he was convicted.

Figure 27

Comparison of Band Pattern in Trimboli with General Population

Population	Probe: 1	2	3	4
Blacks	1 in 1387	1 in 806	1 in 63	1 in 263
Caucasians	1 in 1621	1 in 1010	1 in 77	1 in 441

Frequency of the same banding pattern using all 4 probes:

In the white population = 1 in 54,900,000,000
In the black population = 1 in 18,500,000,000

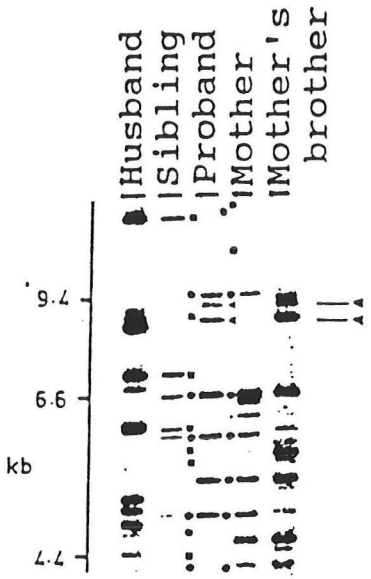
This is only one of many trials in the last few years that has used DNA fingerprinting as evidence. The tests have not always been performed flawlessly. Dr. Eric Lander has summarized many of the problems and concerns of using DNA fingerprinting as evidence in court (Lander, 1989). Among his concerns is the fact that oftentimes the proper controls for the analysis have not been performed. For example, in the Trimboli case, he would argue that the DNA from the suspect and the criminal material should have been mixed together prior to fractionating on a gel to make sure that the bands were exactly the same size in both samples. Internal controls should be run so that the blots are hybridized with a probe that reveals an invariant band, especially in cases where there are concerns about aberrant migration. It is important that the reference population be accurate. It may be that there are subpopulations within racial groups where allele frequencies deviate significantly from the general population. For some companies, the measurements made to determine a match are less precise than those made to determine the frequencies in the population. This results in an exaggerated low probability of there being a match in the general population. Despite these problems, recently the Office of Technology Assessment (OTA), which is responsible for the review of new analytical techniques, has affirmed the validity and the accuracy of DNA fingerprinting. The FBI has standardized the techniques and begun processing material routinely for use in litigation. Computer database banks for DNA fingerprints will soon complement the pre-existing data banks for "real" fingerprints.

The analysis of VNTR loci has also been used successfully utilized for the following purposes:

1) To perform genetic counseling - nonpaternity becomes an issue for the genetic counselor. For example, there was a case of a woman who was heterozygous for a mutation in the β-globin gene and had a child with β-thalassemia. She and her husband sought

genetic counseling so that prenatal diagnosis could be performed. Her husband was found to be homozygous for a normal β-globin gene. DNA fingerprinting was performed on the entire family and the child's biological father was determined to be the mother's brother (Wells, 1988) [Figure 28].

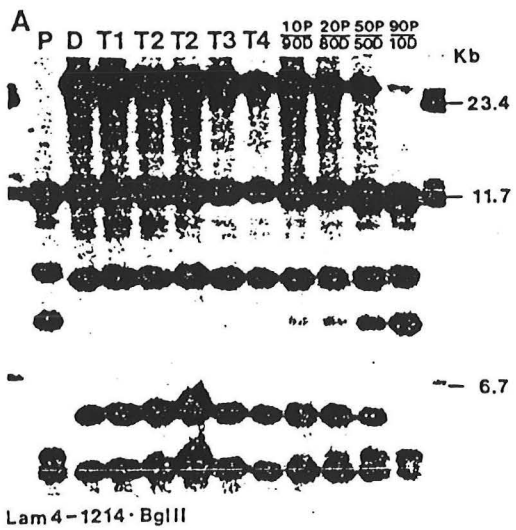
Figure 28 DNA Fingerprinting with Multilocus Probe



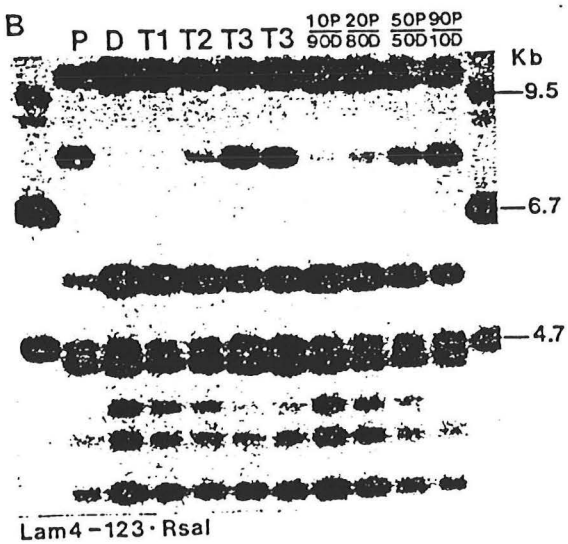
2) To determine the origin of bone-marrow tissue status-post bone marrow transplant. Many bone marrow transplants are between siblings and they can not be distinguished 25% of the time using a single genetic marker. DNA fingerprinting facilitates greatly the ability to determine whether a bone marrow transplant has been successful (Knowlton et al., 1986; Thein et al., 1986; Min et al., 1988, Palkkala et al., 1988). In Figure 29, P= patient, D= donor, and T_1, T_2, T_3 and T_4 are different time intervals after transplant. Note in Panel A that only the donors pattern of fragments persist. In panel B, after a period of time the patient's fragments are again visible which is evidence of recurrence of disease.

Figure 29 Genomic Blot of Bone Marrow Recipient and Donor

Successful transplant



Unsuccessful transplant

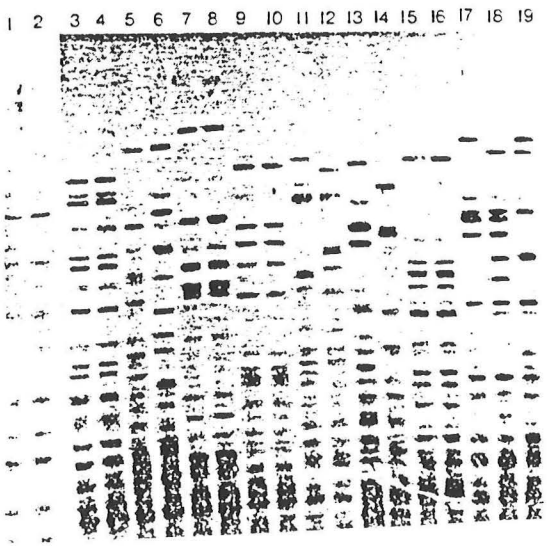


Another case where DNA fingerprinting was used was in a Pakistani boy who presented at age 3 with acute lymphocytic leukemia. He was treated and relapsed two years later. A bone marrow transplant was performed using his brother's marrow. The brother was identical for 35 blood antigens. On day 21, after a stormy hospital course, purged autologous marrow was infused for failure to engraft. Five days later, his granulocyte count was 500/cumm. Were the cells in his blood his brother's from the bone marrow transplant or the patient's from the autologous marrow infusion? DNA fingerprinting was performed and disclosed that the circulating cells had an identical DNA fingerprint to his own cell (Thien et al., 1986). Therefore, bone marrow engraftment of the brother's cells had not taken place.

DNA fingerprints helped clarify a very unusual case involving a woman who had a stillbirth. Karyotype analysis was performed on her lymphocytes. A total of 200 mitoses were studied and she was found to be 46,XY! She was tested for the H-Y antigen and was positive! She had a twin brother and DNA fingerprint analysis was performed using DNA from her blood cells, her brother's blood cells, and her fibroblasts. The DNA fingerprints from her blood cells matched those of her twin and were different from her fibroblasts. Karyotyping of skin fibroblasts was done and she was 46,XX and therefore she is chimeric. Her twin's blood cells must have populated her bone marrow in utero.

3) To determine whether twins are identical (Hill, 1985). In 30% of twins the sexes are discordant and so the twins are obviously fraternal. Approximately 20% of twins share the same placental and chorion and these are invariably monozygotic. In the remaining 50% of twins, it is not obvious whether they are identical or fraternal. This question can easily be answered using DNA fingerprinting as demonstrated in Figure 30. In each pair of lanes, DNA from sets of twins has been restricted with an enzyme, and probed with a minisatellite probe. Some pairs have an identical banding patterns (lanes 3 & 4) and others do not (lanes 5 & 6).

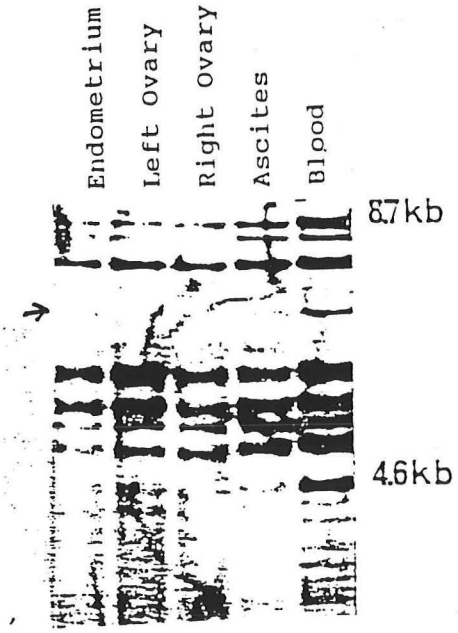
Figure 30 Genomic Blot from Twin Pairs Probes with Myoglobin VNTR



4) To screen for DNA rearrangements in tumor tissues (Thein et al., 1987; Lothe et al., 1988, Lagoda, et al., 1989) and determine which tumors are clonal in nature.

An example where DNA fingerprinting has aided in the diagnosis and treatment of an individual with cancer is shown in Figure 32. A 68 year old woman presented with a six month history of lower abdominal swelling. On laparotomy she was found to have bilateral ovarian tumors. She also had an endometrial tumor. She had multiple metastatic lesions involving the omentum and peritoneal cavity. Was the endometrial tumor a metastatic lesion or another primary tumor? DNA fingerprinting of the tumor tissue from the ovaries, omental implant and endometrial tumor was performed and all the tissues were found to have an identical banding pattern that was different from the pattern seen in the "normal" blood cells [Figure 31]. There is a band missing in the tumor tissues that is present in the blood cells (which is demarcated by the arrow in Figure 31). This is commonly seen when DNA fingerprinting is performed on tumors because of the high frequency of associated chromosomal rearrangements and mutations. Since all the tumor tissues had the same DNA fingerprint, the endometrial tumor is most certainly a metastatic lesion from the ovary.

Figure 31



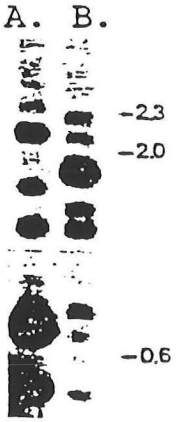
5) To map genes that cause disease. The high degree of polymorphism at VNTR loci make them ideal DNA markers for genetic mapping (Donis-Keller, 1987; Nakamura et al., 1987; Page et al., 1987).

6) To subtype bacteria and trace the outbreak of an infectious disease. Due to variations in DNA sequences in bacteria, when bacterial DNA is cut, a characteristic fragment pattern is seen (Viering and Fine, 1989). DNA fingerprinting was used to analyze

the epidemiology of an outbreak of rheumatic fever (Cleary et al., 1988). In 1986 there was an outbreak of rheumatic fever in various geographically distinct regions of the United States (Utah, Minnesota, New York, Colorado). The streptococcus A strain was the same type (M18) in all locales. This same strain previously had been responsible for an outbreak of rheumatic fever at Lowry Air Force Base in ~1968. Were the same strains responsible for the outbreak in Utah, Minnesota, and Denver or were they different strains? Was it the same strain that had been responsible for the outbreak over 20 years earlier? Based on the analysis of the DNA fingerprints of the bacteria, the same bacterial strain was responsible for all the outbreaks. Now efforts can be directed to identify the genetic determinants that make this particular substrain so virulent.

DNA fingerprinting can also be performed on other organisms, including Plasmodium falciparum. Figure 32 shows that the DNA fingerprints from a chloroquine-sensitive and chloroquine-resistant strain of P. falciparum are quite different.

Figure 32 DNA Fingerprints of Two Strains of Plasmodium falciparum

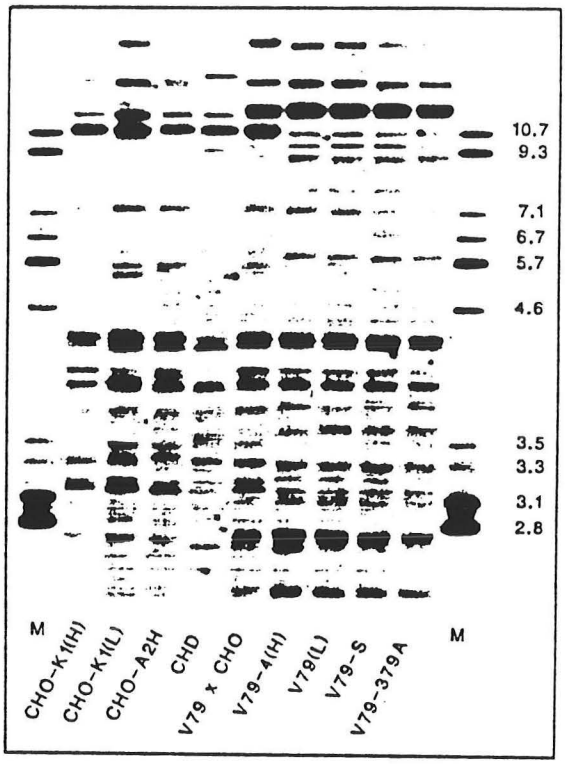


A. Chloroquine-sensitive
Haiti 135

B. Chloroquine-resistant
Indochina I

7) To determine if there has been contamination of tissue culture cell lines (Masters et al., 1988). Some studies performed on the same cell line in different labs have generated different results due to tissue culture cross-contamination. The identity of two cell cultures can be determined by DNA fingerprinting. Figure 33 shows the DNA fingerprints of various cell lines.

Figure 33 DNA Fingerprints of Various Cultured Cell Lines

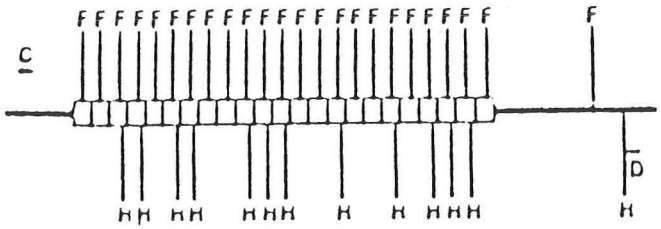


One of the major difficulties of using VNTR loci as genetic markers, particularly for paternity testing and suspect identification, is the problem of how to interpret novel bands. Erroneous conclusions can be made in cases where a spontaneous mutation has occurred producing a band of a different size. As stated previously, the average spontaneous mutation rate for VNTR loci is about 0.4% per gamete (Jeffreys et al., 1985). For some probes, like the λMS31 probe, the frequency approaches 5% (Jeffreys et al., 1988). The rate of new mutations at these loci are in general higher than for point mutations (Jeffreys, 1985b). As might be predicted, the spontaneous mutation rates tend to correlate directly with the degree of polymorphism (Jeffreys, 1988), but must be computed for each VNTR locus used.

When various VNTRs have been sequenced, it was found that the repeats are not completely homologous in sequence. Base pair substitutions frequently shared between multiple repeats of a VNTR. Jeffreys developed a method to assay for these sequence differences between the repeat units by performing a partial digestion of the repeated sequence using an enzyme that cuts within some, but not all, repeats, as shown in Figure 34 (F and H demarcate different restriction enzyme sites). This system is capable of distinguishing over 10^{70} alleles at one VNTR site (Jeffreys et al., 1990). He calls these variations in patterns minisatellite variant repeats (MVR). He used this method to analyze the generation of new alleles at VNTR loci.

Figure 34 Minisatellite Variant Repeats (MVRs) Mapping

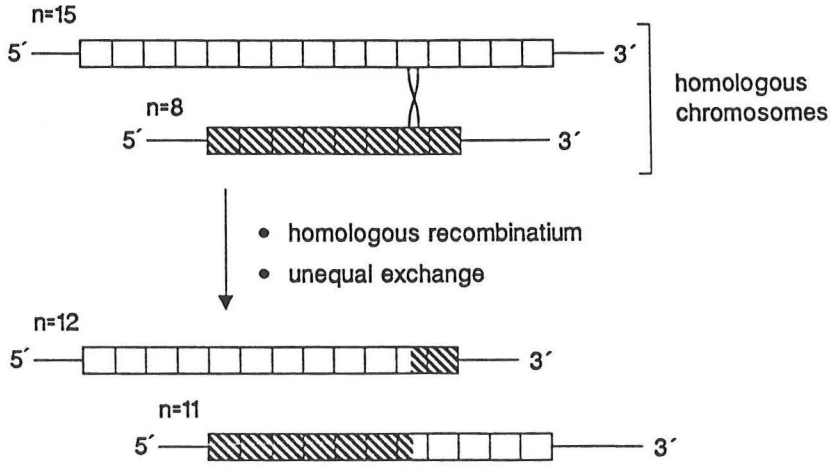
A. Restriction Map of VNTR



B. Partial Digest of VNTR

The mechanism responsible for the VNTR length polymorphism was thought to be due to homologous recombination with unequal exchange which is shown schematically in Figure 35. Subsequently, the homologous chromosomes pair. After the first division of meiosis the chromosomes misalign during pairing and a cross-over event occurs within the repeated sequences, alleles of new length can be generated.

Figure 35 **GENERATION OF LENGTH POLYMORPHISM IN VNTR**



The observations that suggested that homologous recombination with unequal exchange was the mechanism responsible for the generation of length polymorphisms in VNTRs are the following:

1) The most polymorphic VNTR loci tend to be longer in length and have repeats that are more homologous in sequence.

2) VNTRs usually vary in size by several repeats.

3) Many VNTR sequences share a common core sequence that is homologous to the Chi sequence of E. coli and bacteriophage. This sequence is thought to be a hot spot for recombination (Jeffreys, 1985).

4) At least one VNTR has been shown to increase the rate of recombination in vitro (Wahls et al., 1990) and some VNTR have been found to be hotspots for recombination (Steinmetz et al., 1986).

5) In situ studies of VNTR find them clustered at crossover sites during meiosis (Chandley and Mitchell, 1988).

6) The frequency of generation of alleles of new length are equal in men and women despite the fact that they differ markedly in the number of cell divisions prior to the differentiation of germ cells (oocytes have 24 and sperm 400 divisions) (Jeffreys et al., 1988).

7) Only a few cases of somatic or germinal mosaicism have been identified. (Kelly et al., 1989; Nurmberg et al., 1989; Jeffreys et al., 1990).

However, recent molecular characterization of new mutant alleles have shown no evidence of homologous recombination (Wolff et al., 1988; Jeffreys et al., 1990). When Jeffreys identified 64 new deletions in sperm and 42 somatic mutations, he found no evidence of recombination. Rather, by using the MVRs he was able to demonstrate that there was no exchange between the two parental chromosomes. He concluded from these studies that the generation of new alleles occurs within the same parental chromosome, perhaps between sister chromatids during mitosis. Each VNTR represents the maternal or paternal chromosome lineage and thus can be used much in the same way that mitochondrial DNA is used to follow a single parental lineage (Jeffreys et al., 1990). The mechanism responsible for the generation of alleles of different length awaits further delineation.

Why do we, as well as other species, have these sequences in our genome? Minisatellite have been identified in other mammals (cats, dogs, cows), birds, higher plants, fungi and protozoa. The fact that VNTRs have been found in most organisms suggests they perform some function. The sequence of the X gene of hepatitis B virus is homologous to some human VNTR loci. The X gene region in the hepatitis virion is thought to be the site of integration of viral DNA into the genome of the host. This has raised the question as to whether some of the VNTR sequences may have a viral origin (Nakamura, 1987).

Finally, I will end on a seemingly bucolic note. Figure 36 shows the DNA fingerprints from two highly studied adult sparrows, and their eleven offspring from four different broods (Burke, 1987). Despite all appearances of being happily coupled love birds, there is one offspring who clearly has a different banding pattern than the "father" (demarcated by an "X"). Birds of a feather do more than flock together! This is the first reported case of paternity testing in house sparrows and it looks like even sparrows have their milkmen!

The applications of DNA fingerprinting to the study of animal, as well as human, sociobiology and population genetics is only in its infancy.

Figure 36 DNA Fingerprints of House Sparrow Family

References

Armour, J.A.L., Z. Wong, V. Wilson, N.J. Royle and A.J. Jeffreys. 1989. Sequences flanking the repeat arrays of human minisatellites: Association with tandem and dispersed repeat elements. Nuc. Acids Res. 17: 4925-4935.

Baird, M., I. Balazs, A. Giusti, L. Miyazaki, et al. 1986. Allele frequency distribution of two highly polymorphic DNA sequences in three ethnic groups and its application to the determination of paternity. Am. J. Hum. Genet. 39: 489-501.

Balazs, I., M. Baird, M. Clyne, and E. Meade. 1989. Human population genetic studies of five hypervariable DNA loci. Am. J. Hum. Genet. 44: 182-190.

Bell, G. I., M.J. Selby, and W.J. Rutter. 1982. The highly polymorphic region near the human insulin gene is composed of simple tandemly repeating sequences. Nature 295: 31-35.

Boerwinkle, E., W. Xiong, E. Fourest, and L. Chan. 1989. Rapid typing of tandemly repeated hypervariable loci by the polymerase chain reaction: Application to the apolipoprotein B 3' hypervariable region. Proc. Natl. Acad. Sci. USA 86: 212-216.

Burke, T. and M.W. Bruford. 1987. DNA fingerprinting in birds. Nature 327: 149-152.

Capon, D.J., E.Y. Chen, A.D. Levinson, P.H. Soeburg, and D.V. Goeddel. 1983. Complete nucleotide sequences of T24 human bladder carcinoma oncogene and its normal homologue. Nature 302: 33-57.

Cawood, A.H. 1989. DNA Fingerprinting. Clin. Chem. 35: 1832-1837.

Chandley, A.C., and A.R. Mitchell. 1988. Hypervariable minisatellite regions are sites for crossing-over at meiosis in man. Cytogenet. Cell Genet. 48: 152-155.

Cleary, P.P., E.L. Kaplan, C. Livdahl, and S. Skjold. 1988. DNA fingerprints of Streptococus pyogenes are M type specific. J. Infect. Dis. 158: 1317-1323.

Cohen, J.E. 1990. DNA fingerprinting for forensic identification: Potential effects on data interpretation of subpopulation heterogeneity and band number variability. Am. J. Hum. Genet. 46: 358-368.

Donis-Keller, H., P.Green, C. Helms, S. Cartinhour, et al. 1987. A genetic linkage map of the human genome. Cell 51: 319-337.

Dover, G.A. 1990. Mapping 'frozen accidents'. Nature 344: 812.

Farber, C.M., M. Georges, G. DeBock, A. Verhest, P. Simon, M. Verschraegen-Spae, and G. Vassart. 1989. Demonstration of spontaneous XX/XY chimerism by DNA fingerprinting. Hum. Genet. **82**: 197-198.

Fowler, S.J., P. Gill, D.J. Werrett, and D.R. Higgs. 1988. Individual specific DNA fingerprints from a hypervariable region probe: alpha-globin 3' HVR. Hum. Genet. 79: 142-146.

Gelehrter, T.D. and F.S. Collins. 1989. Principles of Medical Genetics. Williams & Wilkins, Baltimore, MD.

Gill, P., Jeffreys, A.J., and D.J. Werrett. 1985. Forensic application of DNA 'fingerprints'. Nature 318: 577-579.

Goodbourn, S.E.Y., Higgs, D.R., Clegy, J.B., and Weatherall, D.J. 1983. Molecular basis of length of polymorphism in the human-zeta-globin gene complex. Proc. Natl. Acad. Sci. USA 80: 5022-5026.

Gyllensten, U.B., S. Jakobson, H. Temrin, and A.C. Wilson. 1989. Nucleotide sequence and genomic organization of bird minisatellites. Nucleic Acids Research 17: 2203-2214.

Helminen, P., C. Ehnholm, M.-L. Lokki, A. Jeffreys and L. Peltonen. 1988. Application of DNA "Fingerprints" to paternity detemrinations. The Lancet, March 12.

Higuchi, R., C. H. von Beroldingen, G.F. Sensabaugh, and H.A. Erlich. 1988. DNA typing from single hairs. Nature 332: 543-546.

Hill, A.V.S. and A.J. Jeffreys. 1985. Use of minisatellite DNA probes for determination of twin zygosity at birth. Lancet i: 1394-1395.

Horn, G.T., G. Richards, and K.W. Klinger. 1989. Amplification of a highly polymorphic VNTR segment by the polymerase chain reaction. Nucleic Acids Research 17: 2140.

Huang, L.-S. and J.L. Breslow. 1987. A unique AT-rich hypervariable minisatellite 3' to the ApoB gene defines a high information restriction fragment length polymorphism. J. Biol. Chem. 262: 8952-8955.

Jeffreys, A.J. 1979. DNA sequence variants in the Gy, Ay, α- and β-globin genes of man. Cell 18: 1-10.

Jeffreys, A.J., V. Wilson, and S.L. Thein. 1985a. Hypervariable 'minisatellite' regions in human DNA. Nature 314: 67-73.

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

Jeffreys, A.J., N.J. Royle, V. Wilson, and Z. Wong. 1988. Spontaneous mutation rates to new length alleles at tandem-repetitive hypervariable loci in human DNA. Nature 332: 278-281.

Jeffreys, A.J., V. Wilson, R. Neumann, and J. Keyte. 1988. Amplification of human minisatellites by the polymerase chain reaction: Towards DNA fingerprinting of single cells. Nucleic Acids Research 16: 10953-10971.

Jeffreys, A.J., R. Neumann, and V. Wilson. 1990. Repeat unit sequence variation in minisatellites: A novel source of DNA polymorphism for studying variation and mutation by single molecule analysis. Cell 60: 473-485.

Kan, Y.W. and A. M. Dozy. 1978. Polymorphism of DNA sequences adjacent to human β-globin structural gene: Relationship to sickle mutation. Proc. Natl. Acad. Sci. USA 75: 5631-5635.

Kelly, R., G. Bulfield, A. Collick, M. Gibbs, and A.J. Jeffreys. 1989. Characterization of a highly unstable mouse minisatellite locus: Evidence for somatic mutation during early development. Genomics 5: 844-856.

Kimberling, W.J., P.R. Fain, J.B. Kenyon, D. Goldgar, E. Sujansky and P.A. Gabow. 1988. Linkage heterogeneity of autosomal dominant polycystic kidney disease. N. Engl. J. Med. 319: 913-918.

Kirby, L.T. 1990. DNA Fingerprinting. An Introduction. Stockton Press, New York, New York.

Knott, T.J., S.C. Wallis, R.J. Pearse, L.M. Powell, and J. Scott. 1986. A hypervariable region 3' to the human apolipoprotein β gene. Nucleic Acids Research 14: 9215-9216.

Knowlton, R.G., V.A. Brown, J.C. Braman, D. Barker, J.W. Schumm, C. Murray, T. Takvorian, J. Ritz, and H. Donis-Keller. 1986. Use of highly polymorphic DNA probes for genotypic analysis following bone marrow transplantation. Blood 68: 378-385.

Lagoda, P.J.L., G. Seitz, J.T. epplen, and O.-G. Issinger. 1989. Increased detectability of somatic changes in the DNA from human tumours after probing with "synthetic" and "genome-derived" hypervariable multilocus probes. Hum. Genet. 84: 35-40.

Lander, E.S. 1989. DNA fingerprinting on trial. Nature 339: 501-505.

Li, H., U.B. Gyllensten, X. Cui, R. K. Saiki, H.A. Erlich, and N. Arnheim. 1988. Amplification and analysis of DNA sequences in single human sperm and diploid cells. Nature 335: 414-417.

Lothe, R.A., Y. Nakamura, S. Woodward, T. Gedde-Dahl, Jr., and R. White. 1988. VNTR (Variable number of randem repeats) markers show loss of chromosome 17p sequences in human colorectal carcinomas. Cytogenet. Cell Genet. 48: 167-169.

MacDonald, M.E., S.V. Cheng, M. Zimmer, J.L. Haines, A. Poustka, B. Allitto, B. Smith, W.L. Whaley, D.M. Ramano, J.Jagadeesh, R. H. Myers, H.Lehrach, J.J. Wasmuth, A.M. Frischauf, and J.F. Gusella. 1989. Clustering of multiallele DNA markers near the Huntington's Disease gene. J. Clin. Invest. 84: 1013-1016.

Marx, J.L. 1988. DNA fingerprinting takes the witness stand. Science 240: 1616-1618.

Masters, J.R.W., P. Bedford, A. Kearney, S. Povey, and L.M. Franks. 1988. Bladder cancer cell line cross-contamination: Identification using a locus-specific minisatellite probe. Br. J. Cancer 57: 284-286.

Min, G.L., J. Hibbin, C. Arthur, J. Apperley, A. Jeffreys, and J. Goldman. 1988. Use of minisatellite DNA probes for recognition and characterization of relapse after allogeneic bone marrow transplantation. Brit. J. Haematology 68: 195-201.

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

Nakamura, Y., M. Carlson, K. Krapcho, M. Kanamori, and R. White. 1988. New approach for isolation of VNTR markers. Am. J. Hum. Genet. 43:854-859.

Nürnberg, P., L. Roewer, H. Neitzel, K. Sperling, A. Pöpperl, J. Hundrieser, H. Pöche, C. Epplen, H. Zischler, and J.T. Epplen. 1989. DNA fingerprinting with the oligonucleotide probe $(CAC)_5/(GTG)_5$: somatic stability and germline mutations. Hum Genet. 84: 75-78.

O'Hara, P.J. and F.J. Grant. 1988. The human factor VII gene is polymorphic due to variation in repeat copy number in a minisatellite. Gene 66: 147-158.

Page, D.C., K. Bieker, L.G. Frown, S. Hinton, M. Leppert, J.-M. Lalouel, M. Lathrop, M. Nystrom-Lahti, A. de la Chapelle, and R. White. 1987. Linkage, Physical Mapping, and DNA sequence analysis of pseudoautosomal loci on the human X and Y chromosomes. Genomics 1: 243-256.

Pakkala, S., P. Helminen, U.M. Saarinen, R. Alitalo, and L. Peltonen. 1988. Differences in DNA-fingerprints between remission and relapse in childhood acute lymphoblastic leukemia. Leukemia Res. 12: 757-762.

Parkrey, P.S., J.C. Bear, J. Morgan, B.C. Cramer, P.J. McManamon, M.H. Gault, D.N. Churchill, M. Singh, R. Hewitt, S. Somlo and S.T. Reeders. 1990. The diagnosis and prognosis of autosomal dominant polycystic kidney disease. New Engl. J. Med. 323: 1085-1090.

Pieke, S.A., Kimberling, W.J., Kenyon, J.B. and P. Gabow. 1989. Genetic heterogeneity of polycystic kidney disease: an estimate of the proportion of families unlinked to chromosome 16. Am. J. Hum Genet. 45: Suppl: A58, abstract.

Reeders, S.T., M.H. Breuning, K.E. Davies, R.D. Nicholls, A.P. Jarman, D.R. Higgs, P.L. Pearson, and D.J. Weatherall. 1985. A highly polymorphic DNA marker linked to adult polycystic kidney disease on chromosome 16. Nature 317: 542-544.

Reeders, S.T., M.H. Breuning, G. Corney, S.J. Jeremiah, P. Meera Khan, K.E. Davies, D.A. Hopkinson, P.L. Pearson, and D.J. Weatherall. 1986. Two genetic markers closely linked to adult polycystic kidney disease on chromosome 16. British Medical Journal 292: 851-853.

Renwick, J.H. and S.D. Lawler. 1955. Genetic linkage between ABO and Nail-Patella loci. Ann. Hum. Genet. 19: 312-331.

Rogstad, S.H., B.L. Herwaldt, P.H. Schlesinger and D.J. Krogstad. 1989. The M13 repeat probe detects RFLPs between two strains of the protozoan malaria parasite Plasmodium falciparum. Nuc. Acids Res. 17: 3610.

Royle, N.J., R. Clarkson, Z. Wong, and A.J. Jeffreys. 1987. Preferential localisation of hypervariable minisatellites near human telomeres. Cytogenet. Cell Genet. 46: 685-686.

Royle, N.J., R.E. Clarkson, Z. Wong, and A.J. Jeffreys. 1988. Clustering of hypervariable minisatellites in the proterminal regions of human autosomes. Genomics 3: 352-360.

Saiki, R.K., D.H. Gelfand, et al. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239: 487-491.

Simmler, M.C., C. Johnson, C.Petit, F. Rouyer, G. Vergnaud, and J. Weissenbach. 1987. Two highly polymorphic minisatellites from the pseudoautosomal region of the human sex chromosomes. EMBO 6: 963-969.

Shen, L.-P. and W.J. Rutter. 1984. Sequence of the human somatostatin I gene. Science 224: 168-171.

Spandidos, D.A. and L. Holmes. 1987. Transcriptional enhancer activity in the variable tandem repeat DNA sequence downstream of the human Ha-ras1 gene. FEB Letters 218: 41-46.

Steinmetz, M., D. Stephan, and K.F. Lindahl. 1986. Gene organization and recombinational hotspots in the murine major histocompatibility complex. Cell 44: 895-904.

Stoker, N.G., K.S.E. Chea, J.R. Griffin, and E. Solomon. 1985. A highly polymorphic region 3' to the human type II collagen gene. Nucleic Acids Research 13: 4613-4622.

Thein, S.L., A.J. Jeffreys, and H.A. Blacklock. 1986. Identification of post-transplant cell population by DNA fingerprint analysis. Lancet ii: 37.

Thein, S.L., A.J. Jeffreys, H.C. Gooi, F. Cotter, J. Flint, N.T.J. O'Connor, D.J. Weatherall, and J.S. Wainscoat. 1987. Detection of somatic changes in human cancer DNA by DNA fingerprint analysis. Br. J. Cancer 55: 353-356.

Viering, T.P. and D.P. Fine. 1989. Genetic Analysis of Streptococcus pneumoniae serotypes with the use of DNA fingerprinting. J. Infect. Dis. 160: 76-82.

Wahls, W.P., L.J. Wallace, and P.D. Moore. 1990. Hypervariable minisatellite DNA is a hotspot for homologous recombination in human cells. Cell 60: 95-103.

Wolff, R.K., Y. Nakamura, and R. White. 1988. Molecular characterization of a spontaneously generated new allele at a VNTR locus: No exchange of flanking DNA sequence. Genomics 3: 347-351.

Wolff, R.K., R. Plaetke, A.J. Jeffreys, and R. White. 1989. Short Communication: Unequal crossingover between homologous chromosomes is not the major mechanism involved in the generation of new alleles at VNTR loci. Genomics 5: 382-384.

Wong, Z., V. Wilson, I. Patel, S. Povey, and A.J. Jeffreys. 1987. Characterization of a panel of highly variable minisatellites cloned from human DNA. Ann. Hum. Genet. 51: 269-288.