DNA Fingerprints in Health and Disease

Certain highly variable regions of human DNA can be exploited to produce a pattern of DNA fragments that is as specific to an individual as is a set of fingerprints

WHAT started out as a more or less routine investigation into the sequence structure of an intron in the human myoglobin gene quite unexpectedly has led during the past year to the development of genetic markers that are as unique to each individual as are the whorls on the finger pads. Appropriately, therefore, Alec Jeffreys and his colleagues at Leicester University, England, who invented the technique, call the markers "DNA fingerprints."

The probability of unrelated individuals sharing the same DNA fingerprint is a minuscule 5×10^{-19}, which for all practical purposes is zero. Even for siblings the statistical gulf is still enormous, with a probability of one in 100 million that they will share precisely the same pattern of markers.

With a power of discrimination of this magnitude, it is not surprising that forensic scientists view DNA fingerprinting with eager anticipation. "It has long been the ambition of the forensic scientist to be able to identify the origin of blood and body-fluid stains with the same degree of certainty as fingerprints," says Barbara Dodd of the London Hospital Medical College. She predicts that "DNA fingerprinting will revolutionize forensic biology."

Indeed, Jeffreys and his colleagues have already demonstrated that their technique can work with 4-year-old bloodstains and can separate the DNA fingerprint in sperm nuclei from the cellular debris of vaginal fluid, both of which can be important in forensic investigations. But the real power of the technique is that for the first time it offers the possibility of positive identification through genetic tests, not just exclusion of identity.

The potential practical applications of DNA fingerprinting outside forensic science are many, and include determining whether same-sex twins are identical or dizygotic and establishing paternity and other familial relationships (see box). For human geneticists, however, the technique offers an attractive method for seeking the location of defective genes in inherited diseases, a task that requires good markers with which to map the genome. The DNA fragments that make up the fingerprints are "the best genetic markers so far characterized for the human genome," comments Ray White of the Howard Hughes Medical Institute at the University of Utah.

For genetic markers to be useful they must be readily identifiable by routine laboratory techniques and, most important, they must be highly variable. The more variability a marker displays, the more informative it can be in comparisons between different individuals. Human genetics is currently undergoing a revolution, primarily because of the discovery of the first marker of this sort half a dozen years ago and the steady accumulation of more as techniques for finding them improve. Jeffreys' discovery is part of that revolution.

"DNA fingerprinting will revolutionize forensic biology."

Most genetic markers are restriction fragment length polymorphisms, of which several hundred have so far been discovered. When genomic DNA is chopped up with one or more restriction enzymes a set of fragments of various sizes is produced. If every individual was genetically identical, the same pattern would be produced. However, because mutations occasionally destroy or otherwise alter the site at which a restriction enzyme cuts, sometimes fragments of different sizes are generated. Hence the term restriction fragment length polymorphism, or RFLP, which is typically pronounced "ruflup." With a radioactively labeled probe of a DNA fragment in question, it is easy to check if an individual possesses the normal form or, say, the longer form, simply by running a gel.

Unfortunately most RFLP's come in just two forms, and are therefore not very informative for genetic analysis. For instance, a maximum of only 50% of the population would be heterozygous for such a marker, and 25% would be homozygous.

Markers start to become really useful when they have half a dozen or so different forms. Given such a polymorphic marker, it is possible to establish whether or not a particular locus is associated with the inheritance of a genetic disease, simply by comparing restriction fragment data from many different, affected families. By means of this approach it has now been possible to pinpoint the general genomic location of several inherited diseases, the first of which was Huntington's disease.

Jeffreys' fingerprint technique also gives a way of tracking down disease loci, but there are differences as well as similarities with the conventional RFLP approach.

For instance, although the DNA fragments that constitute the fingerprint are RFLP's, their lengths vary because of a variation in the number of short repeated sequences within the fragments, not through an alteration of the restriction enzyme cutting sites.

Second, compared with conventional RFLP's, which represent single markers in search of a disease locus, the fingerprint technique effectively monitors as many as 20 loci simultaneously. This factor, notes Kay Davies, of the John Radcliffe Hospital, in Oxford, England, potentially makes the technique "of great value in mapping disease loci in the human genome."

Third, many of the fragments in the fingerprint are much more variable than conventional RFLP's. In some cases this hypervariability pushes the mean heterozygosity in the population to close to 100%.

Ironically, the extreme variability built into the fingerprint technique means that in searching for disease loci, it is not possible to pool and compare members of different families. The fragment patterns of different families are simply too disparate to allow informative comparison: a single, large family containing several generations is what is required.

The DNA fingerprint technique is based on the existence of regions of DNA that are made up of tandem repeats of short sequences, known as minisatellites, a very large number of which are scattered throughout the human genome. These minisatellites, which Jeffreys and his colleagues first found in an intron in the human myoglobin gene, do not constitute a true "family" of sequences: specifically, they are not directly derived from each other in the way that a family of transposable sequences might be. They are, however, related in the sense that they are based on very similar "core" sequences, which are in the region of 15 bases long and constitute the basis of the repeat unit.

Within the total genomic population of minisatellites, there are several different sets, which are united by sharing the same (or

very nearly the same) core sequence. Jeffreys and his colleagues realized the potential for individual identification offered by the minisatellites and have typically been using two such sets in their fingerprinting work.

Specifically, they use DNA probes, each of which is based on a tandem repeat of one of the core sequences. As with conventional RFLP technology, the probe is made radioactive and is applied to a gel separation of a restriction enzyme digest of an individual's genomic DNA. The probe hybridizes to any fragment that contains members of the particular set of ministatellites represented by the core sequence in the probe.

A good probe can light up as many as 80 fragments, which have been separated according to size on a gel. However, it is the large fragments—between 4 and 20 kilobases long—that are most useful for genetic analysis. The reason for this is that these longer restriction fragments often contain large numbers of repeats and, because of various mechanisms such as unequal crossing-over, are subject to more length variation. In other words, they represent highly polymorphic markers of the sort that human geneticists value most. The number of variants in some cases is as many as 20, which gives a mean heterozygosity of greater than 96%.

In any case, using a single probe, more than a dozen hypervariable minisatellites in the 4- to 20-kb range would be produced for any individual, which represent his DNA fingerprint. The probability that another, unrelated individual would share exactly the same pattern is 3×10^{-11}. Add the products of a second probe, and the probability shrinks further, to 5×10^{-19}.

In their initial exploration of these hypervariable regions, Jeffreys and his colleagues examined the DNA fingerprints of 54 related individuals in an extensive Asian family. Examination of the fragment patterns showed that most of the larger fragments were passed from parents to only some of the offspring, which implies a high degree of heterozygosity. And, with just one exception, all of the fragments in the offspring could be found in one of the parents, and then further to the grandparents, which indicates stable Mendelian inheritance of these markers. The single exception was, apparently, a new variant, which on calculation implied a mutation frequency of 0.004.

As yet no inherited disease has been tracked down in the short time that DNA fingerprinting has been available, but Jeffreys and his colleagues are currently collaborating on such a pursuit. Again, the approach is similar to a conventional RFLP search, except that a single minisatellite probe can effectively scan as many as 20 loci simultaneously. With 20 such probes the entire genome could be surveyed in some detail. Once one of the hypervariable fragments is shown to be reasonably close to the disease locus it can be cloned out in the normal way and exploited as a single marker, thereby allowing data from many families to be used.

Jeffreys declines to say to which disease he and his colleagues are applying their technique, but says that "the initial results are encouraging." ■ ROGER LEWIN

A Matter of Maternity

Not long after Alec Jeffreys and his University of Leicester colleagues demonstrated that certain sections of the genome offered a way of generating individual-specific DNA fingerprints, they were approached by a lawyer involved in an immigration case. A Ghanaian boy born in the United Kingdom had emigrated to Ghana to be with his father. Later he decided to return to Britain to rejoin his mother, brother, and two sisters. The immigration authorities suspected that a substitution had been made. Standard blood group and other genetic tests showed a very high probability that the woman and boy were related, but could not exclude the possibility that the woman was the boy's aunt. The authorities therefore refused residence to the boy. The boy's lawyer turned to Jeffreys for help.

Jeffreys agreed to take on the challenge, knowing that the case was about as difficult as it could be. For instance, neither the boy's father nor any of the mother's sisters were available for testing. And, although the mother claimed to be certain that the boy was her son, she was uncertain of the paternity.

The first task was to establish the paternity of the boy. By using two probes, 33.15 and 33.6, on the products of one restriction enzyme, *Hinf* I, Jeffreys and his colleagues generated 80 or so bands in each of the DNA fingerprints of the boy, his putative mother, brother, two sisters, and an unrelated individual. Because some of these DNA markers often approach 100% heterozygosity, the boy's paternity could be inferred by pulling out from the siblings' fingerprints those DNA fragments that were absent in the mother's. Of the 39 bands identified as paternal-specific, about half were to be found in the boy's pattern. "Since DNA fragments are seldom shared between the DNA fingerprints of unrelated individuals, this suggests very strongly that [the boy] has the same father as [the brother and two sisters]," concluded Jeffreys and his colleagues.

What of the boy's relationship to the woman? Removing the paternal-specific fragments from the boy's fingerprint leaves an additional 40, each of which was present in the woman's print. Again, this appeared to be strong evidence that the boy was the woman's son. But what is the probability this might still be wrong? There were 25 DNA fragments shared by the woman and the boy that could be scored confidently. Now, the probability of two unrelated people sharing one fragment is 0.26. So the probability of being unrelated and yet sharing them all is 0.26^{25}, or 2×10^{-15}. Clearly, the boy and the woman must be related.

But what is the probability that the woman is the boy's aunt? If the probability of two unrelated people sharing a single marker is 0.26, then for two siblings it is shortened to 0.62. The probability that the woman is in fact a sister of the boy's mother, yet shares all 25 of the maternal-specific bands, is 0.62^{25}, or 6×10^{-6}. The boy is therefore the woman's son "beyond any reasonable doubt," concluded Jeffreys and his colleagues. Faced with this evidence, the authorities dropped their case against the boy and allowed him to stay with his mother in Britain.

As it happened, the case had been simplified because the boy shared the same father with his siblings. Had another man sired one or more of the siblings, the paternity-specific reconstruction would have become progressively weaker. In addition, the father apparently did not pass on a band solely to the boy and not to the siblings. On average, one in 16 paternal bands might have gone only to the boy, which would have begun to make the analysis less solid, but not substantially so. For instance, if the boy had had five such bands, the chances against the mother/son relationship would have been less than one in 100. "This analysis is therefore robust," concludes Jeffreys, "and would give clear evidence for or against claimed relationship in most such cases." ■ **R.L.**

ADDITIONAL READING

A. J. Jeffreys *et al.*, "Individual-specific 'fingerprints' of human DNA," *Nature (London)* **316**, 76 (1985).

A. V. S. Hill, "Use of minisatellite DNA probes for determination of twin zygosity at birth," *Lancet* **1985-I**, 1394 (1985).