PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Diagnosis of Human Heritable Defects by Recombinant DNA Methods

C. T. Caskey, R. A. Gibbs, J. A. Witkowski and J. F. Hejtmancik

Phil. Trans. R. Soc. Lond. B 1988 **319**, 353-360
doi: 10.1098/rstb.1988.0056

Phil. Trans. R. Soc. Lond. B **319**, 353–360 (1988)
Printed in Great Britain

Diagnosis of human heritable defects by recombinant DNA methods

By C. T. Caskey, R. A. Gibbs, J. A. Witkowski and J. F. Hejtmancik
Howard Hughes Medical Institute and Institute for Molecular Genetics, Baylor College of Medicine, Houston, Texas 77030, *U.S.A.*

Recombinant DNA methods provide highly sensitive means for the detection of DNA alterations that lead to human disease mutations. In this paper I shall illustrate the approaches currently available and discuss new technologies that show promise of replacing the present methods. Medical diagnosis by means of recombinant DNA methods has an expanding role in clinical medicine.

Common disease alleles can arise by expansion of a rare mutational event within a population. The mechanism for the selective advantage of the disease allele is usually not known. However, in the cases of sickling haemoglobinopathies and β thalassaemia, multiple studies support the concept that resistance to infection by the *Plasmodium falciparum* malaria parasite leads to heterozygote advantage (Motulsky 1975). The linking of a polymorphism with such a common disease allele can lead to a method for diagnosis and carrier detection. Kan & Dozy (1978) were the first to discover a restriction fragment length polymorphism (RFLP) and use it for diagnosis of sickle-cell anaemia. Using the restriction endonuclease *Hpa*I, they found a 13.0 kilobase (kb) RFLP highly associated with β^s but not with β^a alleles. This linkage disequilibrium was intensively studied by Chakravarti and Kazazian (Chakravarti *et al.* 1984) and used to classify different β thalassaemia mutations on the basis of their differing RFLP haplotypes. There is an *Ava*II RFLP which has strong linkage disequilibrium with the Z allele responsible for α-1-antitrypsin deficiency, and this RFLP has been used clinically for prenatal diagnosis of this disease (Hejtmancik *et al.* 1986*a*). Even more recently, significant linkage disequilibrium for RFLPs identified with a putative cystic fibrosis cDNA (Estivill *et al.* 1987) has suggested that the cystic fibrosis mutation is a single common allele. A summary of diseases with linkage disequilibrium RFLPs useful for diagnosis is given in table 1.

Table 1. RFLP linkage disequilibrium of common alleles

disease	alleles
sickle-cell anaemia	A, B and C
β thalassaemia	I–IX
α-1-antitrypsin	M and Z
phenylketonuria	wild type and IVS12
cystic fibrosis	wild type and CF

It is possible to detect mutations directly as the loss or gain of a restriction endonuclease recognition site. Unfortunately, this is not a generally useful method because there are only some 125 known restriction endonucleases, each with a different 4–6-base specific recognition sequence (Roberts 1985). This gives a low likelihood that, on a random search basis, one can

find a suitable restriction endonuclease that discriminates between wild-type and mutant sequences. The exception to this rule is the endonuclease *Taq*I, which recognizes the 4–6-base sequence TCGA. Loss of *Taq*I sites at the hypoxanthine phosphoribosyltransferase (HPRT) (Wilson *et al.* 1986), haemophilia A (Youssoufian *et al.* 1986), and ornithine transcarbamylase (OTC) gene loci (Nussbaum *et al.* 1986), are known to result from mutations in the TCGA sequence. There is ample evidence to suspect that CpG sequences are prone to mutations, because they can undergo methylation and thus perturb base pairing (Wallace *et al.* 1986). No other restriction endonucleases have been identified that are as useful as *Taq*I for randomly searching for newly arising point mutations. However, if a disease usually results from a mutation for which extensive sequence information is available, it may be possible to predict an alteration in a restriction endonuclease recognition sequence at the site of the mutation. For example, in the case of sickle-cell disease, the A→T mutational event leads to the loss of an *Mst*II recognition site, i.e. β^s is not cut by *Mst*II whereas β^a is (Chang & Kan 1982). There are no other examples of disease diagnosis by this strategy; this lack indicates the severe limitations of this approach.

Alternative strategies have been developed for the detection of single DNA base changes known to be associated with disease alleles. Synthetic oligonucleotides (19 residues) have been prepared to correspond to either the normal or the mutant DNA sequences. Such allele-specific oligonucleotides (ASOs) have permitted the accurate detection of β^a, β^s and β^c alleles by differential hybridization, but the procedure has suffered the technical problem of high non-specific binding resulting in high background (Conner *et al.* 1983). The ASO detection method was markedly enhanced by the development of the polymerase chain reaction (PCR), which was applied to β^a, β^s and β^c detection by Saiki *et al.* (1986). Two oligonucleotides that hybridized to sequences flanking the region containing both the β^s and β^c mutations were synthesized. These served as primers in successive rounds of annealing and extension across the mutation sites, resulting in a 200000-fold amplification of the region between the oligonucleotides. The genotype of the amplified samples was determined in a simplified dot–blot analysis without time-consuming gel electrophoresis or Southern transfers. Furthermore, DNA from as few as 100 cells was found to be adequate for initiating the PCR. Figure 1 illustrates the progressive improvements of recombinant DNA diagnostic methods for β^s.

The detection of newly mutated alleles is a common diagnostic challenge for a number of diseases, including the Lesch–Nyhan syndrome, OTC-deficiency, haemophilia and Duchenne muscular dystrophy (DMD). We have chosen to develop methods for the rapid detection of point mutations using the HPRT gene locus as a model system. It was known that 85% of HPRT patients with new mutations had no alterations that were detectable by Southern or Northern analysis and were therefore likely to have point mutations (Yang *et al.* 1984). Because the HPRT genomic sequence spans approximately 4000 bp, whereas the protein coding region is only 651 bp, we elected to develop mRNA-based detection methods (Stout & Caskey 1985). Our studies of patients had suggested that the majority of point mutations were in the coding or transcribed sequences. We modified the ribonuclease A (RNAse A) cleavage method of Myers (Myers *et al.* 1985) with the objective of detecting these point mutations. The method is schematically represented in figure 2. The probe consists of a radioactive antisense single-strand RNA synthesized *in vitro* from a plasmid vector. After hybridization of the radioactive probe to unfractionated cellular RNA, the radioactive hybrid complex was isolated on polyuridylic acid-bound affinity paper (mAP; Amersham), thus reducing background

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

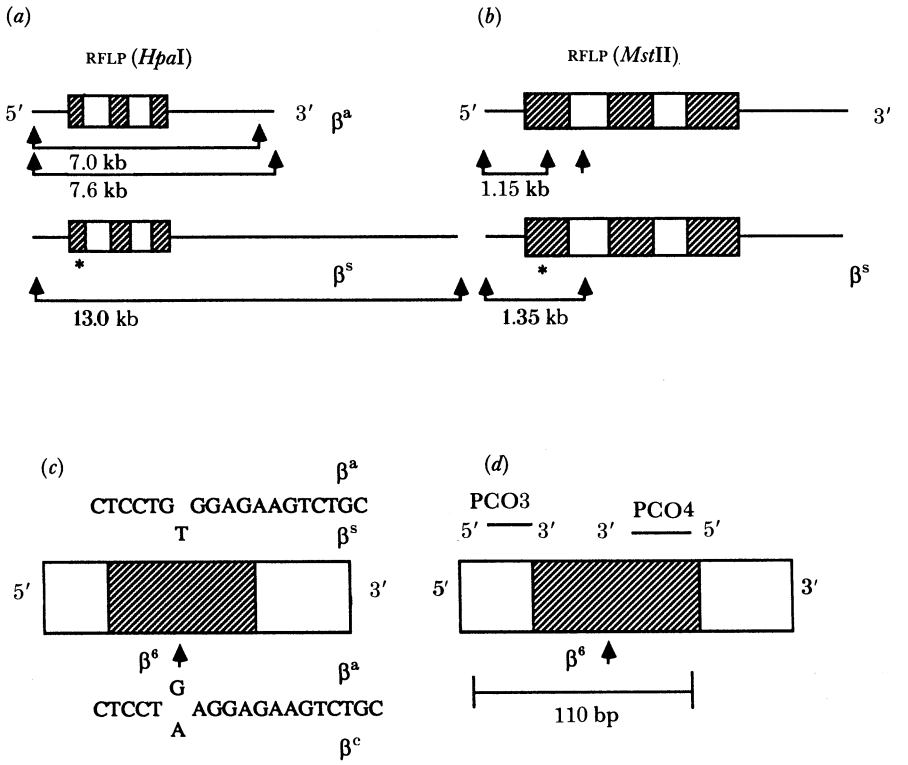

FIGURE 1. Progression of β^s DNA diagnosis. (*a*) *Hpa*I identification of linkage association of the 13.0 kb RFLP with β^s. (*b*) *Mst*II cleavage at codon 6 of β^a, but not β^s, leading to mutation-dependent RFLPs. (*c*) Allele-specific oligodeoxynucleotide (ASO) recognition of β^a, β^s and β^c alleles at codon 6 (β^6). (*d*) Polymerase chain reaction (PCR) primed by PCO3 and PCO4 including the β^6 within the amplified 110 bp (broken line). Exons are indicated by hatched boxes; introns are indicated by open boxes. The asterisk indicates the site of β^s mutation in exon 1. (From Caskey (1987).)

radioactivity from the excess probe. After elution from mAP, the hybrid is subjected to RNAse A cleavage and the fragments resolved by polyacrylamide gel electrophoresis and detected by autoradiography. The method has been successful in identifying and localizing the mutation in 5 of 14 Lesch–Nyhan patients who had been previously classified as having no change by Southern or Northern analysis (Gibbs & Caskey 1987). The RNAse A cleavage procedure therefore offers a one-step scanning method for mutations in the peptide coding region of the HPRT gene. It is possible that with modification of the technique, employing sense single-strand RNA as probe with cDNA synthesized *in vitro*, the sites of mutations could be identified in as many as 80 % of patients.

The localization of mutation sites by RNAse A cleavage has greatly facilitated the subsequent determination of precise nucleotide changes by using other techniques. A method which combines the PCR and RNAse A cleavage is represented schematically in figure 3. Beginning with mRNA, we have synthesized complementary DNA with random hexanucleotides as primers. This more stable source of nucleic acid was then used for PCR amplification of the region containing the RNAse A cleavage site, with two synthetic oligonucleotides which overlapped naturally occurring six-base restriction endonuclease recognition sites. After the PCR, the amplified sequences were cleaved with the appropriate endonucleases, cloned into M13, and subsequently sequenced. This method has been used to

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

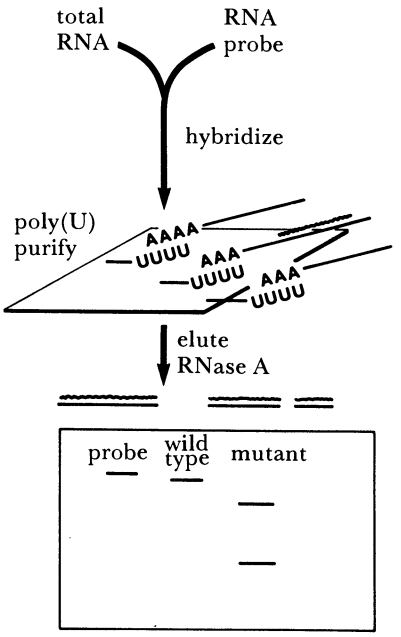

FIGURE 2. Flow diagram for RNase A cleavage analysis with mAP purification. Total cellular RNA extracts were hybridized with radiolabelled RNA probe and the HPRT mRNA; probe hybrids were partly purified by passage through polyuridylic acid [poly(U)]-bound affinity paper. The hybrids were eluted and treated with RNase A to digest single-stranded regions and internal mismatch sites, and the fragments were analysed by denaturing polyacrylamide gel electrophoresis and radioautography. A polyacrylamide gel is shown in schematic form to illustrate the appearance of the probe alone, the probe hybridized to wild-type HPRT mRNA, and the probe hybridized to a mutant HPRT mRNA that yields two fragments. (From Gibbs & Caskey (1987).)

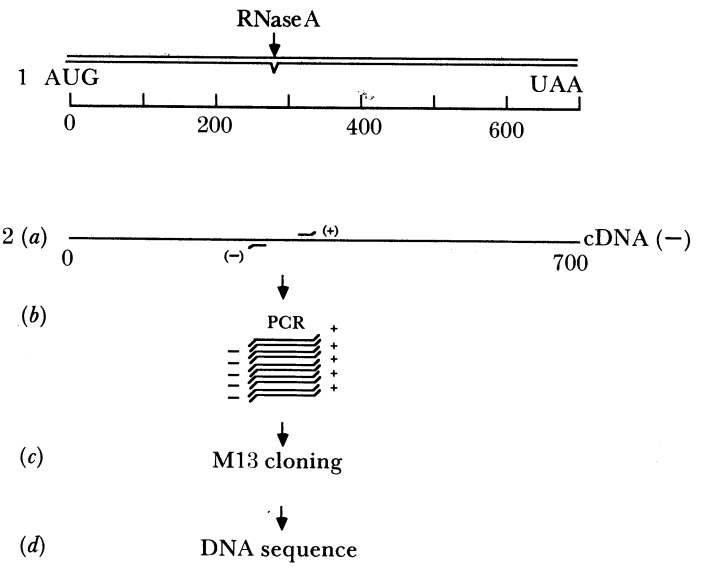

FIGURE 3. Scheme for rapid identification of point mutations in expressed genes. (1) Ribonuclease A (RNase A) cleavage is used to identify and localize a mismatch site, between a wild-type, antisense RNA probe synthesized *in vitro*, and the mutant mRNA (see text). (2*a*) Oligonucleotide primers, which flank the RNase A sensitive site and either overlap natural restriction endonuclease recognition sites or contain 5′ tails with enzyme recognition sequences, are employed as primers in the polymerase chain reaction to amplify the region of interest. (*b*) The amplified sequences are cleaved with the appropriate restriction endonucleases; (*c*) ligated into DNA sequencing vectors; and (*d*) the precise DNA sequence is determined.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

identify the OTC point mutation in the OTC-deficient sparse fur (*spf*) mouse (Veres *et al.* 1987), but has general applicability. These studies, although simple, are highly dependent on sequence knowledge of the gene undergoing mutation.

As detailed DNA sequence information is available for only a few human disease genes, linkage analysis with genetic markers that show tight genetic linkage to the disease remains the preferred, and only, diagnostic approach. Despite the need for cooperation of multiple family members and possible errors arising from non-paternity, this method has proved very powerful. It has been applied with the greatest experience and success in carrier detection and prenatal diagnosis of DMD (Hejtmancik *et al.* 1986*b*; Witkowski & Caskey 1988).

The DMD gene maps to the Xp21 region of the X chromosome. Probes close to and included in the DMD locus have been isolated by many but principally by the laboratories of Kunkel (Monaco & Kunkel 1987), Worton (Ray *et al.* 1985), Davies (Davies *et al.* 1983), and Mandel and Pearson (Hofker *et al.* 1985). These collective efforts provided RFLP-associated Xp21 probes that are within and flank the DMD locus. The gene is known to be 2×10^6 bp at a minimum (Koenig *et al.* 1987) and is prone to both new mutation and recombination.

Our programme has served as a referral laboratory for carrier detection and prenatal diagnosis of DMD by means of genetic linkage analysis. This programme has a genetic counsellor, who acts as an intermediary between the referring centre and our laboratory; a medical director to interpret intake clinical data and perform linkage interpretation; a laboratory director, who is responsible for staff (three technicians) and quality control of laboratory procedures, and a secretary familiar with computer-based reporting and maintaining a database. The laboratory studied 286 families with DMD or Becker's muscular dystrophy and 156 pregnancies over the past 20 months. There were approximately 20 cases in which family members needed for a study refused to cooperate, and some 15 cases of incorrect paternity were identified.

The application of linkage to diagnosis of carriers of DMD is well illustrated in figure 4 (*a*, *b*). It is clear that carrier detection was most accurate for females within families with multiple affected males. In families where the disease has occurred as an isolated event, linkage analysis is less useful. The identification of the gametic origin of the new mutational event is critical to the accurate carrier determination. This mutational analysis is severely limited by a small family size; in our family studies we were able to identify the origin of 37 mutational events of 135 studied.

The application of linkage to prenatal diagnosis (PND) of DMD is medically more successful, because one is trying to render a disease risk assessment for the pregnancy. Because the pregnancy is an additional generation, one can use linkage to clarify pregnancy disease risk even in the setting of a mother whose carrier risk is uncertain. For example, the identification of a female or male fetus carrying a grandpaternal rather than grandmaternal haplotype predicts that the pregnancy risk is low for DMD. The summary in table 2 indicates remarkably good success in the application of PND for DMD via linkage. An estimated 50% of mothers requesting PND were found to have low carrier risk and thus could avoid PND. The remainder required detailed linkage study for accurate diagnosis, and we are still accumulating data on the outcome of these 52 pregnancies with male fetuses. Thus far, we have follow-up data on 25 of the 36 males born and there have been two DMD males carried to term. The first was known to have 70% risk of disease but was complicated by a pregnancy twinning with a female fetus. The second was a case of misdiagnosis due to failure of linkage analysis. The pregnancy was

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

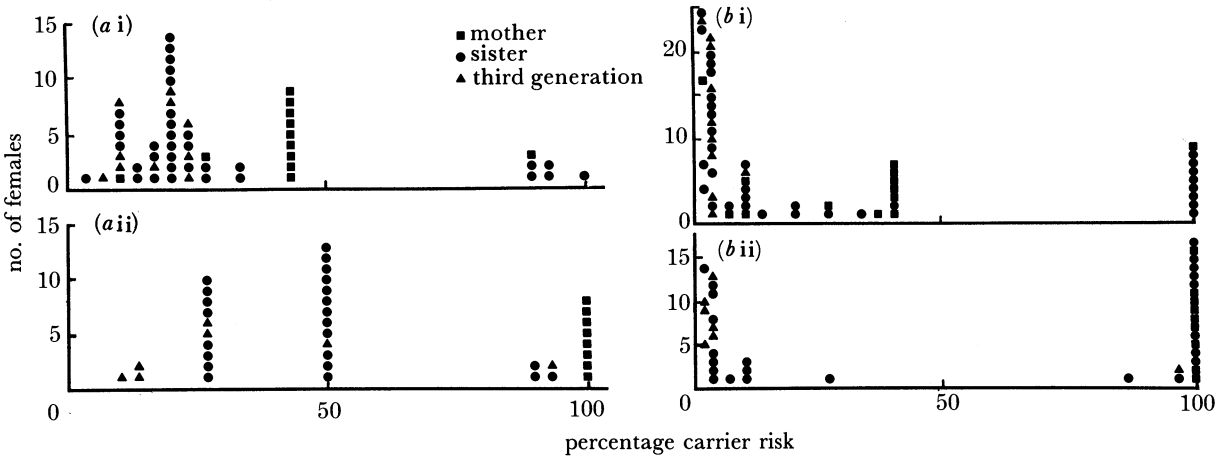

FIGURE 4. Effect of RFLP analysis on carrier risk assessment in DMD. (*a*) The carrier risk based on pedigree and serum creatine kinase information; (*b*) the carrier risks of the same women after incorporating data from linkage analysis with RFLPs. The upper panels (i) show the data for female relatives of isolated cases of DMD; and the lower panels (ii) show data for females in families with a history of DMD.

TABLE 2. CARRIER RISK BEFORE AND AFTER DNA DIAGNOSIS

percentage carrier risk...		less than 5%			5–95%			over 95%		
relative[3]...		M	S	O	M	S	O	M	S	O
familial[1]	before[2]	0	0	0	0	23	7	9	0	0
	after	0	8	6	0	6	0	9	9	1
isolated[1]	before[2]	0	1	0	12	32	10	0	1	0
	after	0	15	9	11	11	1	1	8	0

[1] Familial: families with clear evidence of inheritance of DMD. Isolated: families with a single affected boy.
[2] Before: numbers of women in each category based on pedigree and CK only. After: numbers of women in each category based on DNA analysis, pedigree and CK information.
[3] M: mother of affected boy; S, sister; O, other female relatives.

predicted to be 95% likely to be a normal male, because the nearest marker available for use had 5% recombination with DMD and flanking markers were not informative. There was presumably a crossover in this case that could not be detected. Thus, DNA linkage analysis has proved to be highly reliable and useful for DMD.

Recently, full-length cDNA probes have been made available for DNA diagnosis of mutations at the DMD locus (Koenig *et al.* 1987). These cDNA probes detect deletions in approximately 50% of DMD patients, compared with about 10% detected by using genomic probes. We anticipate that the cDNA probes will prove extremely useful in the clinical situation and we are currently reassessing the status of our isolated cases of DMD, numbering some 60 boys. Our laboratory is now shifting its focus to gene-based diagnostic methods and we have succeeded in utilizing PCR for rapid detection of deletions in this large gene of 2×10^6 bp. We expect that there will be an accelerated application of novel research techniques to the clinical situation, leading to more effective prevention and avoidance of genetic disease.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

C. T. C. is an Investigator with the Howard Hughes Medical Institute; R. A. G. is supported by a postdoctoral fellowship from the Muscular Dystrophy Association; J. A. W. and J. F. H are supported by grants from the Muscular Dystrophy Association and the March of Dimes Birth Defects Foundation.

References

Caskey, C. T. 1987 Disease diagnosis by recombinant DNA methods. *Science, Wash.* **236**, 1223–1229.

Chakravarti, A., Buetow, K. H., Antonarakis, S. E., Waber, P. G., Boehm, C. D. & Kazazian, H. H. 1984 Non-uniform recombination within the human betaglobin gene cluster. *Am. J. hum. Genet.* **37**, 984–997.

Chang, J. C. & Kan, Y. W. 1982 A sensitive new prenatal test for sickle-cell anemia. *New Engl. J. Med.* **307**, 30–32.

Conner, B. J., Reyes, A. A., Morin, C., Hakura, K., Teplitz, R. L. & Wallace, R. B. 1983 Detection of sickle-cell β^s-globin allele by hybridization with synthetic oligonucleotides. *Proc. natn. Acad. Sci. U.S.A.* **80**, 278–282.

Davies, K. E., Pearson, P. L., Harper, P. S., Murray, J. M., O'Brien, T. O., Sarfarazi, M. & Williamson, R. 1983 Linkage analysis of cloned DNA sequences flanking the Duchenne muscular dystrophy locus on the short arm of the human X chromosome. *Nucl. Acids Res.* **11**, 2303–2312.

Estivill, X., Farrall, M., Scambler, P. J., Bell, G. M., Hawley, K. M. F., Lench, N. J., Bates, G. P., Kruyer, H. C., Frederick, P. A., Stanier, P., Watson, E. K., Williamson, R. & Wainwright, B. J. 1987 A candidate for the cystic fibrosis locus isolated by selection for methylation-free islands. *Nature, Lond.* **326**, 840–845.

Gibbs, R. A. & Caskey, C. T. 1987 Identification and localization of mutations at the Lesch–Nyhan locus by ribonuclease A cleavage. *Science, Wash.* **236**, 303–305.

Hejtmancik, J. F., Ward, P. A., Mansfield, T., Sifers, R. N., Harris, S. & Cox, D. W. 1986*a* Prenatal diagnosis of α1-antitrypsin deficiency by restriction fragment length polymorphisms, and comparison with oligonucleotide probe analysis. *Lancet* ii, 767–770.

Hejtmancik, J. F., Harris, S. G., Tsao, C. C., Ward, P. & Caskey, C. T. 1986*b* Carrier diagnosis of Duchenne muscular dystrophy using restriction fragment length polymorphisms. *Neurology* **36**, 1553–1562.

Hofker, M. H., Wapenaar, M. C., Goor, N., Bakker, E., Van Ommen, G. B. & Pearson, P. L. 1985 Isolation of probes detecting restriction fragment length polymorphisms from X-chromosome-specific libraries: potential use for diagnosis of Duchenne muscular dystrophy. *Hum. Genet.* **70**, 148–156.

Kan, Y. W. & Dozy, A. M. 1978 Polymorphism of DNA sequence adjacent to the human β-globin structural gene: relationship to sickle mutation. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5631–5635.

Koenig, M., Hoffman, E. P., Bertelson, C. J., Monaco, A. P., Feener, C. & Kunkel, L. M. 1987 Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. *Cell* **50**, 509–517.

Monaco, A. P. & Kunkel, L. M. 1987 A giant locus for the Duchenne and Becker muscular dystrophy gene. *Trends Genet.* **3**, 33–37.

Motulsky, A. G. 1975 Glucose-6-phosphate dehydrogenase and abnormal hemoglobin polymorphisms – evidence regarding malarial selection. In *The role of natural selection in human evolution* (ed. F. Salzaro), pp. 271–291. Amsterdam: North-Holland.

Myers, R. M., Larin, Z. & Maniatis, T. 1985 Detection of single base substitutions by ribonuclease cleavage at mismatches in RNA:DNA duplexes. *Science, Wash.* **230**, 1242–1246.

Nussbaum, R. L., Boggs, B. A., Beaudet, A. L., Doyle, S., Potter, J. L. & O'Brien, W. E. 1986 New mutation and prenatal diagnosis in ornithine transcarbamylase deficiency. *Am. J. hum. Genet.* **38**, 149–158.

Ray, P. N., Belfall, B., Duff, C., Logan, C., Kean, V., Thompson, M. W., Sylvester, J. E., Gorski, J. L., Schmickel, R. D. & Worton, R. G. 1985 Cloning of the breakpoint of an X;21 translocation associated with Duchenne muscular dystrophy. *Nature, Lond.* **318**, 672–675.

Roberts, R. J. 1985 Restriction and modification enzymes and their recognition sequences. *Nucl. Acids Res.* **13**, r165–r200.

Saiki, R. K., Bugawan, T. L., Horn, G. T., Mullis, K. B. & Erlich, H. A. 1986 Analysis of enzymatically amplified β-globin and HLA-DQα DNA with allele-specific oligonucleotide probes. *Nature, Lond.* **324**, 163–166.

Stout, J. T. & Caskey, C. T. 1985 HPRT: gene structure, expression, and mutation. *A. Rev. Genet.* **19**, 127–148.

Veres, G., Gibbs, R. A., Scherer, S. E. & Caskey, C. T. 1987 The molecular basis of the sparse fur mouse mutation. *Science, Wash.* **237**, 415–417.

Wallace, R. B., Petz, L. D. & Yam, P. Y. 1986 Application of synthetic DNA probes to the analysis of DNA sequence variants in man. *Cold Spring Harb. Symp. quant. Biol.* **51**, 257–261.

Wilson, J. M., Stout, J. T., Patella, T. D., Davidson, B. L., Kelley, W. N. & Caskey, C. T. 1986 A molecular survey of hypoxanthine–guanine phosphoribosyltransferase deficiency in man. *J. Clin. Invest.* **77**, 188–195.

Witkowski, J. A. & Caskey, C. T. 1988 Duchenne muscular dystrophy: DNA diagnosis in practice. In *Current neurology* (ed. S. H. Appel), pp. 1–36. Chicago: Medical Yearbook Publishers.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Yang, T. P., Patel, P. I., Chinault, A. C., Stout, J. T., Jackson, L. G., Hildebrand, B. M. & Caskey, C. T. 1984 Molecular evidence for new mutation at the *hprt* locus in Lesch–Nyhan patients. *Nature, Lond.* **310**, 412–414.

Youssoufian, H., Kazazian, H. H., Phillips, D. G., Aronis, S., Tsiftis, G., Brown, V. A. & Antonarakis, S. E. 1986 Recurrent mutations in hemophilia A give evidence for CpG mutation hotspots. *Nature, Lond.* **324**, 380–382.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES