REVIEW ARTICLE

Detection of single base changes in nucleic acids

Richard G. H. COTTON

Murdoch Institute, Royal Children's Hospital, Melbourne, Victoria 3052, Australia

INTRODUCTION

Ever since the detection of the first mutations in *Escherichia coli*, and subsequently in human DNA (Flavel *et al.*, 1978), enormous efforts have been expended in defining changes in nucleic acid sequences of many organisms. This activity has been pursued in order to: (a) understand mechanisms of disease, (b) understand the structure and function of enzymes and nucleic acids (e.g. tRNA), (c) track micro-organisms in the environment, (d) track disease genes in people for diagnostic purposes, (e) define mutagenic substances, and (f) discover evolutionary relationships between species.

Faced with the tedium and repetitiveness of cloning and sequencing a gene containing a mutation which may take 6 months to define, there has been increasing interest in shortening this process and the last decade has seen some of the fruits of this activity. It has proven difficult to develop methods which will simply detect all or most point mutation types (see below) and hence all or most mutations in a given stretch of nucleic acid. The end point of these studies should be a method which will guarantee detection of all point mutations with a minimum of manipulation of the genetic material before application of a method of extreme simplicity.

This review traces the history and evolution of methodology which has increased the rate of definition of single base mutations. It is not intended as a review of current methods for studying human disease as this is covered elsewhere (Landegren *et al.*, 1988b; Caskey, 1987). Particular methods have been reviewed recently (Myers *et al.*, 1988; Myers & Maniatis, 1986; Lerman *et al.*, 1984). The techniques covered fall into three broad categories: (a) methods for cloning (amplification) and sequencing, (b) methods using hybrid molecules (heteroduplexes) between normal and mutant molecules, and, (c) methods using homoduplex molecules.

Methods for analysis of deletions or insertions will be mentioned only in passing as the larger of these are readily detected.

There is a difference in the methodology depending on whether the position of the mutation is known. When the position of the mutation is unknown the method must screen kilobases of nucleic acid. If the mutation position and type are known the operation is relatively simple. These different categories are addressed.

Isolation of genetic material

Of course the genetic material usually has to be isolated to detect mutations. However, alteration of the phenotype of a cell line can indicate that a mutation has occurred in a particular gene without nucleic acid isolation. One example of this is the detection system which

assays for mutation of the hypoxanthine phosphoribosyltransferase (HGPRTase) gene by alteration of the growth in selective medium (Morley *et al.*, 1983; Albertini *et al.*, 1982). However the aim of this review is to discuss methods allowing description of position and type of mutational change.

The techniques for isolation of nucleic acids have been standard over many years (Maniatis *et al.*, 1982). This applies to microbial DNA, mRNA, viral RNA and DNA and eukaryotic DNA. If larger quantities of these molecules or their copies (in the case of RNA) are needed for mutation characterization, amplification has usually been by cloning into plasmids and other vectors (Maniatis *et al.*, 1982). However, this amplification has been achieved one or two orders of magnitude more quickly by a different approach which relies on enzymic rather than plasmid amplification of nucleic acids and is called the polymerase chain reaction (PCR) (Saiki *et al.*, 1986). This technique has already had a dramatic impact on the time needed to characterize mutations (see below). In simple terms, synthetic oligonucleotides are made complementary to the 5' ends of the sense and antisense strands of the target nucleic acid and a polymerase enzyme is used to synthesize multiple copies of the target by repeated cycles of annealing and synthesis in the presence of excess oligonucleotides. This amplification can be made very efficient by application directly to genomic DNA in a cell extract (Saiki *et al.*, 1986). Other enzymes can be used for amplification (reviewed in Keohavong *et al.*, 1988) and it is claimed that modified T7 DNA polymerase ('Sequenase') will allow amplification of a 2 kb stretch which could not be amplified by the widely used *Taq* polymerase (Keohavong *et al.*, 1988).

However, the simplest methods are those which can be applied directly to the isolated genetic material and several examples of these are available (see below and Table 1). The quest for such direct methods will no doubt continue in order to conserve time and expenditure. Another reason is because the PCR method can introduce errors (Saiki *et al.*, 1988a; Dunning *et al.*, 1988); thus it is wise to analyse at least two separate amplified products from the nucleic acid of interest.

Types of mutation characterization methods

Almost without exception the ultimate complete characterization of the mutation will rely on sequencing which is based on strategies developed over 10 years ago (Maxam & Gilbert, 1977; Sanger *et al.*, 1977). This will define the new base and its exact position in the gene. Thus this is the ultimate reference method. However, because of its time-consuming nature, workers have attempted to avoid sequencing kilobases of mutant gene

Table 1. Screening methods for detection of point mutations

Method	DNA type	DNA stage achieved	Mutation position defined	Percentage mutations detected	Duplex	Maximum bases screened/test	Strongest point	Weakest point
Sequencing	Homoduplex	Genomic	Yes	100	DNA	200-400	Position and change defined	Time-consuming
Restriction enzymes	Homoduplex	Genomic	No	< 50	DNA	Tens of kb	Homoduplex method	Limited coverage; expensive
Denaturing gradient gels	Homoduplex	Genomic	No	50-100	DNA	1000 to tens of kb	Homoduplex method; isolates mutant	Special equipment
Denaturing gradient gels	Heteroduplex	Genomic	No	50-100	DNA · RNA, RNA · RNA, DNA · DNA	1000	Isolates mutant	Special equipment
Solution melting	Heteroduplex	Genomic	No	to 100	RNA · RNA, RNA · DNA	200-300	Detects mutations in high melting domain; simple	Limited screen length
S1 nuclease	Heteroduplex	Plasmid	Yes	0?	DNA · DNA	1000		May not detect point mutations
RNAase	Heteroduplex	Genomic	Yes	60-70	RNA · DNA, RNA · RNA	1000	Cleavage method	RNA probe production; limited coverage
Carbodi-imide/ABC nuclease	Heteroduplex	Plasmid	Yes	?100	DNA · DNA	to 1000	Cleavage method; potentially detects all	Two-step method
Carbodi-imide	Heteroduplex	Plasmid	No	?100	DNA · DNA	to 1000	Potentially detects all	Non-cleavage method
Chemical cleavage of mismatch	Heteroduplex	Plasmid	Yes	?100	DNA · DNA, DNA · RNA	1000	Cleavage method; may detect all; mutation position defined	Toxic reagent

Table 2. Diagnostic methods for detecting mutations

Method	Current use	Strongest point	Weakest point
Sequencing with PCR	Thalassaemias		Sequencing time-consuming and expensive
Restriction enzymes	Widespread in linkages	Genomic; direct	Electrophoresis needed
Restriction enzymes (with PCR)	Thalassaemias		
Oligonucleotides	Widespread	Simple as dot blot	
Ligation	New	Simple, non-radioactive	
Allele-specific amplification	New	Simple; non-radioactive	
Sandwich	New	Simple	

which may only differ from wild type or reference type at a single position. Methods were developed to screen hundreds or thousands of base sequences for differences between mutant and wild type genes and are summarized in Table 1. These methods will be referred to as 'screening methods'. Once the base change and its position is known a different set of methods (Table 2), usually based on the principles used in the screening methods, is used to assay for the presence or absence of this mutation in the future. This type of method is often used in the diagnosis of inherited disease and will be referred to as 'diagnostic methods'.

These two classes, particularly the screening methods (Table 1), can be subdivided into those which rely on homoduplex nucleic acid or heteroduplex nucleic acid. The latter method involves mixing mutant and wild type nucleic acid, heating to melt them and then allowing them to reanneal so that some of the product is a heteroduplex. Obviously methods which function on homoduplexes are advantageous as they avoid this step. All methods requiring melting or reannealing of nucleic acid rely on the principles of complementarity of duplexes. In the case of single base changes between stretches of nucleic acid, heteroduplex formation will create a single base pair mismatch. It is on the presence of this structure that many methods rely.

Melting and reannealing of duplexes

It is widely known that the stability of the DNA double helix relies on the complementarity of the bases G·C and A·T and that G·C base pairs are more stable than A·T base pairs (Saenger, 1984; von Hippel & McGhee, 1972). Also it was found that different regions of a random piece of DNA melted at different temperatures which are dependent on base pair content (Wada *et al.*, 1976). Many methods including methods of mutation detection rely on these findings.

When one of the bases of a pair is replaced by a base not usually paired with the complementary base, it was shown as early as 1966 (Szer & Shugar, 1966) that stability of the helix was altered and this was confirmed in the mid 1970s (Engel & von Hippel, 1974, 1978; Gill *et al.*, 1974; Ehrlich *et al.*, 1975). These findings were put on a more systematic basis by Wallace and coworkers who showed that synthetic oligonucleotides of 11, 14 and 17 bases were less stable when they contained a single

base pair mismatch in the duplex (Wallace *et al.*, 1979). The extension and application of these findings are discussed below.

Mismatch structure

The elucidation of mismatch structure (by n.m.r. and crystal studies) has been the target of intensive activity in recent years (Lomant & Fresco, 1975; Kennard, 1988; Patel *et al.*, 1987), particularly since it was realized that mismatches are the target of repair enzymes (Kramer *et al.*, 1984) which can eliminate or fix a mutation by altering a mismatched base pair, created by the mutation, to a matched base pair. The structural studies have focused on those mismatches considered most important as the targets of repair enzymes. Thus G·T (Benevides *et al.*, 1989; Gorenstein *et al.*, 1988; Quignard *et al.*, 1987; Kalnik *et al.*, 1988b), C·A (Hunter *et al.*, 1986, 1987; Kalnik *et al.*, 1988a; Gao & Patel, 1987) and A·G (Prive *et al.*, 1987) mismatches have been most studied. The conclusions from these studies are that: (a) wobble pairing exists in the G·T and C·A mismatches (Kennard,

Table 3. Effect of mismatches on formation and stability of mismatches

Most stable are ranked in order to the least stable found in these studies. Mismatches with equivalent values are bracketed.

Aboul-El <i>et al.</i> (1985)		Wernteges <i>et al.</i> (1986) (ranked by by temperature)	Ikuta <i>et al.</i> (1987) (50 °C)
50 °C	25 °C		
T·G	T·G	G·T	G·T
G·G	G·G	T·G	G·A
G·T	G·A	A·G	T·T
A·G	G·T	T·T	C·A
G·A	A·G	C·T	A·A
A·A	C·T	G·G	C·T
A·C	T·C	G·A	
C·A	A·A	T·C	
C·T	T·T	A·A	
T·C	C·A	A·C	
T·T	C·C	C·A	
C·C	A·C	C·C	

1988; Patel *et al.*, 1987), and (b) that studies in crystals and in solution (n.m.r.) give rise to different conclusions regarding disturbance to the base stacking and the backbone. It appears that little disturbance of surrounding base pairs occurs in crystals but in solution the double helix is disturbed particularly in C·A mismatches (Kennard, 1988; Patel *et al.*, 1987), but this disturbance is not long-range.

Another class of studies has examined the effect of the presence of all possible mismatches both on the stability (Ikuta *et al.*, 1987; Werntges *et al.*, 1986) and re-hybridization rate (Aboul-ela *et al.*, 1985) of heteroduplexes containing them. These studies are summarized in Table 3. It can be seen that generally the G mismatches are the most stable and pyrimidine·pyrimidine, C·A and A·A tend to be the least stable. The different individual rankings between the three studies presumably reflects different sequence contexts between the studies.

MUTATION DETECTION METHODS WHERE THE MUTATION POSITION IS UNKNOWN (SCREENING METHODS)

Homoduplex methods

Sequencing. As mentioned above, this is the final reference method and barely a screening method although at one time this was the only method available. The recent coupling with the PCR reaction (Saiki *et al.*, 1986) has made screening with sequencing much quicker and hence more popular but screening of say 10 kb in this manner would be a considerable undertaking. Mutations have been readily detected following amplification of PCR and subsequent sequencing of exons (Hughes *et al.*, 1988). This method also has well-known artefactual errors as well as operator error. The main rate-limiting step is the transformation of the M13 sequencing vector in the dideoxy method. However direct sequencing of PCR products can now be achieved without the need for M13 subcloning (Wong *et al.*, 1987). If compression is found to make sequence unreadable further sequencing is necessary. However, the recent automation of sequencing (L. M. Smith *et al.*, 1986) promises to improve the rate and quality of sequencing.

Restriction enzymes. As restriction enzymes have a very specific target sequence (usually four to eight bases), a change in this sequence as a result of a mutation will be readily detected by a change from cleavage to non-cleavage (or vice versa) and mutations have been detected in this way (Maddalena *et al.*, 1988; Carothers *et al.*, 1988; Greever *et al.*, 1981). However, as not all mutations alter restriction enzyme sites the coverage of the screen is obviously incomplete. It has been suggested to be less than 50% (Landegren *et al.*, 1988a). Despite this, harmless mutations (polymorphisms or RFLPs) have been detected after screening using a battery of restriction enzymes and used in tracking disease carrying genes in families (Wolfe, 1988; Kan & Dozy, 1978). This practice is widely used today, but the limitation mentioned is evidenced by the fact that often in the diagnosis of a particular disease, informative and hence useful polymorphisms cannot be found in particular families. A more serious problem occurs when polymorphisms for diagnosis may not be detectable at all. For example, a search using 35 enzymes with a pyruvate dehydrogenase

probe revealed no useful polymorphisms (H. Dahl, personal communication).

Restriction enzymes have been effective in studying viral variants (Griffiths & Grundy, 1987; Tenover, 1988) as enough variation occurs between isolates to be sure that differences will be found with a small number of enzymes. However, although a pattern of difference is found between strains this pattern means very little biologically.

Denaturing gradient gels. Fischer & Lerman (1979) found that when DNA homoduplexes were electrophoresed at constant temperature into an increasing concentration of denaturant in the gel, the duplex slowed its migration rate dramatically at a particular denaturant concentration. Later experiments showed fragments could be separated according to properties of their sequence (Fischer & Lerman, 1980). Further, duplexes differing only by a single base pair were shown to be separable on the system (Fischer & Lerman, 1983). The principle underlying this separation has been discussed in various reviews (Myers & Maniatis, 1986; Myers *et al.*, 1987). In simple terms, random DNA fragments of 100–1000 bp contain two regions, one of which melts at a higher temperature than the other (high and low melting domains respectively). Base changes in the low-melting domain cause a change in the denaturant concentration at which the first domain melts. Thus the slowing of migration of the duplex carrying the mutation occurs at a different point in the gradient compared with the reference duplex. Thus the method allows screening of a considerable number of bases for mutation, but has the limitation that mutations in the highest melting domains will be undetectable. However, many strategies have been developed to overcome these limitations. One is to place an artificial very-high-melting domain on one end (Myers *et al.*, 1985a,b). This comprises a GC-rich sequence called a GC clamp and it is suggested that nearly all mutations will be detected. More recently the GC clamps have been added via the oligonucleotides used for PCR amplification of DNA (Sheffield *et al.*, 1989). Another strategy was based on the destabilizing effect of mismatches on the lower melting domain. Heteroduplexes were made containing mismatches (Myers *et al.*, 1985d) and this allows the detection of more mutations. This approach has also been shown to be applicable to DNA·RNA and RNA·RNA heteroduplexes (Smith *et al.*, 1986). Noll & Collins (1987) applied the method directly to genomic DNA to detect polymorphisms. More recently Gray extended the length of the screen with the method and the percentage coverage of mutations. First, genomic *Drosophila* DNA is digested with up to five restriction enzymes. The fragments are run on the denaturing gel system (Fischer & Lerman, 1979) and blotted onto a membrane which is probed with a 5–7 kb probe. This allowed screening of a large stretch of DNA for polymorphisms and the multiple enzymes ensured that most mutations lay in a low-melting domain with at least one enzyme. This allowed 79% single base changes to be detected in a series of *Drosophila* mutants. This technique may not be so easily applied in human systems due to the presence of repetitive elements. However, the system has been applied to the detection of polymorphisms in human factor VIII by using a cDNA probe (M. R. Gray, personal communication).

A similar technique has been developed in parallel over a similar time. In 1979, double-stranded viral RNA was shown to vary in thermal stability (Diaz-Ruiz & Kaper, 1979). Since that time a routine procedure (Rosenbaum & Riesner, 1987) has been developed to examine variability. This involves electrophoresis into a temperature gradient and is thus analogous to the chemical denaturant method above. The method has been applied to the analysis of a variation in cucumber mosaic virus and peanut stunt virus (Po *et al.*, 1987).

One important use of this system is the ability to purify out a minority population of mutant DNA in a population, allowing its isolation and subsequent sequencing for the purposes of characterization.

Heteroduplex methods

Melting methods. The observation by Ito & Joklik (1972) that heteroduplexes between two RNA virus strains differing by a putative single base moved slower in a 7 M-urea gel than a homoduplex must have been an inspiration for the examination of heteroduplexes in a denaturing gradient electrophoresis system (Smith *et al.*, 1986; Myers *et al.*, 1985d) (see above). In fact it was shown clearly that melting of heteroduplexes was much easier than that of homoduplexes (Lerman *et al.*, 1984).

A variation of this technique without the use of a gradient was developed by Jones *et al.* (1985). RNA·DNA heteroduplexes between radiolabelled SP6 RNA and genomic DNA from sickle and normal individuals were shown to be separable on agarose gels.

In an attempt to simplify the methodology, Smith *et al.* (1988) developed the solution melting method. Heteroduplexes were exposed to differing denaturant concentration (salt or formamide) in different tubes which are exposed to the same temperature for a fixed time. After chilling samples, they are electrophoresed to assay for strand dissociation. This has been applied to viral RNA·RNA duplexes (Smith *et al.*, 1988) and RNA·DNA duplexes (Latham & Smith, 1989). It is suggested that this method detects mutations in the high-melting domain (Smith *et al.*, 1988). In this method the RNA probe must be produced by the SP6 expression system (Melton *et al.*, 1984). With RNA·DNA duplexes all eight mismatches tested (A·A, U·T, C·C, G·C, G·A, U·C, U·G and C·A) were detected in a 130 bp-long melting domain (Latham & Smith, 1989).

Enzyme methods: S1 nuclease. Shenk *et al.* (1975) reported cleavage of heteroduplexes of DNA contain putative differences of single bases. Little use was made of this observation; however, Dodgson & Wells (1977) found that cleavage of single base pair mismatches was extremely slow, but when mismatches were of two or more bases the rate of cleavage was greatly increased. Thus it appears that this method is not suitable for detecting single base changes.

Ribonuclease A. It had been found earlier (reviewed in Brownlee, 1972) that many ribonucleases were found specific for single-stranded RNA. Inspired by the findings of Shenk *et al.* (1975), Freeman & Huang (1981) formed duplexes between genomic RNA of vesicular stomatitis virus and mRNA of various strains and their revertants. Digestion with RNAase A, T1 and T2 in many cases revealed cleavage at the presumed points of mismatch due to putative single base differences between the

Table 4. Cleavage of single-base-pair mismatches in RNA·DNA heteroduplexes by ribonuclease A

Adapted from Myers *et al.* (1985c). The base in the RNA is given first in the mismatch.

Mismatch	Number tested	Number cleaved
C·A	22	22
C·C	4	4
C·T	9	9
U·G	11	7
U·C	3	3 (poorly)
U·T	4	4
G·A	7	1
G·T	14	1
G·G	6	0
A·C	12	4
A·A	6	2
A·G	3	2

strands. However, in some cases there was no digestion and it was presumed that particular mismatches were resistant to cleavage (see below).

The development of the SP6 system for RNA transcript production (Melton *et al.*, 1984) allowed a much readier examination of the ribonuclease system. Winter *et al.* (1985) and Myers *et al.* (1985c) reported further studies with RNA·RNA heteroduplexes and RNA·DNA heteroduplexes using RNAase A (and RNAase T1 in the former case). Of the four mismatches studied by Winter *et al.* (1985), U·C, A·G and A·A were cleaved well and U·U less well. They reported a similar pattern in RNA·DNA duplexes for the equivalent mismatch, except U·T mismatches were not cleaved. Myers *et al.* (1985c) studied a total of 101 single base pair mismatches in RNA·DNA heteroduplexes. Their findings are summarized in Table 4. It is quite clear that many mismatches were not cleaved, particularly where a purine appeared in the RNA. These data led the authors to conclude that four of 12 possible mismatches (i.e. C·A, C·C, C·T and U·T) could be detected efficiently, and thus 60% of all mutations could be detected by use of probes of both senses. This is a minimum estimate as some other mismatches are cleaved. Uncleaved mismatches were not cleaved with ribonuclease T1 or T2.

The most obvious limitation of this method appeared to be that not all mismatch types are detected. Extra cloning is needed for probe production by the SP6 system and only up to 1 kb can be studied at a time. Despite this, useful applications have appeared (Forrester *et al.*, 1987; Gibbs & Caskey, 1987; Dunn *et al.*, 1988; Carothers *et al.* 1988; Lopez-Galindez *et al.*, 1988). For example, Gibbs & Caskey (1987) found distinctive RNAase cleavage patterns in five of 14 Lesch-Nyhan syndrome patients for hypoxanthine phosphoribosyl-transferase mRNA.

ABC nuclease. This enzyme does not recognize single base pair mismatches (Thomas *et al.*, 1986) but it does after they have been reacted with carbodi-imide (Thomas *et al.*, 1986). Thus this is a hybrid method (i.e. enzyme and chemical) and the carbodi-imide method will be reviewed below. However, the strength of this hybrid method is that a cleavage method is produced together

with its attendant analytical advantages. The main limitation is that this enzyme is not widely available.

Chemical methods. Chemical methods have been used extensively over two decades to identify single-stranded regions of RNA molecules (such as tRNA) for determination of a secondary structure (Cramer, 1971). It is thus surprising that only recently have chemicals been shown to be capable of identifying single base pair mismatches. Chemicals have the advantage that they can work at a wider range of pH and ionic strengths than enzymes and do not have their instability problems and requirements for 'clean' nucleic acids.

Carbodi-imide. In the study of tRNA structure, this chemical was shown to react with one example each of a G·T and a T·C mismatch (Novack *et al.*, 1986). The reactive bases were G and T. Data are quoted but not given for G·G and T·T mismatches. Mismatched A and C could be detected using a probe of opposite sense, and thus all mismatches and hence all mutations are potentially detectable. As this is not a cleavage method (except as noted above), a positive reaction is detected by decreased mobility on a non-denaturing gel. Any overhangs likely to contain a G or T have to be made blunt with nuclease to avoid false positives.

Chemical cleavage of mismatch ('CCM'). Many chemicals react with bases of nucleic acids to produce a lesion which is cleavable by subsequent treatment with an alkali such as piperidine, and this is the principle behind the sequencing method of Maxam & Gilbert (1977). We screened a number of chemicals which might react with mismatched bases so that the DNA duplex might be later cleaved at that point by piperidine. It was found that hydroxylamine (H) would react with mismatched C bases and that osmium tetroxide (OT) would react with mismatched T bases (Cotton *et al.* 1988). These chemicals are referred to as the HOT chemicals. So far 21 T mismatches (5 T·C, 12 T·G, 3 T·T) and 25 C mismatches (9 C·T, 14 C·A, 2 C·C) have been cleaved (Cotton *et al.*, 1988; Cotton & Campbell, 1989). No mismatches resist cleavage and no false positive cleavage has been found. Mismatched A and G bases can be detected by use of the probe of opposite sense such that these are converted to mismatched T and C in the probe respectively. This allows one to detect the altered base and its position. The new base can only be determined by use of probes of both senses of the other or mutant DNA.

The chemical cleavage of mismatch method is also applicable to DNA·RNA duplexes. Viral variation in a RNA virus has been assayed (Cotton & Wright, 1989) and mutations detected in DNA·mRNA hybrids (Bate-man *et al.*, 1989; Lamande *et al.*, 1989). In these studies the mismatches shown to be cleaved so far in this system are 9 C·A, 1 T·C, 1 C·C, 2 C·U. No mismatches have failed to be cleaved and there have been no false positives.

A special advantage of this technique is that all mismatches react at approximately the same rate, in contrast with the differing rates of cleavage shown by the RNAase A method (Myers *et al.*, 1985c). Thus if an end-labelled probe is used and the heteroduplex is partially cleaved, a signal for each mismatched T and C in the duplex will be observed (Cotton *et al.*, 1988). This results in a pattern of difference between, for example, viral

isolates which reflects exactly the base sequence differences (Cotton & Wright, 1989).

In the case of single-stranded nucleic acids (e.g. mRNA), only half the mismatches will be detected without copying to form cDNA. However as matched bases near to mismatches are reactive (Cotton *et al.*, 1988; Cotton & Campbell, 1989), some of the remainder of the mismatches will be detectable indirectly when matched C and T bases are next to mismatched A and G in the probe.

MUTATION DETECTION METHODS WHERE THE MUTATION POSITION IS KNOWN (DIAGNOSTIC METHODS)

Once the position of a mutation is known it is a very different, and relatively simple, proposition to assay for it compared to the methods used for screening (see above). As the assay is usually diagnostic it needs to be as simple as possible to contain costs. There are several such systems, some of which are based on the principles and methods discussed above in the screening methods. This problem of the need for simplicity has been addressed by Nicholls & Malcolm (1989) and Landegren *et al.* (1988a).

Systems discussed above which tend to be used in this mode are the PCR and direct sequencing (Wong *et al.*, 1987) or PCR and restriction enzyme analysis (Chehab *et al.*, 1987) for mutation detection and restriction enzyme analysis looking for polymorphisms known to be useful from previous studies. As technology becomes more advanced these methods may be dismissed for these purposes (see below).

Oligonucleotide hybridization. The foundation for clinical use of this strategy was the finding in 1981 (Wallace *et al.*, 1981) that hybridization conditions could be found that allowed a 14-base-pair oligonucleotide to bind when perfectly matched and not bind when a single base pair mismatch occurred in the duplex. This system has been used and developed extensively up to the present. It was shown to be applicable to genomic DNA directly (Conner *et al.* 1983; Orkin *et al.*, 1983; Pirastu *et al.*, 1983; Le Gall *et al.*, 1986). A further step was the assay of DNA on dots (dot blot) (Kafatos *et al.*, 1979) after amplification by PCR (Saiki *et al.*, 1986). This was applied to the *ras* (Lyons *et al.*, 1988) and HLA DQ (Horn *et al.*, 1988) systems. The amplification was able to provide so much DNA for analysis that a nonradioactive assay was possible using the peroxidase system (Saiki *et al.*, 1988b). The sickle mutation has been detected by slot blot (Bowles *et al.*, 1986) after amplification of DNA directly from Guthrie blotter blood spots (Jinks *et al.*, 1989) and a phenylketonuria mutation has been detected in a similar manner (Lyonnet *et al.*, 1988).

Ligation assay. T4 polynucleotide ligase has long been known to join DNA fragments (reviewed in Engler & Richardson, 1982). Earlier studies also showed that where a mismatch occurred at a base on either side of the ligation point, efficient ligation occurred (Tsiapalis & Narang, 1970; Sgaramella & Khorana, 1972; Wiaderkiewicz & Ruiz-Carrillo, 1987; Goffin *et al.*, 1987).

Fig. 1. Principles involved in some of the methods of mutation detection

The left-hand column represents the normal situation and the right the mutant with the mutation marked ★. (a) Denaturation: mutation causes earlier melting of duplex. (b) Cleavage methods: cleavage at points of mismatch. (c) Oligonucleotides: mutation prevents binding of oligonucleotide. (d) Ligation method: ligation prevented by mismatch at one end of oligonucleotide. (e) Allele-specific amplification: PCR product formation inhibited by mismatch at one end of oligonucleotide. (f) Sandwich assay: binding prevented by enzyme cutting at mutation.

However, little evidence was provided of differences in the rate of ligation between DNA containing and not containing a mismatch at the ligation site. Recently, conditions have been found that allow only minimal ligation when a mismatch occurs either side of the point of ligation, paving the way for a simple method of assaying for the presence of a particular mutation (Landegren *et al.*, 1988a; Alves & Carr, 1988; Wu & Wallace, 1989b).

The system used has been a continuous strand of DNA to which two adjacent oligonucleotides are hybridized with 3' and 5' bases of each abutting, so that if joined the region would be complementary to the continuous strand (Fig. 1d). Under conditions of high salt and spermidine various bases (Landegren *et al.*, 1988a) were tried at the 3' end of one oligonucleotide which is at the junction so that all possible Watson-Crick pairs were assayed as well as all the possible 12 mismatched base pairs. It was found that T4 ligase under defined conditions was able to ligate the oligonucleotides when the four normal base pairs occurred at the junction but not when the 12 mismatches occurred at this point. Assay was by virtue of the fact that the 5' end of one oligonucleotide was labelled with biotin and the 3' end of the other with either ^{32}P or a fluorophor. The ligated product was captured and assayed for radioactivity on streptavidin-agarose beads or assayed on h.p.l.c. (a DNA sequencer) respectively. There

was no evidence that surrounding sequence influence the reaction.

A similar system was developed by Wu & Wallace (1989a) with high salt and spermidine. They extended the above work by testing the effect of mismatches on either side of the ligation point and showed that discrimination could be achieved in either case from matched pairs on either side. Only a limited number of mismatches were tried in this study. Assay was by electrophoresis and autoradiography. The system designed by Alves & Carr (1988) differed in that the ligation reaction occurred after the abutting oligonucleotides had been hybridized to the target *in situ* on the gel and assay was by removal of unligated and retention of ligated probe at high-stringency washing.

The first two systems were applied to the sickle cell mutation at the genomic DNA level whereas the latter was similarly applied to the Ha-ras oncogene.

This system has been developed further by Wu & Wallace (1989b) by the addition of an amplification step, thus increasing sensitivity. This amplification has been achieved by sequential rounds of template-dependent ligation. Thus two pairs of oligonucleotides are used, one pair complementary to the upper strand and the other to the lower strands of the target sequence. After the first round of ligation the reaction is heated to dissociate the ligation product and an additional round of ligation

performed and so on. This is referred to as the Ligation Amplification Reaction (LAR) and has been applied to the detection of the sickle mutation. However attention has to be paid to suppression of blunt ended ligation. It was noted that the reaction would be more convenient if a heat-stable ligase could be used.

Allele-specific amplification. This is a simple and effective extension of the PCR reaction, developed by Wu *et al.* (1989). One of the pair of oligonucleotides is replaced by two oligonucleotides covering the mutated base. One of these would be perfectly matched on hybridization and the other would create a mismatch at the 3' terminal with the same DNA (Fig. 1e). Conditions were established such that PCR product was produced in the former case but not in the latter case. This system was applied to the sickle allele of globin and found to be rapid and could be used without the use of radioactivity. Twelve individuals were assayed 'blind' and identified correctly. This assay avoids further probe hybridization, ligation or restriction enzyme cleavage and should be amenable to elimination of the gel separation step. Only A·A and T·T mismatches have been studied and the authors suggest that G·T mismatches should be avoided in probe design due to their relative stability. Similar approaches have been described independently (Gibbs *et al.*, 1989; Newton *et al.*, 1989).

Sandwich assay. The sandwich assay was first described in 1977 (Dunn & Hassell, 1977; reviewed in Nicolls & Malcolm, 1989). It was first used in diagnosis in 1983 to look for the presence of adenovirus (Ranki *et al.*, 1983). Further applications to infectious organisms are reviewed in Nicholls & Malcolm (1989). The analysis of sequence variations was reported in 1985 (Langdale & Malcolm, 1985). Two pieces of DNA either side of the sickle mutation are produced and one is ligated to a solid support (Fig. 1f). The other is radiolabelled. The target DNA, either normal or sickle-cell-derived, is digested with the enzyme *DdeI* which cuts only the sickle-cell-derived DNA. Thus if the enzyme does not cut the DNA as in the normal, the radioactivity is attached to the matrix and can be counted after mixing, annealing and washing. This is not so if the duplex is cut. Whilst simple, this method has not yet been widely used diagnostically in the area of inherited disease.

COMPARISON OF SCREENING METHODS

Screening methods (Table 1)

The methods for screening are obviously in a state of flux at present and in several years time the technologies available are likely to be very different. Obviously if it were possible for sequencing to be simplified and automated so that it was a trivial operation, it would be the method of choice. Some steps have been made towards this ultimate goal in recent years in the form of the PCR reaction and automation of sequence readout. However for the next few years at least it is likely that methods that can quickly screen long tracts of nucleic acids, so that the position of all mutations and hence the region which needs to be sequenced can be identified, will be of value. The low yield restriction enzyme method is likely to be used for some time together with the Gray version of the denaturing gradient electrophoresis. Of the hetero-duplex methods, at a disadvantage because hetero-duplexes have to be formed, the cleavage methods with

near 100% detection of mutations have to be favoured. These are the chemical-based methods, but the carbodi-imide/ABC nuclease method has not been tried on a range of mismatches in contrast to the chemical cleavage of mismatch method. A particular advantage of these methods is that end-labelled probe can be used to obtain a pattern of difference between related pieces of nucleic acid. If one needs to obtain such patterns dozens could be obtained in a day or two using nucleic acid isolated directly from infected cells in the case of viruses (Cotton & Wright, 1989). Similarly with current methodology it may be that it is easier to screen 10 kb of bases with a screening method than to sequence it all.

Diagnostic methods (Table 2)

These methods are very much in a state of flux at the moment with at least two or three novel approaches being reported each year. This situation has been partly brought about by the advent of the PCR technique. It seems unlikely that the final method of choice will avoid the PCR technique due to its potential to avoid radioactive labelling methods and perhaps electrophoresis. However, the allele-specific amplification appears to have considerable promise at this stage. The final arbiter will be speed, simplicity and low cost.

It would appear that restriction enzymes will be used to determine linkage to disease genes in families for diagnosis purposes at least into the near future.

Drs. Susan Forrest, David Howells and Julian Mercer are thanked for criticism of the manuscript.

REFERENCES

- Aboul-Ela, F., Koh, D., Tinoco, I., Jr. & Martin, F. H. (1985) *Nucleic Acids Res.* **13**, 4811–4824
- Albertini, R. J., Castle, K. L. & Borchering, W. R. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6617–6621
- Alves, A. M. & Carr, F. J. (1988) *Nucleic Acids Res.* **16**, 8723
- Bateman, J. F., Lamande, S. R., Dahl, H. M., Chan, D., Mascara, T. & Cole, W. G. (1989) *J. Biol. Chem.*, in the press
- Benevides, J. M., Wang, A. H.-J., van der Marel, G. A., van Boom, J. H. & Thomas, G. J. (1989) *Biochemistry* **28**, 304–310
- Bowles, N. E., Richardson, P. J., Olsen, E. G. J. & Archard, L. C. (1986) *Lancet* **i**, 1120–1123
- Brownlee, G. G. (1972) in *Laboratory Techniques in Biochemistry and Molecular Biology* (Work, T. S. & Work, E., eds.), pp. 67–99, American Elsevier, New York
- Carothers, A. M., Urlaub, G., Grunberger, D. & Chasin, L. A. (1988) *Somatic Cell Mol. Genet.* **14**, 169–183
- Caskey, C. T. (1987) *Science* **236**, 1223–1229
- Chehab, F. F., Doherty, M., Cai, S., Kay, Y. W., Cooper, S. & Rubin, E. M. (1987) *Nature (London)* **329**, 293–294
- Conner, B. J., Reye, A. A., Morin, C., Itakura, K., Teplitz, R. L. & Wallace, R. B. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 278–282
- Cotton, R. G. H. & Campbell, R. D. (1989) *Nucleic Acids Res.* **17**, 4223–4233
- Cotton, R. G. H. & Wright, P. J. (1989) *J. Virol. Methods*, in the press
- Cotton, R. G. H., Rodrigues, N. R. & Campbell, R. D. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4397–4401
- Cramer, F. (1971) *Prog. Nucleic Acid Res. Mol. Biol.* **11**, 391–421
- Diaz-Ruiz, J. R. & Kaper, J. M. (1979) *Biochim. Biophys. Acta* **564**, 275–288

- Dodgson, J. B. & Wells, R. D. (1977) *Biochemistry* **16**, 2374-2379
- Dunn, A. R. & Hassell, J. A. (1977) *Cell* **12**, 23-36
- Dunn, J. M., Phillips, R. A., Becher, A. J. & Gallic, B. L. (1988) *Science* **241**, 1797-1800
- Dunning, A. M., Talmud, P. & Humphries, S. E. (1988) *Nucleic Acids Res.* **16**, 10393
- Ehrlich, M., Ehrlich, K. & Mayo, J. A. (1975) *Biochim. Biophys. Acta* **395**, 109-119
- Engel, J. D. & von Hippel, P. H. (1974) *Biochemistry* **13**, 4143-4158
- Engel, J. D. & von Hippel, P. H. (1978) *J. Biol. Chem.* **253**, 927-934
- Engler, M. J. & Richardson, C. (1982) in *The Enzymes* (Boyer, P. D., ed.), vol. 15, pp. 3-29, Academic Press, New York
- Fischer, S. G. & Lerman, L. S. (1979) *Cell* **16**, 191-200
- Fischer, S. G. & Lerman, L. S. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 4420-4424
- Fischer, S. G. & Lerman, L. S. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1579-1583
- Flavell, R. A., Kooter, J. M., De Boer, E., Little, P. F. R. & Williamson, R. (1978) *Cell* **15**, 25-41
- Forrester, K., Almoguera, C., Han, K., Grizzle, W. E. & Perucho, M. (1987) *Nature (London)* **327**, 298-303
- Freeman, G. J. & Huang, A. S. (1981) *J. Gen. Virol.* **57**, 103-117
- Gao, X. & Patel, D. J. (1987) *J. Biol. Chem.* **262**, 16973-16984
- Gibbs, R. A. & Caskey, C. T. (1987) *Science* **236**, 303-305
- Gibbs, R. A., Nguyen, P. & Caskey, C. T. (1989) *Nucleic Acids Res.* **17**, 2437-2448
- Gill, J. E., Mazrimas, J. A. & Bishop, C. C. (1974) *Biochim. Biophys. Acta* **335**, 330-348
- Goffin, C., Bailly, V. & Verly, W. G. (1987) *Nucleic Acids Res.* **15**, 8755-8771
- Gorenstein, D. G., Schroeder, S. A., Fu, J. M., Metz, J. T., Roongta, V. & Jones, C. R. (1988) *Biochemistry* **27**, 7223-7237
- Greever, R. F., Wilson, L. B., Nallaseth, P. S., Milner, P. F., Bittner, M. & Wilson, J. T. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 5081-5085
- Griffiths, P. D. & Grundy, J. E. (1987) *Biochem. J.* **241**, 313-324
- Horn, G. T., Bugawan, T. L., Long, C. M. & Erlich, H. A. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6012-6016
- Hughes, M. R., Malloy, P. J., Kieback, D. G., Kesterson, R. A., Pike, J. W., Feldman, D. & O'Malley, B. W. (1988) *Science* **242**, 1702-1705
- Hunter, W. N., Brown, T., Anand, N. N. & Kennard, O. (1986) *Nature (London)* **320**, 552-555
- Hunter, W. N., Brown, T. & Kennard, O. (1987) *Nucleic Acids Res.* **15**, 6589-6606
- Ikuta, S., Takagi, K., Wallace, R. B. & Itakura, K. (1987) *Nucleic Acids Res.* **15**, 797-811
- Ito, Y. & Joklik, W. K. (1972) *Virology* **50**, 202-208
- Jinks, D. C., Minter, M., Tarver, D. A., Vanderford, M., Hejtmancik, J. F. & McCabe, E. R. B. (1989) *Hum. Genet.* **81**, 363-366
- Jones, F. S., Grimberg, J. I., Fisher, S. G. & Ford, J. D. (1985) *Gene* **39**, 77-82
- Kafatos, F. C., Jones, C. W. & Efstratiadis, A. (1979) *Nucleic Acids Res.* **7**, 1541-1552
- Kalnik, M. W., Kouchakdjian, M., Li, B. F. L., Swann, P. F. & Patel, D. J. (1988a) *Biochemistry* **27**, 100-108
- Kalnik, M. W., Kouchakdjian, M., Li, B. F. L., Swann, P. F. & Patel, D. J. (1988b) *Biochemistry* **27**, 108-115
- Kan, Y. W. & Dozy, A. M. (1978) *Lancet* **ii**, 910-912
- Kennard, O. (1988) in *Structure and Expression: Vol. 2, DNA and Its Drug Complexes* (Sarma, R. H. & Sarma, H. M., eds.) pp. 1-25, Adenine Press
- Keohavong, P., Wang, C. C., Cha, R. S. & Thilly, W. G. (1988) *Gene* **71**, 211-216
- Kramer, B., Kramer, W. & Fritz, H.-J. (1984) *Cell* **38**, 879-887
- Lamande, S. R., Dahl, H. M., Cole, W. G. & Bateman, J. F. (1989) *J. Biol. Chem.*, in the press
- Landegren, U., Kaiser, R., Sanders, J. & Hood, L. (1988a) *Science* **241**, 1077-1080
- Landegren, U., Kaiser, R., Caskey, C. T. & Hood, L. (1988b) *Science* **242**, 229-237
- Langdale, J. A. & Malcolm, A. D. B. (1985) *Gene* **36**, 201-210
- Latham, T. & Smith, F. I. (1989) *DNA* **8**, 223-230
- Le Gall, I., Millasseau, P., Dausset, J. & Cohen, D. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 7836-7840
- Lerman, L. S., Fischer, S. G., Hurley, I., Silverstein, K. & Lumelsky, N. (1984) *Annu. Rev. Biophys. Bioeng.* **13**, 399-423
- Lomant, A. J. & Fresco, J. R. (1975) *Prog. Nucleic Acid Res. Mol. Biol.* **15**, 185-218
- Lyonnet, S., Caillaud, C., Rey, F., Berthelon, M., Frezal, J. & Rey, J. (1988) *Lancet* **ii**, 507
- Lyons, J., Janssen, J. W. G., Bartram, C., Layton, M. & Mufti, G. J. (1988) *Blood* **71**, 1707-1712
- Maddalena, A., Spence, J. E., O'Brien, W. E. & Nussbaum, R. L. (1988) *J. Clin. Invest.* **82**, 1353-1358
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) in *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 560-564
- Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K. & Green, M. R. (1984) *Nucleic Acids Res.* **12**, 7035-7056
- Morley, A. A., Trainor, K. J., Seshadri, R. & Ryall, R. G. (1983) *Nature (London)* **302**, 155-156
- Myers, R. M. & Maniatis, T. (1986) *Cold Spring Harbor Symp. Quant. Biol.* **51**, 275-284
- Myers, R. M., Fischer, S. G., Maniatis, T. & Lerman, L. S. (1985a) *Nucleic Acids Res.* **13**, 3111-3130
- Myers, R. M., Fisher, S. G., Lerman, L. S. & Maniatis, T. (1985b) *Nucleic Acids Res.* **13**, 3131-3145
- Myers, R. M., Larin, Z. & Maniatis, T. (1985c) *Science* **230**, 1242-1246
- Myers, R. M., Lumelsky, N., Lerman, L. S. & Maniatis, T. (1985d) *Nature (London)* **313**, 495-498
- Myers, R. M., Maniatis, T. & Lerman, L. S. (1987) *Methods Enzymol.* **155**, 501-527
- Myers, R. M., Sheffield, V. C. & Cox, D. R. (1988) in *Genomic Analysis: A Practical Approach* (Davies, K., ed.), pp. 95-139, IRL Press, Oxford
- Newton, C. R., Graham, A., Heptinstall, L. E., Summers, S. J., Kalsheker, N., Smith, J. C. & Markham, A. F. (1989) *Nucleic Acids Res.* **17**, 2503-2516
- Nicholls, P. J. & Malcolm, A. D. B. (1989) *J. Clin. Lab. Anal.* **3**, 122-136
- Noll, W. W. & Collins, M. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 3339-3343
- Novack, D. F., Casna, N. J., Fischer, S. G. & Ford, J. P. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 586-590
- Orkin, S. H., Markam, A. F. & Kazazian, H. H., Jr. (1983) *J. Clin. Invest.* **71**, 775-779
- Patel, D. J., Shapiro, L. & Hare, D. (1987) *Annu. Rev. Biophys. Chem.* **16**, 423-454
- Pirastu, M., Kan, Y. W., Cao, A., Conner, B. J., Teplitz, R. L. & Wallace, R. B. (1983) *N. Engl. J. Med.* **30**, 284-287
- Po, T., Steger, G., Rosenbaum, V., Kaper, J. & Riesner, D. (1987) *Nucleic Acids Res.* **15**, 5069-5083
- Privé, G. G., Heinemann, U., Chandrasegaran, S., Kan, L.-S., Kopka, M. L. & Dickerson, E. (1987) *Science* **238**, 498-504

- Quignard, E., Fazakerley, G. V., van der Marel, G., van Boom, J. H. & Guschlbauer, W. (1987) *Nucleic Acids Res.* **15**, 3397–3409
- Ranki, M., Palva, A., Virtanen, M., Laaksonen, M. & Soderlung, H. (1983) *Gene* **21**, 77–85
- Rosenbaum, W. & Riesner, D. (1987) *Biophys. Chem.* **26**, 235–246
- Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York and Berlin
- Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. G., Horn, G. T., Ehrlich, H. A. & Arnheim, N. (1985) *Science* **230**, 1350–1354
- Saiki, R. K., Bugawan, G. T., Horn, K. B. & Erlich, H. A. (1986) *Nature (London)* **324**, 163–166
- Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988a) *Science* **239**, 487–491
- Saiki, R. K., Chang, C.-A., Levenson, C. H., Warren, T. C., Boehm, C. D., Kazazian, H. H., Jr. & Erlich, H. A. (1988b) *N. Engl. J. Med.* **319**, 537–541
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467
- Sgaramella, V. & Khorana, H. C. (1972) *J. Mol. Biol.* **72**, 427–444
- Sheffield, V. C., Cox, D. R., Lerman, L. S. & Myers, R. M. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 232–236
- Shenk, T. E., Rhodes, C., Rigby, P. W. J. & Berg, P. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 989–993
- Smith, F. I., Parvin, J. D. & Palese, P. (1986) *Virology* **150**, 55–64
- Smith, F. I., Latham, T. E., Ferrier, J. A. & Palese, P. (1988) *Genomics* **3**, 217–223
- Smith, L. M., Saunders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., Heiner, C., Kenk, S. B. H. & Hood, L. E. (1986) *Nature (London)* **321**, 674–679
- Szer, W. & Shugar, D. (1966) *J. Mol. Biol.* **17**, 174–187
- Tenover, F. C. (1988) *Clin. Microbiol. Rev.* **1**, 82–101
- Thomas, D. C., Kunkel, T. A., Casna, N. J., Ford, J. P. & Sancar, A. (1986) *J. Biol. Chem.* **261**, 14496–14505
- Tsiapalis, C. M. & Narang, S. A. (1970) *Biochem. Biophys. Res. Commun.* **39**, 631–636
- von Hippel, P. H. & McGhee, J. D. (1972) *Annu. Rev. Biochem.* **41**, 231–300
- Wada, A., Tachibana, H., Gotoh, O. & Takanami, M. (1976) *Nature (London)* **263**, 439–440
- Wallace, R. B., Shaffer, J., Murphy, R. F., Bonner, J., Hirose, T. & Itakura, K. (1979) *Nucleic Acids Res.* **6**, 3543–3557
- Wallace, R. B., Johnson, M. J., Hirose, T., Miyake, T., Hawashima, E. H. & Itakura, K. (1981) *Nucleic Acids Res.* **9**, 879–894
- Werntges, H., Steger, G., Riesner, D. & Fritz, H.-J. (1986) *Nucleic Acids Res.* **14**, 3773–3790
- Wiaderkiewicz, R. & Ruiz-Carrillo, A. (1987) *Nucleic Acids Res.* **15**, 7831–7848
- Winter, E., Yamamoto, F., Almoguera, C. & Perucho, M. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7575–7579
- Wolfe, H. J. (1988) *Am. J. Clin. Pathol.* **90**, 340–344
- Wong, C., Dowling, C. E., Saiki, R. K., Higuchi, R. G., Erlich, H. A. & Kazazian, H. H., Jr. (1987) *Nature (London)* **330**, 384–386
- Wu, D. Y. & Wallace, R. B. (1989a) *Gene* **76**, 245–254
- Wu, D. Y. & Wallace, R. B. (1989b) *Genomics* **4**, 560–569
- Wu, D. Y., Ugozzoli, L., Pal, B. K. & Wallace, R. B. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 2757–2760