Genotypic selection methods for the direct analysis of point mutations

Barbara L. Parsons *, Robert H. Heflich

Division of Genetic Toxicology, HFT-120, National Center for Toxicological Research, 3900 NCTR Dr., Jefferson, AR 72079, USA

Received 19 January 1997; accepted 23 May 1997

Abstract

Genotypic selection enriches a particular DNA sequence relative to another closely-related DNA sequence based only on a change of one or a few bases. This review is a survey of the genotypic selection methods that have the sensitivity to detect rare point mutations. These methods are primarily being used to study mutations caused by environmental mutagens; however, the ability to detect and measure very minor DNA sequence populations is likely to further research efforts in many fields. The approaches for allele-selection have intrinsic strengths and weaknesses, and vary greatly in sensitivity. The most sensitive method is Restriction Fragment Length Polymorphism/Polymerase Chain Reaction (RFLP/PCR) by which mutant fractions as low as 1 mutant allele in 10^8 wild-type alleles can be detected. The RFLP/PCR approach is presented as a prototype genotypic selection method. Genotypic selection methods are categorized in terms of those that (1) selectively destroy the abundant or wild-type allele, (2) selectively amplify the rare or mutant allele, or (3) spatially separate the alleles. Issues relevant to the further development of genotypic selection methods include initial DNA pool size, strategies to eliminate the bulk of extraneous DNA, the use of an internal copy number standard in quantitative PCR, the fidelity of thermostable DNA polymerases, and the effective use of PCR in linking two or more genotypic selection techniques. We conclude that proficient genotypic selection requires more than one allele-enrichment technique with at least one of these preceding a high-fidelity PCR amplification step. © 1997 Elsevier Science B.V.

Keywords: Genotypic selection; Pool screening; Mutation detection; Sensitivity; Point mutation; Polymerase chain reaction

Abbreviations: ACB-PCR, allele-specific competitive blocker PCR; ACRS, amplification-created restriction site; ADPL, allele discrimination by primer length; AIRS, artificial introduction of restriction site; ARMS, amplification refractory mutation system; ASA, allele-specific amplification; CDCE, constant denaturant capillary electrophoresis; CDGE, constant denaturant gel electrophoresis; CMSA, competitive mobility shift assay; COP, competitive oligonucleotide priming; DGGE, denaturing gradient gel electrophoresis; dNTPs, deoxynucleotide triphosphates; LAR, ligase amplification reaction; LCR, ligase chain reaction; LDR, ligase detection reaction; MAMA, mismatch amplification mutation assay; MASA, mutant allele-specific amplification; MNU, N-methyl-N-nitrosourea; MSPA, mutation-specific PCR assay; MS-PCR, mutagenically-separated PCR; MutEx, MutS/exonuclease; PAMSA, PCR amplification of multiple specific alleles; PASA, polymerase amplification of specific alleles; Pfu, Pyrococcus furiosus; PCR, polymerase chain reaction; PNA, peptide nucleic acid; Q-PCR, quantitative PCR; RFLP, restriction fragment length polymorphism; RSM, restriction site mutation (method); SNuPE, single nucleotide primer extension; SSCP, single-strand conformation polymorphism

* Corresponding author. Tel.: +1 (870) 543 7496; fax: +1 (870) 543 7393; e-mail: bparsons@nctr.fda.gov
1. Introduction

‘Genotypic selection’ is a term that refers to the DNA-based enrichment of a particular allele. Methods for genotypic selection, therefore, are a subset of the many procedures developed for either mutation detection or genotyping. Genotypic selection, as a term or as a category of mutation detection methods, is not widely recognized. Sometimes the term is used in reference to a few sensitive mutation detection methods [1–3], the most prominent of which is RFLP/PCR [2]. Most often, however, techniques that enrich a rare allele are presented without any mention of genotypic selection. This review is based on the idea that the genotypic selection of point mutations is a broad area that includes ‘‘any technique that enriches a particular DNA sequence relative to another closely-related DNA sequence based only on a difference of one or a few bases’’. This may be a useful framework for thinking about genotypic selection because additional sensitive mutation detection schemes are likely to be developed by coupling the less-sensitive allele-specific selection methods that already exist but are not thought of as genotypic selection. Genotypic selection methods have also been used to analyze length mutations and translocations but these topics are beyond the scope of this review. Naturally, this review will place a greater emphasis on those methods that provide a relatively high enrichment and allow a rare or mutant allele to be detected and characterized in the presence of an abundant or wild-type allele.

2. Applications for genotypic selection

The suitability of a genotypic selection method for a particular application is directly related to the sensitivity of the assay (Fig. 1), the specific change in the DNA sequence being detected, and the pool size of target sequences available. Generally, these methods are useful for detecting mutations in a small target of known sequence, primarily point mutations at one or a few basepairs and small insertions or deletions. These techniques are applicable to any situation where there is a need to detect and quantify a specific DNA sequence that represents only a minor fraction of a pool of closely-related sequences.

<table>
<thead>
<tr>
<th>Method</th>
<th>Sensitivity</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSCP</td>
<td>$10^{-1}$</td>
<td>Analysis of tumor tissue</td>
</tr>
<tr>
<td>SSCP</td>
<td>$10^{-2}$</td>
<td>Population screening for genetic polymorphisms</td>
</tr>
<tr>
<td>ASA, PASA</td>
<td>$10^{-3}$</td>
<td>Detection of mutagen-induced mutations</td>
</tr>
<tr>
<td>mismatched PCR</td>
<td>$10^{-4}$</td>
<td>Detection of spontaneous mutations</td>
</tr>
<tr>
<td>PNA clamping, DGGE</td>
<td>$10^{-5}$</td>
<td></td>
</tr>
<tr>
<td>Blocker-PCR, CDCE, MfB</td>
<td>$10^{-6}$</td>
<td></td>
</tr>
<tr>
<td>MS-PCR</td>
<td>$10^{-7}$</td>
<td></td>
</tr>
<tr>
<td>ACG-PCR, Gap-LCR</td>
<td>$10^{-8}$</td>
<td></td>
</tr>
<tr>
<td>MAMA, 2X CDCE</td>
<td>$10^{-9}$</td>
<td></td>
</tr>
<tr>
<td>MutEx/PCR/SNucPE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCR/RFLP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RFLP/PCR</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Consequently, the potential applications for genotypic selection technology are varied and many.

2.1. Risk assessment in animal models

Measurement of mutations induced in animal model systems is considered an important means of testing the mutagenic potency of potential genotoxics. This is because the measurement of mutation in vivo takes into account the many different biological parameters involved in producing a mutation, such as route of administration, mutagen activation, DNA adduction, repair, and relevant apoptotic pathways. However, measuring the relatively infrequent mutational events that result from toxicant treatment at biologically relevant doses is a challenging undertaking [4]. Toward this end, phenotypic selection assays have been employed that rely on the alteration of a selectable protein function to permit the clonal amplification required for detection of rare mutational events (e.g., hprt [5] or LacI [6]). Unfortunately, these methods are limited in terms of the target sequence, tissue, and timing of mutational analysis, and often involve tedious clone by clone analysis [7,8]. Genotypic selection has the potential
to analyze point mutations in any gene sequence of interest that result from in vivo toxicant treatment. Pools of DNA can be analyzed from any tissue at any time after treatment provided only that the mutations are fixed by DNA replication and an adequate DNA pool size can be isolated. The ability to directly measure mutations in genes relevant to cancer, like oncogenes and tumor suppressor genes, is a significant advantage for cancer risk assessment. A sensitive genotypic selection method would eliminate the need for the clonal expansion of transformed cells in order to detect mutations and, consequently, eliminate the need for tumor promoters associated with some models of carcinogenesis [9–12]. Because genotypic selection does not require the use of viable cells, it is particularly well-suited to the analysis of mutations in archival tissue [1]. Furthermore, genotypic selection may prove to be a useful approach for studying germ cell mutagenesis and augmenting the specific-locus tests currently used to study mutations induced in embryos treated in utero.

2.2. Biomonitoring

Biomonitoring is used here to describe a broad group of applications related to the detection of somatic mutations in human tissues. Genotypic selection could be used to determine the spontaneous mutation frequency in germline and somatic cells in various populations [1,13]. The specificity intrinsic to the molecular techniques used for allele-selection generally precludes mutation detection in more than one or a few basepairs at a time. Although this small target size is a major disadvantage of these approaches, direct information regarding the spontaneous mutation frequencies of particular sequences could, nevertheless, be very useful. Spontaneous mutation frequency determined by genotypic selection could be compared to current estimates of spontaneous mutation frequency that are based on the analysis of a few phenotypically-selectable genes or the frequency of various genetic diseases [14,15].

The impact of environmental and occupational toxicant exposure upon the generation of human somatic mutation is of great interest. In order to evaluate this impact, efforts have been focused toward the determination of mutational spectra [16]. Genotypic selection has been proposed as the most efficient means of determining mutational spectra in humans, although DNA pool size considerations may require the use of material from biopsy or autopsy (see Section 6.1) [8,16,17]. Also, mutagen-induced base substitutions can be analyzed in human tissue culture cells using genotypic selection methods. In fact, this has been one of the most successful applications of genotypic selection [18–25].

2.3. Mechanisms of carcinogenesis and mutagenesis

Carcinogenesis is a multistep process involving the accumulation of mutations in genes that control cell growth and proliferation: namely, oncogenes and tumor suppressor genes. A finite number of such mutations appears to be the most important determinant in the genesis of a tumor [26–30]. Most oncogene and tumor suppressor gene mutations have been identified in tumor tissue, where the transformed phenotype has resulted in clonal amplification. This type of analysis has provided information about which mutations are most often present in the various tumor types of different organs. Early and late events in tumorigenesis, however, can rarely be distinguished by the analysis of tumor tissue. Early events in tumorigenesis, the timing of mutation induction, repair, and latency, can be better addressed using sensitive genotypic selection methods in animal model systems [11,31–34]. Further insight into the mechanisms of tumorigenesis can be obtained by measuring these parameters in different strains of animals with differing tumorigenic responses [12].

Genotypic selection should also be useful for studying basic mechanisms of mutagenesis. More accurate measurements of mutation frequencies and rates may be made than are possible when measuring mutations selected on the basis of a phenotype. Sequence changes can be measured in genes for which a convenient selection system does not exist and in the more than 90% of the mammalian genome that has no apparent protein coding function. Detecting all sequence changes without the filter of a phenotype should increase our knowledge of the types of mutations formed spontaneously and the types of mutations produced by different DNA damages. Also, genotypic selection may provide a clearer assessment of the effects of sequence context, tissue kinetics, and repair status on mutagenesis, and all
these measurements can be made using endogenous DNA sequences. One obvious shortcoming of this approach to studying mutation is that the biological significance of the sequence changes detected by genotypic selection must be determined by other means.

Other developing applications of genotypic selection include population screening for alleles associated with human diseases [8,35–37], understanding the role of mitochondrial DNA point mutations in aging [3,38–42], measurement of DNA polymerase fidelity [43,44], screening for point mutations associated with pathogenicity in attenuated viral vaccine preparations [45], and the study of genetic diversity within populations, from microbes to man.

3. Sensitivity required for the various applications of genotypic selection methods

The sensitivities of genotypic selection methods vary by many orders of magnitude (see Section 5 and Fig. 1). Those with greater sensitivity are necessarily more complicated in terms of methodology. Therefore, the appropriate choice of a genotypic selection method depends on the sensitivity required for a particular application. Two applications that require little sensitivity are genotyping and the analysis of tumor tissue. Methods that have been used only for genotyping (to detect an allele that is present at a ratio of no less than one in two) are numerous and beyond the scope of this review. The analysis of tumor tissue, however, is an application in which the mutant allele may represent only a minor fraction. Excised tumors are often mixtures of normal and transformed cells, and the transformed cells are often heterozygous for a given mutation [46]. Nonetheless, a genotypic selection technique with a sensitivity of ~1 in 10 is often sufficient for this particular application in which a mutant allele has been enriched by clonal tumor growth.

At the other end of the spectrum, in terms of the sensitivity required, is the detection of spontaneous mutations. Estimates of spontaneous mutation frequencies fall between >1 × 10⁻⁴ and ~5 × 10⁻⁸ per locus, depending on the endpoint that was measured [15]. Therefore, a mutation frequency between 1 × 10⁻⁷ and 5 × 10⁻¹¹ per basepair can be calculated by assuming that there are 1000 mutable basepairs per locus. It should be noted, however, that Cha et al. [31] used a genotypic selection approach to demonstrate that the frequency of a spontaneous ras mutation in rats was 1 × 10⁻⁵. This unexpectedly high frequency emphasizes two points. The precise sensitivity required to detect a particular spontaneous mutation cannot be known in advance. Also, the spontaneous mutation frequency is likely to vary greatly between sequences and it is possible that it will be measurable only at mutational hotspots. Clearly, the detection of spontaneous mutations warrants the use of the most sensitive genotypic selection methods available. Presently, there are few examples where the spontaneous mutation frequency for a particular basepair of a gene in a particular species is known. Hopefully, through the use of sensitive genotypic selection methods this type of information will become more readily available.

The detection of induced mutations in laboratory animals is expected to require an intermediate sensitivity. The spontaneous mutant frequency in rats is ~5 × 10⁻⁹ for the hprt locus, or roughly 5 × 10⁻⁹ per basepair [15]. If a particular mutagen treatment caused a 100-fold increase in this mutant frequency, then the induced mutant frequency would be 5 × 10⁻⁷ per basepair. In order to detect these mutations, a few copies of mutant sequence must be detected amid a huge pool of wild-type sequence. Therefore, genotypic selection schemes must effect a significant enrichment of the rare mutant sequence relative to the abundant wild-type sequence.

4. RFLP/PCR: the prototype genotypic selection scheme

The only genotypic selection method that has achieved the sensitivity expected to detect most spontaneous mutations is the Restriction Fragment Length Polymorphism/Polymerase Chain Reaction (RFLP/PCR) method [2,47–49], or the conceptually similar Restriction Site Mutation (RSM) method [25,50]. These approaches measure mutations that alter a restriction endonuclease recognition sequence. Digestion by the restriction enzyme reduces the amount of wild-type sequence, and PCR with primers flanking the restriction site amplifies the uncut mu-
Fig. 2. Genotypic selection by Restriction Fragment Length Polymorphism/Polymerase Chain Reaction (RFLP/PCR) (based on Pourzand and Cerutti [2]).

The number of plaques corresponding to any particular mutation is compared to the number of plaques corresponding to the copy number standard and the original mutation frequency is calculated.

The principal limitation on the sensitivity of this assay is the efficiency of the restriction digest. Not every restriction site can be analyzed by this procedure because restriction enzyme digestion does not always approach completeness [25,52]. Of interest in this regard is Taq I endonuclease, which has been used in several RFLP/PCR studies [7,25,52,53]. Taq I reduces the background of uncut wild-type sequence and consequently increases sensitivity because it is thermostable and digests wild-type sequence before, during, and after PCR. A secondary limitation on assay sensitivity is the background of PCR-induced mutations created by the amplification of uncut wild-type sequence. Pourzand and Cerutti [19] increased the sensitivity of RFLP/PCR by substituting Pyrococcus furiosus (Pfu) DNA polymerase for Taq polymerase during the initial rounds of PCR. Unlike Taq, Pfu polymerase possesses an error-correcting, 3′-5′ exonuclease activity [54].

Steingrimsdottir and colleagues [25] have uncovered additional factors that affect assay performance. For instance, it is apparent that certain types of DNA damage can produce mutant signals in the assay, presumably by blocking restriction enzyme digestion and subsequently causing an alteration in DNA sequence during PCR. It also appears that some cells produce a high background signal that is unrelated to the spontaneous mutant frequency. The source of this background signal has not been identified.

An intrinsic constraint upon the RFLP/PCR assay is that not every gene sequence of interest encompasses a restriction enzyme cleavage site. Thus, the range of sequences that can be studied using RFLP/PCR is somewhat limited. Also, not all sites that are amenable to RFLP/PCR are of equal interest. While the Taq I endonuclease is particularly well-suited for genotypic selection, and Taq I recognition sequences can be found in the human H-ras and p53 genes, these restriction sites do not lie in sequences important for tumor-related mutations in these genes [7,52,53]. However, there are a number of restriction endonuclease sites that do have significance to carcinogenesis. Restriction sites that have been used for RFLP/PCR include an Msp I site in human H-ras that contains the first two bases of codon 12 [19,23], Msp I and Hae III sites in the human p53 gene that, respectively, contain the tumor mutational hot spot codons 248 and 249 [18,20–22,24,55], and an Mnl I site in rat H-ras codon 12 [12]. Another approach to utilizing available restriction endonuclease recognition sequences was taken by Palombo et al. [56] who measured N-methyl-N-
nitrosourea- (MNU)-induced mutations in two reporter gene sites whose sequence was expected to be a hot spot and a cold spot for mutagenesis.

Most of the RFLP/PCR studies have been conducted in vitro, using cultured mammalian cells as targets for mutagenesis. The approach, however, has also been applied to measuring mutations in *Xenopus*, rodents, and humans [12,18,57,58]. Although the in vivo studies suggest the potential usefulness of RFLP/PCR for biomonitoring and risk assessment, these analyses were conducted with varying levels of sophistication. Many used no mutant standards and the levels of sensitivity were as low as $10^{-3}$.

5. Allele selection techniques for the detection of point mutations

Allele-selection techniques can be placed into three categories: (1) those that preferentially destroy the wild-type allele, (2) those that preferentially amplify the mutant or rare allele, and (3) those that spatially separate the mutant from the wild-type allele.

5.1. Methods that preferentially destroy the abundant or wild-type allele

5.1.1. RFLP / PCR

Restriction digestion, as has already been mentioned, can be a very efficient way in which to destroy a wild-type allele that contains a restriction site. The use of this selection in RFLP/PCR has been described (see Section 4).

5.1.2. PCR / RFLP

In this approach, PCR of a target sequence is followed by restriction enzyme digestion of the products using an enzyme that cuts at the site of mutation. PCR/RFLP has been used extensively to detect mutations in *ras* genes [59], as well as in the genotyping of a number of human diseases [60,61]. Usually the mutation is detected by the loss of a restriction site from the wild-type sequence (Fig. 3A), but mutations also have been detected by the sequence change forming a new restriction site (Fig. 3B). However, the latter situation precludes further characterization of the mutation by sequencing the intact PCR product. A major advantage of PCR/RFLP is the possibility of creating a restriction site where none exists naturally by using a partially mismatched PCR primer that abuts the sequence of interest (primer-mediated RFLP, also called Amplification-Created Restriction Site (ACRS) or Artificial Intro-

![Fig. 3. Genotypic selection by Polymerase Chain Reaction/Restriction Fragment Length Polymorphism (PCR/RFLP). The strategies depicted include situations where: (A) the wild-type sequence contains a restriction site, (B) the mutation being detected has formed a new restriction site, and (C) a base change introduced by PCR creates a restriction site in the wild-type sequence. Basepairs corresponding to a restriction site are indicated by rectangles. PCR primers are depicted by arrows. Basepairs not corresponding to the sequence of a restriction site are indicated by ovals. Wild-type basepairs are stippled and mutant basepairs are filled.](image)
PCR/RFLP is used mainly as a rapid screen for previously described mutations. When visualization of the digestion products is performed on an ethidium bromide-stained gel, the sensitivity for mutation detection is generally less than 1 in 100 [65–67]. A second round of PCR followed by restriction and non-radiolabeled detection can increase sensitivity to $10^{-3}$ to $10^{-4}$ [68,69]. Versions of PCR/RFLP that are still more sensitive have used radiolabeled probes or high-fidelity T7 DNA polymerase to achieve sensitivities of $10^{-4}$ to $10^{-6}$ [42,70–72].

5.1.3. Mismatch protection from exonuclease cleavage by MutS

Another genotypic selection technique in which the wild-type allele is selectively destroyed is based on the MutS/exonuclease (MutEx) assay that is outlined in Fig. 4 [73]. Mutant and wild-type sequences are denatured and reannealed, thereby converting these alleles into heteroduplexes or homoduplexes, respectively. MutS, the *E. coli* mismatch binding protein, is added and it binds to the mismatches in the heteroduplexed molecules. The exonuclease activity of T7 DNA polymerase is then used to degrade the unbound homoduplexes and, therefore, the degradation of homoduplex accomplishes the selective degradation of the abundant, wild-type allele. While the MutEx assay was initially developed to map unknown mutations, it was later adapted to screening pools of alleles for a known but rare mutation at codon 61 of the mouse H-ras gene [74]. The MutEx selection step was followed by PCR in order to amplify the selected population of DNA sequences. Then single nucleotide primer extension (SNuPE) [75] was used for mutant and wild-type base-identification. The overall sensitivity of this assay was 1 in 50,000 [74]. Because SNuPE by itself was shown to have a sensitivity of no more than 1 in 50, it was concluded that the MutEx/PCR approach had enriched the mutant fraction ~1000-fold [74]. An advantage of this approach over RFLP/PCR is that it does not require a restriction enzyme cleavage site in the sequence to be analyzed. However, MutS is known to vary in its binding to different mismatched bases, both in terms of affinity and positioning of the protein on the mismatched basepair [76]. Therefore, it remains to be demonstrated that this approach will provide the same mutant enrichment for all base substitutions. The fact that any base substitution will result in two different mismatches somewhat mitigates this potential problem. This approach can be used to detect mutations in only one basepair at a time. This is a disadvantage of this approach as compared to RFLP/PCR where the mutational target is the length of a restriction endonuclease cleavage site.

Fig. 4. The MutEx approach to allele enrichment. A mutant basepair is indicated by a closed rectangle; the corresponding wild-type basepair is indicated by an open rectangle. The sequences are denatured and reannealed. Mutant alleles (in the form of heteroduplex molecules) are protected from the 3'-5' exonuclease activity of T7 DNA polymerase by the binding of the *E. coli* mismatch binding protein, MutS. Unprotected wild-type alleles (mostly in the form of homoduplex molecules) are degraded to nucleotides. MutS-protected sequences can then be amplified by PCR.

<table>
<thead>
<tr>
<th>Abundant wild-type allele</th>
<th>Rare mutant allele</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denature and reanneal</td>
<td></td>
</tr>
<tr>
<td>Wild-type allele remains primarily as homoduplexes</td>
<td>Mutant allele becomes incorporated into heteroduplexes</td>
</tr>
<tr>
<td>MutS</td>
<td></td>
</tr>
<tr>
<td>3'-5' exonuclease activity of T7 DNA polymerase</td>
<td></td>
</tr>
<tr>
<td>MutS-protected DNA is used as template for PCR</td>
<td></td>
</tr>
</tbody>
</table>

| Wild-type homoduplexes are degraded to nucleotides | |
5.2. Methods that selectively amplify the rare or mutant allele

5.2.1. Allele-specific amplification or allele-specific PCR

This group of techniques uses a PCR primer that has fewer mismatches when hybridized to the rare allele than to the abundant allele as the basis for the genotypic selection [63,77,78]. Mismatched-primer design and reaction conditions that aid in the discrimination of allele-specific amplification (ASA) have been reviewed [79]. ASA requires the use of a polymerase that lacks a 3'-5' exonuclease activity because such activity would remove the 3'-terminal mismatch(es). Generally, 5'-3' terminal mismatches result in the checking of mismatch discrimination in terms of extension of mismatched mismatches . Generally, Taq polymerase has been used. However, there is some evidence that the Stoffel fragment of Taq polymerase provides better discrimination in terms of extension of mismatched termini. The lower processivity of the Stoffel fragment, which also lacks 5'-3' exonuclease activity, is thought to account for this increased discrimination [80,81]. The ASA approach has been modified in different ways in order to increase sensitivity. Unfortunately, each of these modifications has been given a different designation resulting in a bewildering array of acronyms. A number of assays, based on preferential amplification of matched vs. mismatched primers, are not included here because they have been used only for the purpose of genotyping.

Several ASA methods have been applied to the detection of relatively large mutant fractions. These include ASA using the tetra-priming approach [82], Polymerase Amplification of Specific Alleles (PASA) [83,84], Amplification Refractory Mutation System (ARMS) [85], and the mismatched polymerase chain reaction [86]. The priming and ASA strategies used in these different approaches are depicted in Fig. 5. As the name implies, the ASA tetra-priming approach uses four primers (Fig. 5A) [82]. Primers specific for opposite strands of both the mutant and wild-type allele are used. Two outside primers are also included so that different length PCR products are synthesized from the mutant and wild-type templates. The sensitivity of this approach is 1 in 40. PASA is a robust genotyping technique that has also been applied to population screening for base substitutions associated with genetic diseases. PASA was used to detect a mutant allele for phenylalanine hydroxylase associated with phenylketonuria [87] and a factor IX allele associated with hemophilia [35]. In each case, PASA could detect the mutant allele with a sensitivity of 1 in 40. ARMS has been used to screen for a mutant apolipoprotein B allele associated with hypercholesterolemia [36]. Using ARMS, the mutant allele of a single carrier was detected in a pooled blood sample from 50 individuals. The ‘mismatched polymerase chain reaction’ approach was used in the identification of K-ras mutations in DNA from tumor tissue embedded in paraffin sections. Stork et al. [86] estimated a sensitivity of ~1 in 100 for this approach. PASA, ARMS, and mismatched polymerase chain reaction share a similar priming strategy (Fig. 5B).

Mitochondrial DNA point mutations associated with human neuromuscular diseases have been investigated using ASA techniques. Mixtures of mutant and wild-type mitochondrial DNA, termed heteroplasmity, can be present in a given mitochondrion, cell, or tissue. Using an ASA reaction that was followed by restriction endonuclease cleavage, Münscher et al. [38] detected a mitochondrial DNA mutant fraction of 1 in 50. Sensitivities of 1 in 100 and 1 in 1000 have been reported by Seibel et al. [40] and Zhang et al. [39], respectively, for the detection of heteroplasmic point mutations using ASA priming strategies similar to that shown in Fig. 5B.

Mutagenically Separated PCR (MS-PCR) uses two allele-specific primers of different length (Fig. 5C) [37]. Each allele-specific primer has one 3'-terminal mismatch with its specific allele, compared to two with the non-specific allele. Furthermore, the mismatches are in different places and additional mismatches (compared to the DNA template of both alleles) are placed toward the 5' terminus of the longer primer. The 3' terminal mismatches provide specificity in the early rounds of PCR. The additional mismatches make the PCR products generated from the two alleles more distinct than the initial templates. Subsequent rounds of PCR amplification are, therefore, more specific. A sensitivity of 1 in 6400 has been achieved using MS-PCR [37].

Allele-specific Competitive Blocker PCR (ACB-PCR, Fig. 5D) [88] is based on the principles of the ARMS [85] and Competitive Oligonucleotide Priming (COP) [89]. For ACB-PCR, Orou et al. [88]
designed a blocker-primer that has one mismatch in the penultimate position and a perfectly-matched, 3'-terminal dideoxy base (relative to the wild-type allele). The specific primer for the `mutant' allele also has a deliberate mismatch in the 3'-penultimate position (relative to either allele), but its 3'-terminal base matches the rare allele. Therefore, the mutant-specific primer has one mismatch to the mutant template compared to two mismatches to the wild-type template. The inclusion of a competitive non-extendable blocker-primer makes this a relatively robust method with a sensitivity of 1 in 10^8 [88].

There are three reports claiming sensitivities of 1 in 10^6 to 1 in 10^8 using either allele-specific
amplification alone or in conjunction with oligonucleotide specific hybridization [11,90,91]. However, the amount of input DNA in each of these experiments could not contain a sufficient total number of alleles for detection of mutations at the reported mutant fractions (see Section 6.1). While these appear to be relatively sensitive methods, their exact sensitivity is as yet unknown.

The allele-specific selection method with the highest reported sensitivity is the Mismatch Amplification Mutation Assay (MAMA) [92]. MAMA is based on the use of a mutant-specific primer that contains one deliberate mismatch to the mutant allele but two mismatches to the wild-type allele (Fig. 5B). Different double-mismatched primers (relative to the wild-type sequence) were evaluated for their amplification efficiency on wild-type and mutant DNA templates [92]. MAMA was also optimized in terms of cycling and solvent conditions in order to obtain the greatest sensitivity achievable. Using MAMA, Cha et al. [31] were able to detect spontaneous mutations in the 12th codon of the rat H-ras gene that occur at a frequency of $10^{-5}$ and showed that MNU-induced rat mammary tumors arose from preexisting H-ras mutations.

Although ASA techniques comprise a relatively diverse group of methods, they do share some common advantages for genotypic selection. They are relatively simple to perform, rapid, inexpensive, and do not require a restriction endonuclease cleavage site in the mutational target. The small target size, however, must be considered a disadvantage. Some of these methods can only detect a single base substitution, while others are able to detect two different nucleotides at the same position. Some modifications of ASA effectively increase the mutational target size. Primers of different length that match one or another allele have been used together so that multiple nucleotide substitutions at the same base can be detected (MS-PCR, PCR Amplification of Multiple Specific Alleles (PAMSA), and Allele Discrimination by Primer Length (ADPL; Fig. 5C) [37,93,94]. Alternatively, primers specific for different alleles can be labelled with different fluorescent molecules [95]. The application of these approaches to pool screening for rare alleles remains to be demonstrated. Because the literature suggests that sequence context affects the frequency of polymerase misinsertion [86,96,97], the broad applicability of conditions established for any particular sequence is also an issue. Broadly applicable conditions have been used to detect a variety of different base substitutions, but only in the context of large mutant fractions [84]. To establish an assay with high sensitivity will probably require the empirical determination of appropriate conditions for each sequence [92].

### 5.2.2. Blocker-PCR

Like the ASA methods described above, blocker-PCR (Fig. 6) also relies on the specificity of oligonucleotide hybridization to provide allele selection [98]. However, the strategy for blocker-PCR is quite distinct from ASA for two reasons. First, a pair of oligonucleotides is designed so that they perfectly match both strands of the wild-type DNA sequence. These primers each have a 3'-terminal dideoxynucleotide, so that they cannot be extended. Second, these primers hybridize to the mutational target that is located between two perfectly-matched amplification primers. The result is that these wild-type blocker-primers selectively bind to the wild-type DNA sequences and simply block the efficient polymerase elongation through that region of wild-type template DNA. Conditions are chosen such that hybridization of these oligonucleotides to DNA con-

---

**Fig. 6.** Blocker polymerase chain reaction (based on Seyama et al. [98]). Depicted are the mutant template, wild-type template, and the oligonucleotide primers used in blocker PCR. The mutation is indicated by filled rectangles, while the corresponding wild-type bases are indicated by open rectangles. Primers 1 and 2 are the amplification primers. Primers 3 and 4 are the blocker oligonucleotides that each carry a 3'-terminal dideoxynucleotide (indicated by a filled circle). The sequence of the blocker oligonucleotides matches the wild-type allele. Annealing of the blocker oligonucleotides to the mutant template is disrupted by a single central mismatch.
Fig. 7. Peptide nucleic acid (PNA)-mediated PCR clamping (based on Thiede et al. [99]). Depicted are the mutant template, wild-type template, and the oligonucleotide primers indicated by arrows used in PNA-mediated PCR clamping. The mutation is indicated by filled rectangles on the 'DNA strands', while the corresponding wild-type basepair is indicated by open rectangles. The sequence of the PNA molecule (indicated by a hatched rectangle) matches the wild-type allele. Primers 1 and 2 are the amplification primers. The 3'-end of Primer 2 and the PNA molecule each contain short stretches of sequence on the 3'-side of (but not including) the mutation being detected. Molecules drawn outside the DNA strands indicate that annealing to a particular template is not favored.

5.2.3. Peptide nucleic acid-mediated PCR clamping

Peptide Nucleic Acid- (PNA-) mediated PCR clamping (Fig. 7) is a technique that has commonality with both ASA techniques and the blocker-PCR approach [99]. PNA is a oligonucleotide analogue in which bases are attached to an aminoethylglycine oligomer backbone. As in blocker-PCR, PNA molecules are designed so that they will hybridize selectively to the wild-type or abundant allele. Because PNA molecules cannot be extended by DNA polymerases, they act as blocker-primers. The PNA primer is designed so that it covers the target site. One of the amplification primers overlaps with the PNA primer sequence, although it does not include the target sequence. PNA:DNA duplexes are more stable than DNA:DNA duplexes [100], and the PNA will bind to the wild-type sequence more strongly than the amplification primer. If the sequence is mutated, however, the binding of the PNA molecule will be destabilized and the binding of the amplification primer will be favored, thereby providing selective amplification of mutant alleles. Again, this approach is useful for detecting mutations over several basepairs. Using this approach, Thiede et al. [99] were able to detect a codon 12 K-ras mutation with a sensitivity of 1 in 200.

5.2.4. Ligase chain reaction

Two abutting perfectly-matched oligonucleotides annealed to a template molecule are efficiently ligated by the activity of DNA ligase. However, a mispairing at the oligonucleotide junction greatly decreases the efficiency of the joining. This is the basis of the genotypic selection in Ligase Chain Reaction (LCR) [101]. This selection can be used to detect all possible base substitutions because each mismatch causes a decrease in ligation efficiency [102]. In a given pair of oligonucleotides, one oligonucleotide is synthesized such that the terminal base will pair correctly with the particular mutation being detected and mispair with the wild-type allele. The ligation of the oligonucleotides that have annealed to the mutant template DNA generates a new template molecule. The use of thermostable DNA ligase facilitates multiple rounds of denaturation, annealing, and ligation that greatly amplifies the signal from rare mutant alleles [103,104]. Ligase Detection Reaction (LDR) [105] and the Ligase Amplification Reaction (LAR) [106] are strategies that involve two oligonucleotides annealing to a single template strand and result in linear signal amplification (Fig. 8A). The better-known LCR uses two oligonucleotides to anneal to each template strand and results in logarithmic signal amplification (Fig. 8B,C) [105]. Different approaches are available to detect and quantify the ligated product. Oligonucleotide primers can be labelled with 32P, fluorescent dyes, digoxigenin, or biotin, and the ligated product monitored by size using gel electrophoresis [105,107,108]. In more sophisticated strategies, one primer of a pair is labelled with a molecule that binds a solid support and the other primer of a pair is labelled with a reporter molecule [109,110]. Therefore, only the ligated products are recovered and detected. In some cases,
primers that are specific for different base substitutions and vary slightly in length can be included in the same reaction, thereby somewhat increasing the mutational target [105].

Kalin et al. [107] evaluated LCR for the detection of a mouse H-ras mutation at codon 61. An enzyme-linked immunosassay with a digoxigenin marker was used that required 250 molecules of mutant for signal detection. Because the upper limit for DNA template in their LCR was 4 μg, they concluded that the theoretical limit of detection for this approach was 1 in 10^4. In practice, however, they were unable to distinguish the signals from a 1 in 10^2 mutant fraction and a no-mutant control due to a high background. The high background observed was probably due to target-independent ligation. This occurs when complementary primers form duplex molecules with blunt-ends that can then be ligated (Fig. 8B). Designing complementary primers with single basepair overhangs reduces this template-independent ligation (Fig. 8C) [101,105].

Gap-LCR [111] and pLCR [101,112] are approaches that combine the genotypic selections of ASA and LCR (Fig. 8D). Primer pairs are chosen such that there is a gap of one or a few bases between them when hybridized to the template. The complementary primer pairs are designed to have one or a few terminal nucleotide overhangs to prevent blunt-end ligation [101,105]. The mutational target is at the 3'-end of one of the primers that is adjacent to the gap. The sequence of this primer is such that the 3'-terminal or penultimate base matches the mutant allele being detected. The primers are hybridized, then extended with a polymerase that lacks 3'-5' exonuclease function. The gap is designed so that it can be filled without using all four dNTPs because omitting at least one dNTP increases the specificity of the extension reaction. This reaction also contains a thermostable ligase. Correctly-

Fig. 8. Strategies for ligase-mediated amplification of specific alleles. The long lines represent the template molecules. The shorter lines, drawn within the template, indicate the oligonucleotides present in a given reaction. A terminal mismatch between the template and an annealed oligonucleotide is indicated by curving of the ‘oligonucleotide’ line away from the template. (A) The Ligase Detection Reaction (LDR) or Ligase Amplification Reaction (LAR). Two abutting oligonucleotides are used. A terminal nucleotide of one of the oligonucleotides matches the mutant allele but not the wild-type allele. (B) Ligase Chain Reaction (LCR) using oligonucleotides without terminal overhangs. Two complementary pairs of oligonucleotides are used. When oligonucleotides are designed without terminal overhangs, complementary pairs can anneal and be ligated together. Annealed oligonucleotide pairs are depicted below the template. (C) LCR using oligonucleotides with terminal overhangs. Two complementary pairs of oligonucleotides are used. Annealed oligonucleotide pairs are depicted below the template. (D) Gap-LCR and p-LCR. One oligonucleotide of each pair has a 3'-nucleotide that matches the mutant but not the wild-type sequence. These oligonucleotides can anneal to the mutant template and act as primers for gap-filling with a thermostable polymerase. Gap-filling is followed by ligation. In this example, modification of oligonucleotides with a recoverable label or reporter molecule is also depicted.
matched primers, therefore, can be extended and ligated to form the template for additional rounds of this reaction. Abravaya et al. [111] linked this approach to an immunoassay for the detection of a human immunodeficiency virus mutation associated with AZT resistance. They were able to detect this mutation with a sensitivity of 1 in 10000 and showed that the Gap-LCR approach was more discriminating than allele-specific PCR amplification when used alone.

Although LCR generally is being developed and applied as a sensitive sequence detection method, ligase-mediated allele enrichment has some advantages as a genotypic selection method. The assay is relatively simple to perform and may be coupled with other genotypic selection methods [78,113]. Coupling the digestion of wild-type restriction endonuclease recognition sequences to LCR may simplify the RFLP/PCR approach while retaining a high level of sensitivity [113]. In addition, the assay can be used to detect multiple substitutions at a single position by using allele-specific oligonucleotides that differ in length [105].

5.3. Methods that spatially separate the mutant and wild-type alleles

5.3.1. Single-strand conformation polymorphism

Single-Strand Conformation Polymorphism (SSCP) is based on the differential electrophoretic mobility of single-strand DNA molecules that differ by a single base [114–116]. SSCP qualifies as a genotypic selection method because an enriched minor-allele fraction may be obtained by spatial separation from the abundant or wild-type allele. PCR followed by SSCP (PCR–SSCP) has been used primarily for genotyping. In a few instances, however, PCR–SSCP has been evaluated for its sensitivity in detecting mutations in mixed cell populations and sensitivities of 1 in 4 to 1 in 33 have been reported [117–121]. But, even when the mutant fraction is only 1 in 2, PCR–SSCP does not always discriminate between the mutant and wild-type alleles [115,116,122–124]. Clearly, SSCP has little power as a genotypic selection method. Its simplicity may, nevertheless, foster its use as part of more complex genotypic selection schemes. SSCP has been used in conjunction with blocker-PCR and augmented its sensitivity ten-fold [98].

5.3.2. Denaturing gradient gel electrophoresis

Denaturing Gradient Gel Electrophoresis (DGGE) [125,126] is a technique developed by Fischer and Lerman [127] that can identify homoduplex molecules that differ by single basepair substitutions. Duplex DNA molecules that have high- and low-melting domains move through a polyacrylamide gel containing an increasing concentration of denaturant until part of the molecule becomes single-stranded or ‘melts’ [126]. This melting greatly impedes the mobility of the molecule. Because the melting of a particular domain is very sequence dependent, single basepair substitutions alter the electrophoretic mobility of otherwise identical DNA molecules. This technique is even more effective for separating homoduplex DNA from heteroduplex molecules because even single mismatches destabilize a melting domain. Therefore, heteroduplex molecules usually become partially single-stranded (melt) at a lower concentration of denaturant than the corresponding homoduplex sequence [126]. Initially, DGGE could only be used to separate duplexes that had a base substitution in the low-melting domain of its DNA sequence. The addition of a GC-clamp (an artificial, high-melting domain) to the end of a DNA duplex, however, now permits the separation of duplex molecules with base substitutions in high-melting domains [128–131]. The ability to spatially separate nearly identical alleles is the genotypic selection provided by DGGE.

DGGE is a widely-used technique with many applications [132]. In terms of DNA pool screening, however, its major application has been the analysis of 6-thioguanine-resistant mutants in the human hprt gene [133–142]. In this case, a large pool of mutants has already been selected, phenotypically, but those mutations are spread over a large region of DNA. DGGE is then used to select (genotypically) mutants in a more narrowly defined region of DNA from the original pool of mutants. At the same time, DGGE provides the means to measure the frequency of a particular mutation and to extract it from the pool of sequences for identification of the particular mutational event.

This approach has been applied to determine the human mutational spectra of a variety of genotoxi-
cants, including methyl-nitro-nitroso-guanidine, ICR-191, ultraviolet light, oxygen, hydrogen peroxide, benz[a]pyrene-diol epoxide, cisplatin, X-rays, 4-nitroquinoline-1-oxide, chromium (VI), and aflatoxin B₁. Large cultures of ~1 × 10⁸ cells are exposed to the toxicant. The human B lymphoblastoid cell line, TK6, has been used most often, although the h1A2v2 cell line that expresses P450 CYP1A1 has also been used [142]. Cells carrying mutations in the hprt gene are selected by culture in 6-thioguanine. The DNA that is isolated from this culture represents a complex mixture of 10³ to 10⁴ mutants. This DNA is used as template to amplify exon 3 by high fidelity PCR. In some cases, a GC-clamp is added. The PCR product is denatured and reannealed in order to produce heteroduplex molecules that are then analyzed by DGGE. Mutants are visualized as relatively minor bands with altered mobility as compared to the wild-type homoduplex DNA. The use of multiple, treated- and untreated-cultures helps to identify bands as spontaneous or induced mutations. The sensitivity of mutant detection using this approach is about 1 in 10² [133,135,138,142], although detection of a mutant fraction of 1 in 10³ in a reconstruction experiment has been reported [137]. This means that from the pool of 10³ to 10⁴ mutants, only those represented by a mutant fraction of ~1% or more can be detected. This approach, therefore, can only be used to identify mutational hotspots in a given segment of a phenotypically selectable gene [136].

A major advantage of DGGE is the very large target size compared to the other genotypic selection methods described. The suitability of DGGE for a number of applications is based on this large target size. DGGE also has been used in the analysis of mutations in tumor samples [143,144], and DGGE of PCR products of 16S rRNA genes has been used to characterize microbial populations [145–147]. Muyzer et al. [145] demonstrated that a particular sulfate-reducing bacteria could be detected in a population provided it represented 1% of the microbes.

5.3.3. Constant denaturant capillary electrophoresis

Constant Denaturant Capillary Electrophoresis (CDCE) combines the techniques of Constant Denaturant Gel Electrophoresis (CDGE) and replaceable linear polyacrylamide capillary electrophoresis [148]. CDGE is a modification of DGGE. Like DGGE, CDCE is based on a difference in melting temperature between homoduplex DNA and heteroduplex DNA, but CDCE provides improved spatial separation of alleles [149,150]. CDGE uses polyacrylamide gels of a single denaturant concentration. The resolution of homoduplex and heteroduplex molecules in such a gel is a simple function of the distance travelled (i.e. the longer the gels are run, the better the separation) [149]. This is in contrast to DGGE where partially denatured molecules continue to move slowly through the gel into higher concentrations of denaturant which may cause additional conformational change that further affects their migration.

Capillary electrophoresis is, to some extent, replacing slab gel electrophoresis in a variety of mutation detection methods because of its high resolving power and speed [151–154]. For instance, Mitchell et al. [154] performed PCR/RFLP using capillary electrophoresis. The reported advantages of this approach were short analysis time, single base resolution of DNA fragments > 100 bases, small sample requirement, automation, and quantitation. Part of the development of capillary electrophoretic technology has been the use of replaceable linear polyacrylamide matrices [155,156].

The applicability of CDCE to pool screening for mutations was demonstrated by Khrapko et al. [148]. PCR products corresponding to a 206 basepair region of mitochondrial DNA were synthesized and isolated. A primer carrying a single base change was used in one reaction in order to produce a pair of DNA fragments that differed by a single basepair substitution. Some primers were 5'-end-labeled with fluorescein. The PCR products were mixed, denatured, and reannealed to form various heteroduplexes and their separation was monitored and optimized by CDCE. The partial denaturation of molecules takes place in a constant-temperature, denaturing-zone instead of a slab gel. A mutant fraction of 0.03% (1 in 3.3 × 10⁸) was detected. Furthermore, it was determined that as many as 10¹¹ molecules could be loaded onto the capillary and as few as 3 × 10⁴ molecules of double-stranded DNA could be detected using laser-induced fluorescence of the fluorescein-tagged DNA [148]. CDCE was further evaluated for its sensitivity for detecting mutations in codons 12, 13, and 61 of the human N-ras gene [157]. Using similar methodology to that described.
above it was determined that a single round of PCR and CDCE could detect mutant fractions as low as 1 in $10^3$. However, high-fidelity PCR reamplification of heteroduplex-containing CDCE fractions and a second round of CDCE increased the sensitivity between 2- and 10-fold [17,157]. This finding that the sensitivity of mutation detection can be increased by using multiple rounds of CDCE was confirmed and extended in an analysis of human mitochondrial DNA [158]. Using a somewhat complicated series of steps (preparative CDGE, high-fidelity PCR, CDCE, high fidelity PCR, and high resolution CDCE), mitochondrial DNA mutations as rare as 1 in $10^6$ were detected based upon their altered melting temperatures [158]. Krapko et al. [158] also describe mutant quantification based on an internal copy number standard and specific tests for a variety of potential artefacts. The contributions of polymerase errors and unrepaired mutagenic lesions and mismatches in the test DNA samples to the generation of mutant signals were evaluated. Of the genotypic selection methods based solely upon spatial separation of mutant and wild-type alleles, CDCE is clearly the most sophisticated and the most powerful.

5.3.4. Competitive mobility shift assay

The Competitive Mobility Shift Assay (CMSA) [159] is very similar to the liquid hybridization/gel retardation assay described by Kumar and Barbacid [71]. This approach appears to be sensitive and yet relatively simple to perform. A PCR product is synthesized from genomic DNA and an aliquot of this product is then used in a second, asymmetric PCR amplification that produces a single-strand DNA template. Oligonucleotides that contain the mutant or wild-type base near the center of the molecule are synthesized and the mutant oligonucleotide is $^{32}$P-labelled. The single-strand DNA template and the mutant and wild-type oligonucleotides are added to a hybridization reaction. Slow ramping through the temperature range expected for hybridization allows perfectly-matched hybrids to form before any mismatched hybrids can form. The hybrid molecules are then separated and identified by gel electrophoresis. Including wild-type, ‘competitor’ oligonucleotides in the reaction increases the discrimination of the selection. The genotypic selection of the CMSA is based on the specificity of nucleic acid hybridization and the spatial separation of template:oligonucleotide hybrids. Using this approach, Chen et al. [159] reported the detection of human $K$-ras mutations with a sensitivity of 1 in $10^6$. However, this sensitivity seems high for a method initiated by two sequential Taq polymerase amplification steps (see Section 6.3).

5.3.5. Magnetic mismatch binding beads

The spatial separation, enrichment, and isolation of mutant alleles has been performed using magnetic beads with attached mismatch binding protein ($M_2B_2$, Genecheck, Fort Collins, CO). Mixtures of cloned mutant and wild-type human glucokinase DNA were used as the template in a PCR in which one of the primers was $5'$-biotinylated. Heteroduplex molecules were generated by denaturation and reannealing and then purified using the magnetic beads. Using a colorimetric assay based on the biotin-label, it was shown that mutant fractions as low as 1 in 1000 could be measured (Robert Wagner, personal communication).

6. Important considerations for genotypic selection

Genotypic selection methods are proliferating. The power of these assays, generally referred to as sensitivity, lies in their ability to discriminate between nearly identical alleles. In order to evaluate rationally this complex and growing literature, one must be aware of the difficulties and limitations inherent in genotypic selection. In fact, there are a number of issues that must be considered when evaluating, choosing, or designing a genotypic selection scheme. These include initial DNA pool size, the number of mutants required for detection, the use of an internal amplification or copy number standard, polymerase fidelity, and the most effective use of PCR.

6.1. Initial DNA pool size

The appropriate number of copies of target sequence for a genotypic selection assay depends on the sensitivity of the assay. The number of target molecules must be equal to or greater than the number of mutant alleles required for detection divided by the mutant frequency. If the goal is to
measure mutations in a single copy nuclear gene, for example, the expected mutation frequency might be $1 \times 10^{-7}$. Furthermore, the genotypic selection method might require at least 10 mutant molecules in order to produce a detectable signal (for a statistical discussion of initial numbers of mutant molecules, see [7,158]). Under these circumstances, a DNA pool of no less than $10^8$ mutational targets must be used. Given that 1 μg of mammalian DNA is equivalent to $\sim 3 \times 10^5$ copies of the genome [160], this particular genotypic selection must be initiated using at least 333 μg of DNA per sample. Surprisingly, there are a number of examples in the literature where the reported sensitivity of detection is larger than the total number of target sequences analyzed [11,90,91].

The issue of DNA pool size is a significant problem for at least two reasons. Elaborating on the example above, a DNA pool size of $10^8$ (for a single copy target) corresponds to $\sim 0.5$ grams of tissue or 100 ml of blood [17]. These relatively large tissue samples are an obstacle for human mutational analyses, particularly if multiple samples are needed. Depending on the sensitivity of the genotypic selection, therefore, tissue samples obtained through biopsy or autopsy may be required [16]. Furthermore, some rodent tissues may not contain the necessary number of cells [4] and tissues from several animals might have to be pooled. A second problem is that the techniques required for allele-selection often function poorly in the presence of large amounts of extraneous DNA and require some type of gene-specific enrichment step. Preparative gel electrophoresis has been used in conjunction with RFLP/PCR (Fig. 2) [53]. Specific nucleic acid hybridization is another approach for enriching the target sequence. Restriction fragments have been enriched by hybridization with target-specific, biotin-labeled RNA sequences that are subsequently recovered using streptavadin-agarose beads [24]. Hybridization of restriction fragments to target-specific sequences covalently bound to a nylon membrane is another possible approach (unpublished observations).

The use of multicopy targets circumvents some problems associated with the necessity of using a large pool of target sequences. Ribosomal RNA genes, for instance, are present in hundreds of copies per cell [4,8]. Because there are between $10^3$ and $10^4$ mitochondria per cell, a pool of $2 \times 10^3$ mitochondrial genomes can be obtained from only 1 ml of blood [17]. However, using these mutational targets negates a major advantage of genotypic selection namely, the ability to analyze targets with known relevance to carcinogenesis.

6.2. Quantitative PCR and the use of an internal copy number standard

Given the inherent difficulty in developing genotypic selection methods, it is not surprising that relatively little attention has thus far been paid to the quantitative aspects of the PCR used in these methods. The sensitivity of genotypic selection methods are most often demonstrated by performing reconstruction experiments in which decreasing amounts of mutant is added to known amounts of wild-type allele. Only rarely has this approach clearly demonstrated that the assay is quantitative [18,19,56,72,88,111,158]. Sometimes the sensitivity reported is determined by the lowest mutant fraction that can be distinguished from the no-mutant control, even when there is no clear dose-response at lower dilutions [11,98,161]. Furthermore, when a dose response is observed, ten-fold decreases in mutant fraction often result in less than a two-fold decrease in signal [11,12,68,91,98]. Yet, the genotypic selection methods described are clearly intended to be quantitative.

Considerations for quantitative-PCR (Q-PCR) have been reviewed [162,163]. The use of a standard, the contribution of precision and reproducibility in analysis of the variability, and the linearity of amplification are Q-PCR issues that clearly apply to any PCR used in genotypic selection. In fact, the more sensitive and complicated a selection/amplification scheme, the more critical it becomes that the sources of variability are understood. The PCR must be performed in the linear range of amplification [162]. This means measuring the product generated with relatively few PCR cycles and probably requires more sensitive detection methods than ethidium bromide staining (e.g., radiolabeling, fluorescent, or chemiluminescent detection) [163]. Ideally, the range of mutant fractions over which the particular assay is quantitative should be determined and reported along with the lowest mutant fraction that was detectable. Analyzing replicate samples in the
context of a given experiment and between experiments is helpful for determining the precision and reproducibility of a measurement. This is particularly important when small numbers of mutant molecules (e.g., below 10 copies) are being detected because sampling errors are expected. Finally, to obtain an exact, rather than relative, quantitation, a copy number standard must be used. Signal comparison between the mutant test DNA and the copy number standard will provide the most accurate estimate of the mutant fraction if the standard is roughly the same size and is amplified with the same primers as the test DNA [162]. Furthermore, the test DNA and the amplification standard should be in comparable abundance when the PCR is initiated. The use of an internal standard is preferable to a standard measured in parallel because small differences in PCR conditions could affect significantly the amount of product formed [164]. Ideally, the internal amplification standard should be subject to the same genotypic selection and PCR amplification as the mutant target and yet be distinguishable from that target. RFLP/PCR (Fig. 2) [2] and the CDCE approach [158] are the only genotypic selection methods in which an internal amplification standard has been used.

6.3. Fidelity of PCR and the linking of two or more genotypic selection methods

PCR is a valuable tool for genotypic selection. Nevertheless, errors that are produced by in vitro DNA amplification may obscure the detection of rare mutational events [160,165,166]. Once they arise, such polymerase errors continue to be amplified and accumulate in a reaction [160]. Therefore, the fidelity imparted by the reaction conditions and the thermostable DNA polymerases are important considerations.

Information regarding the fidelity of thermostable polymerases has been derived using a variety of methods. Polymerase fidelity has been studied using: (1) a M13mp2 forward mutational assay [167], (2) an M13mp2 reversion assay [167], (3) an assay involving PCR, cloning and plaque screening of a lacIOZa mutational target [54], (4) an assay involving PCR and plasmid cloning of a lacZ mutational target [168], (5) DGGE [43,44,169–171], (6) PCR/RFLP [66], (7) a p53-based PCR fidelity assay [172], (8) a MutHLS-dependent cleavage assay [173], and (9) by DNA sequencing [174]. For each assay, the mutational target and the variety of mutations that can be detected are different, thus making it somewhat difficult to interpret the reported misincorporation rates.

Once a mutant frequency has been determined using one of these assays, a polymerase error rate is calculated. The error rates that have been determined using the techniques described above, and are primarily related to basepair substitution errors, are summarized in Table 1. Although the range of Taq polymerase error rates is quite large, $3.1 \times 10^{-4}$ to $3.2 \times 10^{-6}$ errors per basepair duplicated, the majority of these error rates fall between 1.0 to $2.0 \times 10^{-4}$ errors per basepair duplication. Generally, the thermostable polymerases with 3'-5' proofreading activity have higher fidelity than those without this activity [44,54,140,170–175]. Of these proofreading polymerases, Pfu polymerase has the lowest error rate and UlTma polymerase has the highest [54,140,172,173]. While the relative fidelities of these polymerases is apparent, their absolute fidelities can only be precisely known in the context of a particular assay. The increased fidelity attributed to the 3'-5' exonuclease activity of some thermostable polymerases has clear advantages for genotypic selection. Unfortunately, this activity can also result in the degradation of PCR primers, the sequence of which is often critical for genotypic selection [176]. Introduction of a single 3'-terminal phosphorothioate bond into a PCR primer may circumvent this problem [176,177].

The most sensitive DNA-based mutation detection methods will include PCR and a number of genotypic selection steps. Using PCR as the first step in a genotypic selection scheme will limit the sensitivity of the assay to the fidelity of the polymerase. In RFLP/PCR, there are two different selection techniques used (Fig. 2). The initial selection is the destruction of the wild-type sequence by restriction digestion. The non-digested, mutant sequences are PCR amplified and then cloned. Cloning of individual sequences and amplification in spatially-separated plaques is another technique that provides allele-enrichment. Using this selection scheme, where PCR amplification follows the initial selection, mutations that occur as infrequently as 1 in $10^7$ have been detected [51]. In contrast, 1 in $10^7$ is the best re-
ported sensitivity for a PCR/RFLP method in which Taq polymerase amplification is the first step in the genotypic selection [71]. The theoretical limits of sensitivity in assays initiated by PCR are not more than 1 in $3.3 \times 10^5$ for Taq polymerase, 1 in $4.3 \times 10^5$ for Vent™ polymerase, and 1 in $1.3 \times 10^6$ for Pfu polymerase. These limits are based on the lowest reported error rate for each polymerase and, therefore, are likely to be overestimated. One must also recognize that the limit of polymerase fidelity, and consequently the assay’s sensitivity, is expected to have a sequence-dependent component [178].

The importance of high-fidelity PCR amplification in genotypic selection necessitates care in the selection and optimization of PCR conditions. Factors that have been shown to affect polymerase fidelity include dNTP concentration, MgCl₂ concentration, annealing and extension times, annealing temperature, the addition of solvents (e.g., glycerol, Triton X-100, formamide), cycle number, and pH.

Table 1
Fidelity of thermostable DNA polymerases

<table>
<thead>
<tr>
<th>Polymerase b</th>
<th>3'-5' exonuclease</th>
<th>Error rate c</th>
<th>Assay</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taq</td>
<td>Absent</td>
<td>$2.0 \times 10^{-4}$</td>
<td>DNA sequencing PCR product</td>
<td>[174]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$1.1 \times 10^{-4}$</td>
<td>M13mp2 reversion assay</td>
<td>[167]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$1.1 \times 10^{-4}$</td>
<td>M13mp2 forward assay</td>
<td>[167]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$1.9 \times 10^{-4}$</td>
<td>DGGE</td>
<td>[43]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$0.03 \times 10^{-4}$</td>
<td>M13mp2 reversion assay</td>
<td>[180]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$1.7 \times 10^{-5}$</td>
<td>M13mp2 forward assay</td>
<td>[180]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$8.9 \times 10^{-5}$</td>
<td>DGGE</td>
<td>[44]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$2 \times 10^{-4}$</td>
<td>PCR/RFLP</td>
<td>[66]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$2 \times 10^{-5}$</td>
<td>PCR/cloning/plaque assay of lacZ</td>
<td>[54]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$0.7 \times 2.0 \times 10^{-4}$</td>
<td>DGGE</td>
<td>[170]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$1.3 \times 10^{-4}$</td>
<td>DGGE</td>
<td>[171]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$9.5 \times 10^{-5}$</td>
<td>DGGE</td>
<td>[140]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$3.2 \times 10^{-6}$</td>
<td>MutHLS-dependent cleavage</td>
<td>[173]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$1.1 \times 1.3 \times 10^{-4}$</td>
<td>PCR/plasmid cloning assay of lacZ</td>
<td>[168]</td>
</tr>
<tr>
<td>Klen Taq</td>
<td>Absent</td>
<td>$5.0 \times 5.1 \times 10^{-5}$</td>
<td>PCR/plasmid cloning assay of lacZ</td>
<td>[168]</td>
</tr>
<tr>
<td>Vent™</td>
<td>Present</td>
<td>$1.0 \times 2.0 \times 10^{-5}$</td>
<td>M13mp2 reversion assay</td>
<td>[175]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$4.5 \times 6.6 \times 10^{-5}$</td>
<td>DGGE</td>
<td>[170]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$2.4 \times 10^{-5}$</td>
<td>DGGE</td>
<td>[44]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$5.3 \times 10^{-5}$</td>
<td>DGGE</td>
<td>[171]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$1.0 \times 7.3 \times 10^{-5}$ d</td>
<td>p53-based PCR fidelity assay</td>
<td>[172]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$2.3 \times 10^{-6}$</td>
<td>MutHLS-dependent cleavage</td>
<td>[173]</td>
</tr>
<tr>
<td>Vent™ (Exo-)</td>
<td>Absent</td>
<td>$0.6 \times 1.5 \times 10^{-4}$</td>
<td>M13mp2 reversion assay</td>
<td>[175]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$2.7 \times 10^{-5}$</td>
<td>DGGE</td>
<td>[171]</td>
</tr>
<tr>
<td>Pfu</td>
<td>Present</td>
<td>$1.6 \times 10^{-6}$</td>
<td>PCR/cloning/plaque assay of lacZ</td>
<td>[54]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$2.5 \times 10^{-5}$</td>
<td>DGGE</td>
<td>[140]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$2.0 \times 10^{-6}$ d</td>
<td>p53-based PCR fidelity assay</td>
<td>[172]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$7.9 \times 10^{-7}$</td>
<td>MutHLS-dependent cleavage</td>
<td>[173]</td>
</tr>
<tr>
<td>Ultma</td>
<td>Present</td>
<td>$3.2 \times 8.4 \times 10^{-5}$ d</td>
<td>p53-based PCR fidelity assay</td>
<td>[172]</td>
</tr>
</tbody>
</table>

a Fidelity generally refers to basepair substitution errors, the most confounding type of errors in genotypic selection techniques.

b Both native and recombinant forms.

c Error rates are given in terms of errors per basepair duplicated (errors per base duplicated in the case of the M13mp2 mutational assays).

d Error rates were calculated based on the assumption that the Taq polymerase error rate is $2.0 \times 10^{-5}$. 
DNA damage of the template is another source of background errors and template exposure to DNA damaging agents including ultraviolet light and heat should be minimized. Introduction of DNA adducts may be an integral aspect of a mutagenesis or carcinogenesis protocol, but such adducts also have the potential to confound genotypic selection techniques. Using the RSM or RFLP/PCR approach, for example, wild-type but adducted bases may block restriction digestion and result in a false positive signal. A false positive signal might also be generated on adducted DNA template using the MutEx approach because MutS binds to adducted basepairs in addition to mismatched basepairs. These complications can be mitigated by allowing sufficient time after mutagen exposure for damage repair or dilution by DNA replication. Also, genotypic selection techniques could be used to analyze both strands of the sequence of interest, thereby identifying any discrepancies due to adducts or unresolved mismatches in the DNA.

7. Summary and conclusions

Applications for genotypic selection techniques are likely to expand in the future. The Human Genome Project should provide more examples of genes and gene mutations related to human disease and stimulate the application of DNA-based, pool screening for genetic defects. As our understanding of the underlying genetics of cancer grows, so does the number of targets that should be analyzed by genotypic selection. Genotypic selection also represents a viable alternative for evaluation of xenobiotics perhaps as part of a battery of tests that could replace rodent bioassays. These and other important applications of DNA-based mutation detection will stimulate the continued development of genotypic selection methods.

Acknowledgements

We would like to thank the NCTR library staff for their considerable and very capable assistance. We also thank Dr. Robert Wagner for providing unpublished data on Genecheck’s M2B2, magnetic bead mismatch binding assay.

References


large-scale screening of pooled blood samples, J. Lipid Res. 33 (1992) 1563–1567.


[95] F.F. Cheehab, Y.W. Kan, Detection of specific DNA sequences by fluorescence amplification: a color complemen-


[155] Y.F. Pariat, J. Berka, D.N. Heiger, T. Schmitt, M. Vi-


175. C.M.C. de Noronha, J.I. Mullins, Amplimers with 3’-terminal phosphorothioate linkages resist degradation by Vent polymerase and reduce Taq polymerase mispriming, PCR Methods Appl. 2 (1992) 131–136.


