

REVIEW ARTICLE

Molecular Approaches in the Diagnosis of Primary Immunodeficiency Diseases

Maurizio Costabile,^{1,2*} Alex Quach,^{1,2} and Antonio Ferrante^{1–3}¹School of Pharmacy and Medical Sciences, University of South Australia, Adelaide, Australia; ²Department of Immunopathology, Children's, Youth and Women's Health Services of South Australia, Adelaide, Australia; ³Department of Paediatrics, Children's, Youth and Women's Health Services of South Australia, Adelaide, Australia

For the Immunogenetics Special Issue

Over 120 inherited primary immunodeficiency diseases (PIDs) are known to exist. The genes responsible for many of these diseases have also been identified. Recent advances in diagnostic procedures have enabled these to be identified earlier and appropriately treated. While a number of approaches are available to identify mutations, direct sequencing remains the gold standard. This approach identifies the exact genetic change with substantial precision. We suggest that a sensitive and economical approach to mutation detection could be the direct sequencing of cDNA followed by the confirmatory sequencing of the corresponding exon. While screening techniques such as single-stranded conformation polymorphism (SSCP), heteroduplex analysis (HA), denaturing gradient gel electrophoresis (DGGE), and denaturing high-performance liquid chromatography (dHPLC) have proven useful, each has inherent advantages and disadvantages. We discuss these advantages and disadvantages and also discuss the potential of future sequencing technologies such as pyrosequencing, combinatorial sequencing-by-hybridization, multiplex polymerase colony (polony), and resequencing arrays as tools for future mutation detection. In addition we briefly discuss several high-throughput SNP detection technologies. *Hum Mutat* 27(12), 1163–1173, 2006. © 2006 Wiley-Liss, Inc.

KEY WORDS: immunology; immunogenetics; primary immunodeficiency diseases; mutation detection; dideoxy sequencing; resequencing; SNP detection

INTRODUCTION

The human immune system is responsible for protecting the host against infections caused by bacteria, viruses, fungi, and parasites. This protection is achieved through both nonspecific (innate) and specific (adaptive) immune responses. Defects in either of these components of the immune system leads to disease and may ultimately result in death. This is particularly apparent in primary immunodeficiency diseases (PIDs), which are caused by congenital defects of the immune system [Bonilla and Geha, 2006]. While several diagnostic criteria exist for determining if an individual is suffering from a PID [Conley et al., 1999], the hallmark signs are the presentation of recurrent infections with pathogens to which the host normally displays effective resistance, such as *Pneumocystis carinii*. The laboratory-based diagnosis of PIDs involves enumerating leukocyte cell numbers and evaluating their functional responses. While these approaches are useful in classifying the type of PID, their classification cannot be definitively established without determining the genotype. Since the genes responsible for many PIDs are known, genetic typing is possible [Notarangelo et al., 2006]. However, a large number of different molecular techniques have been used to diagnose PIDs and it can often be difficult to decide which is the most appropriate to use.

We suggest that mutation analysis could commence with the direct sequencing of the cDNA of the gene of interest, and confirming the presence of any mutation in the corresponding genomic exon. We discuss some of the most commonly practiced and well-established techniques used to screen for disease causing mutations. These include single-stranded conformation

polymorphism (SSCP), heteroduplex analysis (HA), denaturing gradient gel electrophoresis (DGGE), and denaturing high-performance liquid chromatography (dHPLC). The advantages and disadvantages of each approach are discussed below. The potential of novel sequencing technologies and single-nucleotide polymorphism (SNP) identification will also be discussed.

CLASSICAL APPROACHES TO PID DIAGNOSIS

The diagnosis of an individual with frequent recurrent infection, an indicator of the presence of a PID, normally involves the assessment of blood leukocyte profiles and a functional analysis. A differential count will determine if the major leukocyte populations are present within a normal reference range. In addition, when combined with flow cytometry [Illloh, 2004], the ratio of individual cell populations can also be determined. For example, the surface expression of CD40 ligand (CD40L) is reduced or absent in X-linked hyper-immunoglobulin M (hyper-IgM) syndrome (MIM# 308230).

The Supplementary Material referred to in this article can be accessed at <http://www.interscience.wiley.com/jpages/1059-7794/suppmat>.

*Correspondence to: Dr. Maurizio Costabile, School of Pharmacy and Medical Sciences, University of South Australia, GPO Box 2471, Adelaide, South Australia, Australia.

E-mail: maurizio.costabile@unisa.edu.au

DOI 10.1002/humu.20412

Published online 7 September 2006 in Wiley InterScience (www.interscience.wiley.com).

One caveat in relation to the use of flow cytometry should be noted. Cell surface proteins can be mutated, but may still be expressed on the surface of the cell, leading to the conclusion that the protein is not mutated. Amino acid substitutions may be present in the cytoplasmic domain of the protein, rendering the receptor nonfunctional, although mutations may also exist in the extracellular domain of the receptor. Flow cytometry is unable to distinguish this difference and diagnosis should be coupled with cell specific functional studies. For example, the lack of in vitro IgG production following CD40L cross-linking in patients with suspected X-linked hyper-IgM syndrome.

Western blot can also be used to determine the presence/absence of a protein and thus aid in the diagnosis of PIDs. The level of protein can allow for variants of disease to be identified. For example, X-linked chronic granulomatous disease (CGD) (MIM# 300481) exists in three variant forms. Mutations in the gene, *CYBB*, which encodes for the protein gp91^{phox}, can lead to this protein being present (yet nonfunctional), reduced, or absent in a patient's neutrophils [Cross et al., 1995; Roos et al., 1996].

Most severe combined immunodeficiencies (SCID) present with individual leukocyte populations being absent or significantly reduced. For example, a significant reduction in circulating T-lymphocytes, abnormal lymphocyte proliferation, abnormal facies, and the absence of a thymic shadow are indicative of DiGeorge syndrome (MIM# 188400) [Wilson et al., 1993]. Individuals with recurrent bacterial infections also have serum immunoglobulin and complement levels routinely measured either by nephelometry or enzyme-linked immunosorbent assay (ELISA) [Ferrante et al., 1986].

The assessment of cellular functional responses also provides additional information about the status of individual leukocyte populations. For example, the proliferation of T- and B-lymphocytes can be assessed in response to mitogenic or antigenic stimuli [Amedei et al., 2001], while neutrophils can have their ability to migrate [Zicha et al., 1997], produce reactive oxygen species [Jirapongsananuruk et al., 2003], and to kill bacteria and fungi assessed in vitro [Wolach et al., 2005]. Each of these investigations provides the physician with the ability to predict the disease, and in turn the gene(s) to be examined.

MOLECULAR APPROACHES TO TYPING PIDs

PID Databases

Significant progress has been made in defining the molecular cause of many of the currently-classified PIDs. Numerous disease specific databases are publicly available (Table 1) [Piiirilä et al., 2006]. These resources relate to an individual or group of diseases, and include the type and spectrum of mutations (Supplementary Table S1; available online at <http://www.interscience.wiley.com/jpages/1059-7794/suppmat>). Relevant research publications that detail the particular molecular approach used to diagnose the particular disease are also highlighted, and should therefore be consulted as a starting point in the analysis of a gene of interest. Screening techniques are used to rapidly identify defects in a gene being investigated and to facilitate the diagnostic process. We suggest that mutation detection for these diseases should commence with the sequencing of the cDNA of the gene under investigation followed by sequencing of the corresponding genomic exon (Fig. 1 and 2), since many of the screening techniques do not display 100% sensitivity.

TABLE 1. Listing of Databases and Societies Relevant to the Study of PIDs

Resource	URL
IMT-Bioinformatics	http://bioinf.uta.fi/
IMT-ID Resource	http://bioinf.uta.fi/idr/
IMT-ID-bases	http://bioinf.uta.fi/base_root/
ESID	www.esid.org/
HGVS	www.hgvs.org/
HGVbase	http://hgvsbase.cgb.ki.se/
MIM	www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=MIM
HGMD	www.hgmd.cf.ac.uk/
HUGO	www.gene.ucl.ac.uk/nomenclature/index.html
IPD	www.ebi.ac.uk/ipd/
Entrez-SNP	www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=snp

ESID, European Society for Immunodeficiency; HGMD, Human Gene Mutation Database; HGVS, Human Genome Variation Society; HGVbase, Human Genome Variation database; HUGO, Human Genome Organisation; IMT-Bioinformatics, Institute of Medical Technology Bioinformatics; IMT-ID bases, Institute of Medical Technology Immunodeficiency Databases; IMT-ID resource, Institute of Medical Technology Immunodeficiency Resources; IPD, Immuno Polymorphism Database; MIM, Online Mendelian Inheritance in Man.

SCREENING METHODS

Mutation detection strategies can be divided into two categories: those that identify specific, previously characterized sequence variations, and those that detect uncharacterized sequence variations. An example of the former is the detection of specific mutations in the *MASP-2* gene (MIM# 605102), which regulates serum levels of the human mannose binding lectin protein (MASP). Levels of this protein are modulated by three known sequence variations clustered in exon 1 [Turner and Hamvas, 2000]. Individuals with low or absent MASP protein should thus be screened for these sequence variations. However, for most genes, searching for previously characterized sequence variation is not useful, as mutational hotspots are uncommon. For example, X-linked CGD is predominantly caused by mutations in the protein gp91^{phox}. By plotting the location of amino acid substitutions vs. their location in the protein, amino acid changes are scattered along the entire length of the protein with no discernable mutation hotspots (Fig. 3). This is common for many diseases, and makes mutation-specific detection technology unsuitable, as it is not uncommon for individuals to possess unique mutation(s). Laboratories must therefore interrogate the entire gene of interest to identify the location and nature of the disease-causing mutation. Since the average size of human genes is 10–15 kb and the number of exons/gene is nine [Sakharkar et al., 2005], mutation scanning techniques, such as SSCP, HA, dHPLC, DGGE, temperature gradient gel electrophoresis (TGGE), and conformation-sensitive gradient electrophoresis (CSGE) are used to rapidly identify products where a mutation may exist. These are then subsequently sequenced to identify the precise genetic change.

At present, 92% of disease-causing mutations are caused by micro-lesions (missense/nonsense, splicing, regulatory, small deletions and insertions), with the remainder caused by gross lesions (repeat variations, gross insertions/duplications, complex rearrangements and gross deletions) [Stenson et al., 2003] (Supplementary Table S1). All of these sequence variations can be identified using the techniques outlined above.

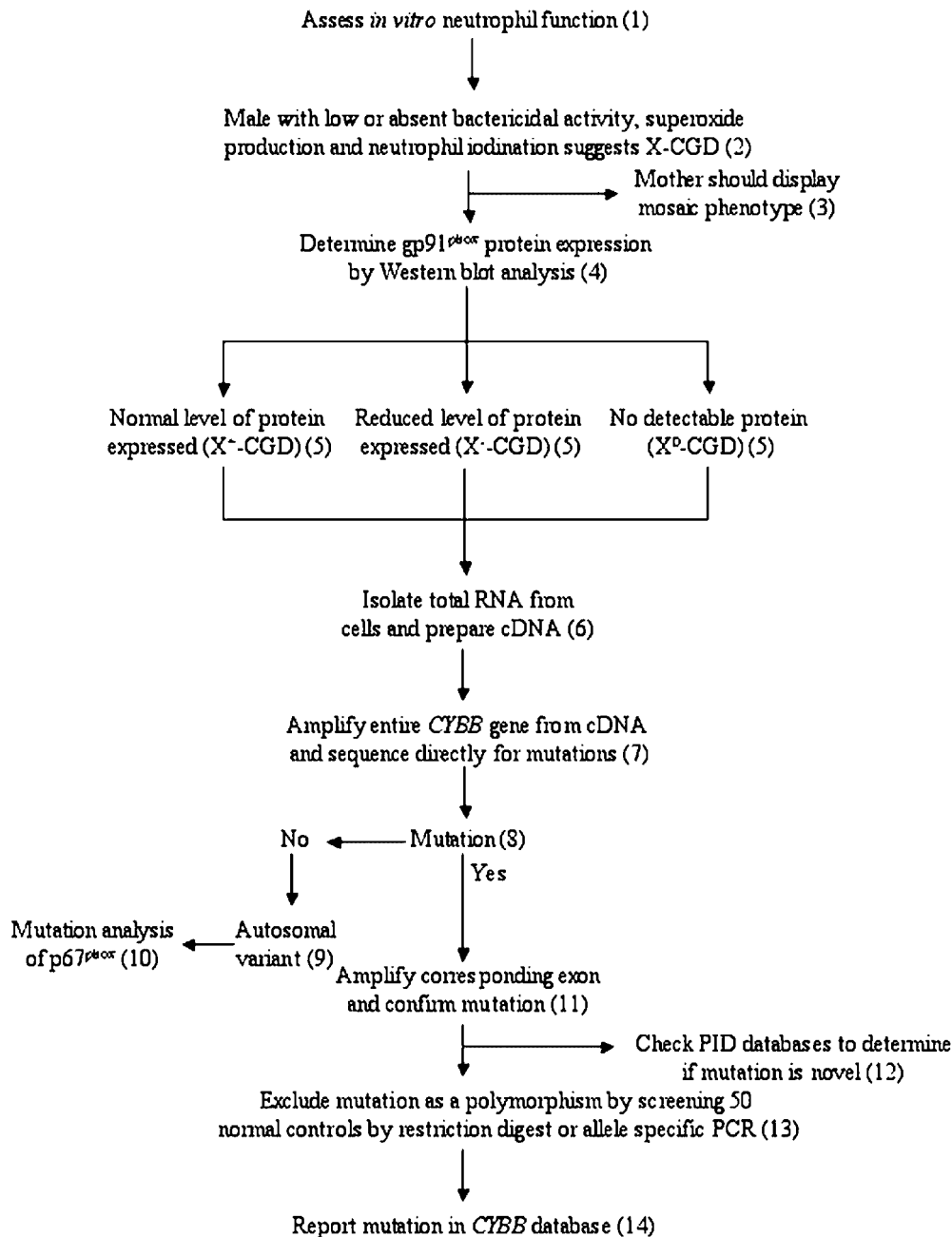


FIGURE 1. Flowchart outlining our current approach in the diagnosis of X-linked chronic granulomatous disease. Initial diagnosis is made by assessing biological cellular function or products (1). An assessment of superoxide production can provide a rapid guide to CGD diagnosis (2). In X-linked diseases, affected females should display a mosaic phenotype that can be seen with the nitroblue tetrazolium test (3). The presence/absence of a protein can then be used to help subclassify the precise variant of CGD (4), of which three variants exist for the X-linked form (5). Total RNA from the affected individual as well as any other family members is isolated and converted into cDNA (6); The *CYBB* gene is then amplified using the cDNA as the template and the products are directly sequenced (7); If a mutation is not found (8), then other variants need to be assessed, such as the autosomal form of CGD (9), and mutation analysis should be performed on the most likely gene involved (10); any mutation should be confirmed at the genomic level, where the corresponding exon should be amplified and sequenced (11); any mutations/polymorphisms should then be screened against other known mutations (12). If the mutation has not been reported previously then it needs to be excluded as a novel polymorphism by either restriction enzyme digestion or by allele specific PCR (13). Novel mutations can then be submitted to the relevant database, such as *CYBB* base for the *CYBB* gene (14).

All screening techniques require the gene of interest be amplified by PCR. Individual exon(s) are amplified either as one product, or if individual exons are very large, in smaller overlapping fragments and then subjected to mutation screening. The theory behind each of these techniques and their advantages and disadvantages in the detection of mutations is discussed below and is summarized in Table 2.

cDNA Analysis

Nearly all PID mutation detection studies commence with the interrogation of genomic DNA using screening techniques. We suggest that direct sequencing of cDNA is an efficient approach for detecting most pathogenic mutations. Since cells from the blood can be easily obtained for analysis, this facilitates

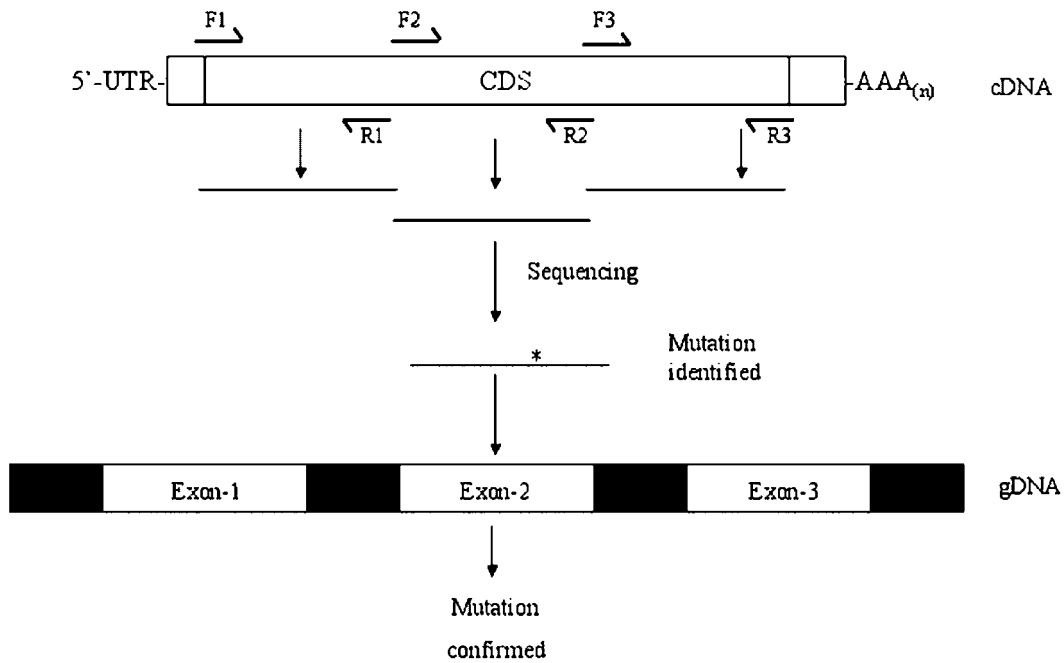


FIGURE 2. Strategy for the detection of mutations in cDNA. The gene of interest is amplified from cDNA using overlapping combinations of primers. These products are then sequenced and any mutations that are identified are then confirmed in the corresponding exon in genomic DNA. *, mutated nucleotide; AAA(n), poly-A tail; cDNA, complementary DNA; CDS, coding sequence; F1–F3, forward 1–3 primer; gDNA, genomic DNA; R1–3, reverse 1–3 primer; UTR, untranslated region.

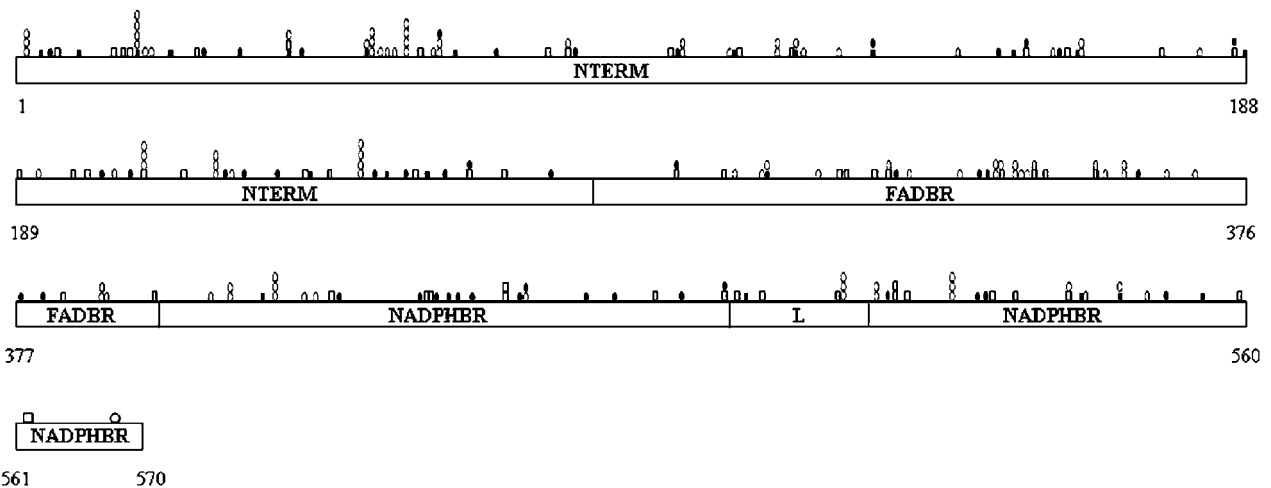


FIGURE 3. Mutation map of the gp91^{phox} protein that is mutant in most cases of X-linked CGD. The location of previously identified amino acid substitutions in the various domains is shown. The results demonstrate the lack of any mutational hotspot, necessitating the complete sequencing of the *CYBB* gene for mutation identification. Insertion (■), deletion (□), missense (○), and nonsense (●) mutations; NTERM, N-terminal; FADBR, flavin adenine dinucleotide (FAD) binding region; NADPHBR, nicotinamide adenine dinucleotide phosphate (reduced form) (NADPH) binding region; L, linker region. Data were obtained from CYBBbase (see Table 1).

the process of RNA isolation. Cost, in terms of time and reagents, is one of the main limiting factors associated with the sequencing of multi-exon genes without a prior screening approach. For example, the *CYBB* gene consists of 12 exons [Royer-Pokora et al., 1986], which require a minimum of 24 primers for their amplification. Each exon must be sequenced in both directions to confirm the sequence. Thus, at least 24 sequencing reactions are needed to screen the entire gene. If the nominal cost/sequencing reaction is US\$10/sample, then the total cost of sequencing alone is US\$240/sample. We now routinely prepare cDNA from patient cells and using overlapping pairs of primers,

amplify and sequence the entire cDNA (Fig. 2). We have been able to amplify the entire *CYBB* cDNA using only five pairs of primers, reducing the cost. Once a mutation is observed, the corresponding exon is then sequenced to confirm its presence at the genomic level (Figs. 1 and 2).

A disadvantage of this approach is that it is dependent upon the presence of suitable mRNA in the patient sample. Some mutations can reduce mRNA production and/or stability and thereby reduce the effectiveness of this approach. However, given the sensitivity of RT-PCR, even small amounts of mRNA can be readily converted into cDNA. Where no mRNA exists, DNA would

TABLE 2. Summary of Major Advantages and Disadvantages Associated with SSCP, HA, DHPLC, CSGE, DGGE, TGGE, and Sequencing

Technique	Advantages	Disadvantages	References
SSCP	Technically simple Inexpensive Uses commonly available reagents and equipment High throughput Does not require use of radioactivity	Low sensitivity Fragment size is a limiting factor Exact nature of nucleotide change is unknown No theoretical basis to predict migration pattern Results can be difficult to interpret High degree of optimization required for every gene to be examined Numerous factors can affect sensitivity Sequencing still required to identify the nature of the mutation Low sensitivity Fragment size is a limiting factor Exact nature of nucleotide change is unknown Results can be difficult to interpret High degree of optimization required for every gene to be examined Numerous factors can affect sensitivity Sequencing still required to identify the nature of the mutation Requires specialized equipment Exact nature of nucleotide change is unknown Difficult to identify heterozygotes unless spiked with a known control More expensive than SSCP Sequencing still required to identify the nature of the mutation	Balogh et al. [2004]; Cotton [1997] Hayashi and Yandell [1993]; Nataraj et al. [1999] Orita et al. [1989]; Sheffield et al. [1993] Nagamine et al. [1989]; Nataraj et al. [1999] Nollau and Wagener [1997] Huber et al. [1993]; O'Donovan et al. [1998] Oefner and Underhill [1995]; Park et al. [2003] Balogh et al. [2004]; Ganguly et al. [1993]; Ganguly [2002] Morelli et al. [2000]
HA	Technically simple Inexpensive Uses commonly available reagents and equipment High throughput Does not require use of radioactivity	F-CSGE requires an automated DNA sequencer	
DHPLC	Can achieve 100% sensitivity and specificity PCR products can be larger than for SSCP Can be high throughput	Fluorescently labeled primers are expensive if F-CSGE is used Specialized data skills are needed DNA mismatch and sequence can affect sensitivity Exact nature of nucleotide change is unknown Exact nature of nucleotide change is unknown Composition of gels can vary Low sensitivity Exact nature of nucleotide change is unknown Can be expensive	
CSGE	Can be automated Does not require the use of radioactivity Software available to predict optimal conditions Technically simple	Requires specialized equipment Only products up to 1,000bp can routinely be sequenced Requires specialized data skills	Ganguly [2002]; Lander et al. [2001]; Sanger et al. [1977] Venter et al. [2001]
TGGE	Inexpensive Higher sensitivity than HA Reproducible	Requires specialized equipment Limited read length Performance of all 4 enzymes crucial for accuracy of process Enzymes can lose activity over time Sample dilution can reduce read length	
DGGE	Does not require denaturing agents Reduced sample handling Simple Can analyze fragments up to 1000bp	Expensive (at present) GC and AT rich regions affect detection Only 1 array for PIDs (ATM)	
Sequencing (DD)	Exact nature of nucleotide change is known Can identify heterozygosity Easily standardized Highly reproducible Can be automated Can be high throughput Exact nature of nucleotide change is known Can be automated Can be high throughput Fast Can be used for SNP analysis Can analyze difficult secondary structures Can sequence entire genes (genomes) rapidly Can screen for SNP in large population studies Future of mutation analysis?		
Pyrosequencing			
Resequencing			

need to be interrogated directly. Also, some pathogenic mutations are found in intronic sequences that would be missed by analyzing cDNA alone. However, these are rare, with the vast majority of pathogenic mutations being found in coding sequences.

Sequencing technology has also improved over the past decade, and 96 capillary machines allow high throughput screening of patient samples. Our approach can enhance this throughput. Given that many PIDs are rare, the workload for a clinical diagnostic laboratory should not be overly excessive to limit the use of this definitive approach.

Single-Stranded Conformation Polymorphism (SSCP)

Historically, SSCP has been the most commonly used screening technique to identify mutations since it was first reported by Orita et al. [1989], including mutations causing PIDs. Examples of PIDs in which SSCP has been used include Bruton's tyrosine kinase (MIM# 300300) [Conley et al., 1998; Vorechovsky et al., 1995], chronic granulomatous disease (MIM# 306400) [Aoshima et al., 1996; Patino et al., 1999], and Wiskott-Aldrich syndrome (MIM# 300392) [Derry et al., 1995; Kondoh et al., 1997].

SSCP involves the PCR amplification of the gene of interest, which is then denatured using heat and formamide to produce single-stranded DNA (ssDNA) [Nataraj et al., 1999]. The fragments are then separated by nondenaturing electrophoresis. During electrophoresis, the ssDNA fold into a nucleotide sequence-dependent conformation. This shape determines the rate of mobility in the gel. Most base changes will alter the physical conformation of the DNA sufficiently that it can be detected as a mobility difference [Humphries et al., 1997].

SSCP is used due to its low cost, use of common reagents and equipment, suitability for high throughput analysis, and reasonable sensitivity. SSCP can also be performed in the absence of radioactive labeling [Cotton, 1997]; however, since only small amounts of DNA are loaded, highly sensitive detection methods are needed. While a radioactive marker can be used, and is more sensitive than either ethidium bromide or silver staining [Sunnucks et al., 2000], fluorescent dyes such as SYBR[®] green (Carlsbad, CA; www.invitrogen.com) are becoming popular alternatives [Law et al., 1996].

A major disadvantage of SSCP is its lack of sensitivity, ranging from 50% [Liu et al., 1996] to 97% when matched to capillary based electrophoresis [Mogensen et al., 2003]. False-negative results can have a drastic effects on a patient's treatment regime [Buckley, 2006].

The size of the PCR products directly affects the sensitivity of the assay. For example, 97% of mutations present in the β -globin, p53, and rhodopsin genes were identified when the PCR products were 155 bp [Sheffield et al., 1993], yet sensitivity fell to 78% when the product was shortened to 135 bp or increased above 155 bp. The choice of gel matrix, electrophoretic conditions, buffer composition, GC content, DNA concentration, location of the sequence variation [Humphries et al., 1997; Liechti-Gallati et al., 1999; Nataraj et al., 1999], concentration of glycerol, and constancy of the temperature during electrophoresis [Balogh et al., 2004; Nataraj et al., 1999] can affect sensitivity. In general, 90% of mutations can be detected in fragments up to 200 bp by varying and optimizing these conditions. The detection rate falls to about 80% for fragments longer than 350 bp and shorter than 200 bp [Hayashi and Yandell, 1993].

Products larger than 1,000 bp can first be digested using a suitable restriction enzyme and then subjected to SSCP analysis. This additional step adds to the cost and complexity of the assay.

In addition, the choice of restriction enzyme(s) used can affect the sensitivity of the assay. For example, digestion with BglIII improved detection of two mutants over digestion of the same fragment with HinfI [Hayashi and Yandell, 1993; Lee et al., 1992]. It should be noted, however, that this effect is mutation specific and not general for these two restriction enzymes.

One reason for the variable sensitivity associated with SSCP is due to the empirical optimization of assay conditions, as it is not possible to predict the electrophoretic patterns expected during SSCP analysis [Balogh et al., 2004]. Poorly standardized conditions can lead to difficult to interpret mobility patterns [Cotton, 1997], giving rise to possible false-negative or -positive results. To achieve optimal sensitivity, genes with exons larger than 200 bp need to be amplified with sets of overlapping primers producing products of a suitable size.

Heteroduplex Analysis (HA)

HA depends on the conformation of duplex DNA resolved in a native polyacrylamide gel. PCR-amplified DNA will consist of either two complementary strands or strands that have a single-base pair mismatch. Mismatched DNA will form heteroduplexes [Nagamine et al., 1989], which have a retarded electrophoretic mobility compared to homoduplexes [Nataraj et al., 1999]. Fragments with an altered mobility are then sequenced.

Nucleotide mismatch is the principle factor influencing separation of heteroduplexes from homoduplexes. For example, G:G/C:C mismatches are more easily identified than A:A/T:T mismatches [Highsmith et al., 1999]. HA is technically simple and cheap to perform [Henderson et al., 1997]. HA has a reported detection rate similar to SSCP [Rossetti et al., 1995], requires significant optimization, and in many cases can be paired with SSCP to increase detection sensitivity, although this leads to an increase in workload per sample.

Factors affecting HA sensitivity include electrophoretic conditions, type of base substitution, base sequence, GC content of the fragment [Nataraj et al., 1999], and fragment size, with the optimal size of 200–600 bp [Morelli et al., 2000; Nollau and Wagener, 1997].

Denaturing High-Performance Liquid Chromatography (dHPLC)

dHPLC involves subjecting PCR products to ion-pair reverse-phase liquid chromatography in a column containing alkylated nonporous particles [Huber et al., 1993; Oefner and Underhill, 1995]. Using a gradient of acetonitrile and partial heat denaturation, heteroduplexes display a reduced column retention time compared to wild-type homoduplexes [Xiao and Oefner, 2001].

dHPLC does not show any significant drop in sensitivity over a size range of 150–700 bp [Liu et al., 1996; O'Donovan et al., 1998] and also displays excellent specificity. For example, 100% of mutations were identified when using dHPLC to analyze the factor IX blood coagulation (MIM# 306900) and neurofibromatosis type 1 genes (MIM# 162200) [O'Donovan et al., 1998]. In contrast only 50% of the mutations were detected using SSCP [Liu et al., 1996]. Similarly, SSCP failed to detect mutations associated with multiple endocrine neoplasia type 1 (MIM# 131100) that were detected by dHPLC [Park et al., 2003]. Software is available that can predict the appropriate conditions (temperature and acetonitrile concentration) for a given nucleotide sequence in order to detect heteroduplexes.

A disadvantage of dHPLC is its requirement for a specialized piece of equipment. While most molecular laboratories have the

equipment required to perform SSCP and HA, dHPLC requires a dedicated piece of equipment. It can also be difficult to identify heterozygotes unless the sample is spiked with a known control. As with the other techniques described, the exact nature and location of the nucleotide change is also unknown.

Conformation-Sensitive, Denaturing, and Temperature-Gradient Gel Electrophoresis (CSGE, DGGE, and TGGE)

CSGE, DGGE, and TGGE involve the electrophoresis of PCR amplified double-stranded DNA with a gradient of denaturant (urea or formamide; DGGE), temperature (TGGE) with a constant concentration of denaturant or without any gradient (CSGE). Heteroduplexes and homoduplexes are then separated according to their different melting behavior, caused by differences in their sequence or by mismatches present in the area of sequence variation. In DGGE, the concentration of denaturing agent, or in TGGE, the temperature of the buffer is increased. The mobility of the DNA is reduced as the DNA begins to denature.

CSGE is a development of HA using a modified polyacrylamide gel [Ganguly et al., 1993] to differentiate homoduplexes from heteroduplexes. The simple protocol, the use of routine laboratory equipment, and the use of PCR products without any further manipulation have been the main incentives for using this technique. The inclusion of fluorescent tags has increased the speed and reproducibility and enhanced the resolving power of the scanning method. The disadvantages include the requirement for access to an automated DNA sequencer, cost of synthesizing fluorescent primers, and need for specialized skills in data analysis and management [Ganguly, 2002]. The size of the mismatch and the sequence context of the mismatch affect the sensitivity of CSGE. CSGE is also less sensitive to mismatches in a GC-rich area than an AT-rich area [Ganguly, 2002].

The main disadvantage of DGGE is that the preparation of denaturing gradient gels is not consistently reproducible, leading to variability [Balogh et al., 2004]. Sensitivity is also an issue, having been reported to be only 82% [Morelli et al., 2000]. Sensitivity of CSGE is reduced when mismatches are located within 50 bp of the ends. TGGE does not require the use of denaturing agents, requires less sample handling, has greater reproducibility, and is cost effective [Balogh et al., 2004].

As with all techniques examining changes in nucleotide composition through variations in mobility, the maximum fragment size for DGGE, TGGE, or CSGE is approximately 1,000 bp [Balogh et al., 2004], while for fluorescent CSGE it is only 250–500 bp [Ganguly, 2002].

Direct Sequencing

The only way to definitively identify if an individual harbors a genetic mutation is to sequence the entire gene. In terms of sequencing, several methods exist; however, only the well established dideoxy chain termination method [Sanger et al., 1977] and pyrosequencing [Ronaghi et al., 1996] are discussed below. In addition, the future prospects of using sequencing by hybridization, multiplex polymerase colony (polony) sequencing, and oligonucleotide arrays are also discussed.

Dideoxy Chain Termination

Sequencing by dideoxy chain termination [Sanger et al., 1977] is based on DNA synthesis occurring in the presence of a mixture of deoxy- and dideoxynucleotides (ddNTPs). When ddNTPs are randomly incorporated into a growing strand of DNA, they cause

amplified products of varying lengths to terminate. These prematurely terminated fragments are analyzed by gel electrophoresis and read manually, or when the ddNTPs are linked to a fluorochrome, they can be automated via capillary electrophoresis (CE).

Dideoxy sequencing has been used for the sequencing of several different bacterial, archaeal, and eukaryotic genomes (www.ncbi.nlm.nih.gov; www.tigr.org), and the human genome [Lander et al., 2001; Venter et al., 2001]. Sequencing is the only true method of identifying uncharacterized nucleotide changes, allowing for the consequence of this change to be determined, such as premature stop codons, missense, or silent mutations. For X-linked diseases, the carrier status of female family members is an important consideration, particularly in providing genetic counseling. With the introduction of improved dye chemistries, the detection of carrier status has been facilitated.

The cost of sequencing is an issue, but has become less so with the improvements in fluorescent technology. Smaller amounts of label can be used for efficient detection of a sample sequence. Sequencing can also be paired with a screening technique. However, this approach is only applicable in laboratories that have the required equipment.

As with all the other methods, the generation of false-positive sequence variations introduced during the PCR amplification of a target is a possibility. However, polymerases with proofreading ability and inherently low error rates are now available. In addition, mutations in PCR products should be confirmed using an alternate technique such as restriction enzyme digestion (if possible) or allele-specific PCR.

Pyrosequencing

Pyrosequencing is a technique that detects the release of pyrophosphate (PPi) during DNA synthesis [Ronaghi et al., 1996]. A series of four enzymatic reactions are linked to the release of visible light, whose generation is proportional to the number of incorporated nucleotides. The cascade starts with a nucleic acid polymerization reaction in which inorganic PPi is released as a result of nucleotide incorporation by DNA polymerase. The released PPi is then converted to ATP, which ultimately leads to the release of light. Since nucleotides are added in a stepwise fashion and their order is known, the sequence of the template can be determined.

This technology has recently been used to successfully sequence a bacterial genome in only 4 hr [Margulies et al., 2005] as well as detecting SNPs in the mannose-binding lectin [Roos et al., 2006]. It can also be used for microbial typing, resequencing, tag sequencing, and analysis of DNA sequences with difficult secondary structures [Ronaghi, 2001].

A major disadvantage of pyrosequencing is the limited read length [Ahmadian et al., 2006] typically used for mutation detection [Ahmadian et al., 2000; Garcia et al., 2000]. In addition, the performance of all four enzymes is crucial for the accuracy of this DNA sequencing technology. Enzyme activity is lost over time, while sample dilution, as each nucleotide is subsequently added, can also reduce enzyme activity and have a negative impact on read length [Ronaghi, 2001].

Combinatorial Sequencing-by-Hybridization

Resequencing is the sequencing of DNA whose sequence has previously been annotated. Resequencing is achieved through standard sequencing technologies. Recently, new technologies have been applied to the large-scale resequencing of genes and

genomes. One approach is combinatorial sequencing-by-hybridization (cSBH) [Cowie et al., 2004; Drmanac et al., 1993].

In cSBH, a universal probe is attached to a glass slide while a fluorophore-labeled probe is free in solution. Unlabeled PCR-amplified target DNA is mixed with DNA ligase and the labeled probe in solution, and allowed to hybridize to the slide-attached probe. When both probes hybridize to the target DNA at contiguous complementary positions, they are covalently joined by the DNA ligase, creating one long-labeled probe attached to the solid surface. The combinatorial process generates all possible probes that are complementary to the target. A standard array reader scores the fluorescent signal at each array position, and software generates a sequence readout of the template PCR product. This approach is available through Callida Genomics (Sunnyvale, CA; www.callidagenomics.com).

This approach has been used to screen for mutations in the APC gene (MIM# 175100), the cause of the autosomal dominant form of colorectal cancer [Cowie et al., 2004]. A total of 23 clinically diagnosed individuals with familial adenomatous polyposis were screened, and with a maximal read length of 3.6 kb, bases were sequenced with an accuracy of 99.97%. This approach has also been used to interrogate the neurofibromatosis type 1 (*NF1*) gene (MIM# 162200) [Schirinzi et al., 2006]. A total of 30 patients were examined, and readability and accuracy were found to be 99.97% and 99.99%, respectively. A 10-fold reduction in uncalled bases and reduced background was achieved when dual labeling was used. Base substitutions, insertions, and small deletions were confirmed in all patients.

Oligonucleotide Resequencing Arrays

Oligonucleotide resequencing arrays rely on the hybridization of fragmented end-labeled DNA to sets of tiled, overlapping oligonucleotide 25mer probes. For each base being interrogated there are four probes, one representing the reference sequence and three that vary at the central base with one of the other three nucleotides [Mockler et al., 2005]. Probes corresponding to both the sense and antisense strand are synthesized on the array, and signal intensities from the probes are analyzed to give homozygous or heterozygous base calls [Chee et al., 1996; Warrington et al., 2002]. A mismatch at the middle base will reduce or prevent binding of the labeled target and thus result in a reduced hybridization signal. At present, custom re-sequencing arrays are available from Affymetrix (Santa Clara, CA; www.affymetrix.com).

Using this system, Hacia et al. [1998] screened for all possible mutations in the ataxia-telangiectasia (MIM# 208900) gene. Using this approach, 94.4% distinct heterozygous and 100% distinct homozygous sequence variants were detected accurately, while five false-positive calls were identified. Karaman et al. [2005] extended this study using oligonucleotide arrays to determine the hybridization properties of probes complementary to small insertions and deletions. Probes complementary to two base substitutions displayed the highest average specificity followed by those complementary to single-base substitutions, single-base deletions, and single-base insertions [Karaman et al., 2005].

Resequencing arrays have also been applied to detect sequence variations in multiple autosomal recessive retinal disease [Mandal et al., 2005] and cancer-associated genes [Tengs et al., 2005]. Mandal et al. [2005] sequenced 11 genes (155 exons) associated with autosomal recessive retinitis pigmentosa (MIM# 600105). Over 97% of the sequence was determined with greater than 99% accuracy and reproducibility. Similarly, Tengs et al. [2005] found

92% of the homozygous changes and 98% of heterozygous loci were identified correctly.

Wong et al. [2004] used a resequencing array to interrogate the 30-kb severe acute respiratory syndrome (SARS) virus genome. The method was highly reproducible, accurate (>99.99%), and capable of identifying known and novel variants of SARS. Similarly, Zwick et al. [2005] resequenced multiple isolates of *Bacillus anthracis* with 92.6% accuracy.

A major advantage of array-based resequencing is that no prior knowledge of the variation in a site is needed. Comparisons can be made directly to a reference genomic sequence. Thus, a single analysis allows for SNPs to be identified [Mockler et al., 2005]. One disadvantage of resequencing technology relates to the composition of the probes. Probes with a high AT content [Hacia et al., 1998] or GC content [Mandal et al., 2005; Zwick et al., 2005] can affect hybridization intensities, preventing bases from being efficiently detected. This problem exists even in organisms with a low G+C content such as *B. anthracis* (~34%) [Zwick et al., 2005]. This technology is also dependent on the availability of a finished quality reference sequence. Also, since resequencing arrays only compare a new sequence to a reference sequence, they are not as useful for identifying novel sequences or sequence rearrangements that may occur in an experimental sample.

SNP Detection

As mentioned above, some diseases are associated with a set of well-characterized mutations, such as the MASP protein [Turner and Hamvas, 2000]. In these cases, assays based on high-throughput SNP detection can be used. Examples of two such techniques include single-nucleotide extension (SNE) and oligonucleotide ligation assay (OLA)/SNPlex.

In SNE, primers are designed to anneal immediately 3' to the polymorphic site of interest. A DNA polymerase is then used to extend the primer with a single labeled nucleotide complementary to the nucleotide at the variant site. The labeled product can then be detected by CE [Syvänen 1999]. This approach allows for highly-specific detection of SNPs and has been used in numerous reports [Murphy et al., 2003].

SNPlex is a platform modification of OLA produced by Applied Biosystems (Foster City, CA; www.appliedbiosystems.com) in which oligonucleotide ligation/PCR and CE is used to analyze biallelic SNP genotypes. This approach is well suited for SNP genotyping in which throughput and cost efficiency are essential. During SNPlex, allele-specific oligonucleotide (ASO) probes and locus specific oligonucleotide (LSO) probes hybridize to the target sequence. These two oligonucleotides only bind when there is a perfect matching sequence at the SNP site. At the same time, universal linkers are ligated to the distal termini of the probes. The linkers have both universal primer binding sequences as well as sequences complementary to the ASO and LSO probes. A unique code is placed at the 5' end of each ASO facilitating the identification of the target sequence. All probes are designed to function under the same hybridization conditions. Unligated probes, linkers, and genomic DNA are removed by enzymatic digestion and then the purified ligation products are amplified using one set of universal primers, one of which is biotinylated. Biotinylated amplicons are then captured using streptavidin. Fluorescently-labeled universal probes then bind to unique 5' termini, which are then separated by CE. Tobler et al. [2005] used the SNPlex genotyping system to test 48 SNPs in 44 genomic control DNA samples, as well as 521 SNPs in 92 individuals. This approach demonstrated an overall call rate, precision, and

concordance with TaqMan assays of 99.32, 99.95, and 99.84%, respectively.

Future Sequencing Technologies

Several novel high-throughput multiplex sequencing technologies have been developed with the aim of sequencing the genome for \$1,000. Sequencing-by-synthesis (SBS) pioneered by Solexa (Hayward, CA; www.solexa.com) uses four fluorescently-labeled modified nucleotides to sequence millions of fragments of DNA or cDNA in parallel, one base at a time. The modified nucleotides [Bai et al., 2003; Li et al., 2003] possess a reversible termination property [Ruparel et al., 2005] that allows each sequencing cycle to occur simultaneously in the presence of all four nucleotides.

Another high-throughput sequencing technology is polony sequencing, which can amplify 1,000–10,000 template molecules in parallel and independently [Shendure et al., 2005; Zhang et al., 2006]. Although still a technique requiring further optimization, Shendure et al. [2005] were able to resequence 30.1 million bases of an *E. coli* strain at a median raw accuracy of 99.7% in just 2.5 days. Using the same approach Zhang et al. [2006] performed long-range multilocus haplotyping spanning of five SNPs spanning 158 Mb of chromosome 7.

DISCUSSION

The diagnosis of PIDs is complemented through a combination of cellular-, immunochemical-, and molecular-based approaches. While cellular-based assays enable an assessment of the phenotype manifested, a molecular approach provides a definitive result to the origin of the immunodeficiency.

Fortunately, recent mutation scanning techniques have facilitated the molecular diagnosis of PIDs. But nevertheless, we need to be aware that all approaches discussed, apart from dHPLC, suffer from a lack of sensitivity that can lead to the generation of false negative results. The only limitation with dHPLC is its dependency upon the use of DNA sequencing to identify the exact genetic change. DNA sequencing is still the only technique that can provide a definitive answer to variations in gene sequences which may or may not result in genetic disease. However, sensitivity is a limitation for any mutation detection method and is one of the key advantages in using direct sequencing.

The amplification of all of the individual exons of a gene followed by DNA sequencing makes this approach not cost effective. As a result alternative strategies to reduce the need for entire sequencing of genes have been developed and discussed above. Our approach is aimed at reducing the total number of samples that need to be sequenced by analyzing cDNA fragments and confirming any changes in individual genomic exons.

As technology improves and the cost associated with sequencing reduces, it will be feasible for laboratories to sequence all the genes associated with PIDs in one event, thereby identifying genes of potential interest.

ACKNOWLEDGMENTS

We thank Professor Grant Sutherland for critical reading of the manuscript.

REFERENCES

Ahmadian A, Lundeberg J, Nyren P, Uhlen M, Ronaghi M. 2000. Analysis of the p53 tumor suppressor gene by pyrosequencing. *Biotechniques* 28:140–144; 146–147.

- Ahmadian A, Ehn M, Hober S. 2006. Pyrosequencing: history, biochemistry and future. *Clin Chim Acta* 363:83–94.
- Amedei A, Romagnani C, Benagiano M, Azzurri A, Fomia F, Torrente F, Plebani A, D'Elisio MM, Del Prete G. 2001. Preferential Th1 profile of T helper cell responses in X-linked (Bruton's) agammaglobulinemia. *Eur J Immunol* 31:1927–1934.
- Aoshima M, Nunoi H, Shimazu M, Shimizu S, Tatsuzawa O, Kenney RT, Kanegasaki S. 1996. Two-exon skipping due to a point mutation in p67-phox-deficient chronic granulomatous disease. *Blood* 88:1841–1845.
- Bai X, Li Z, Jockusch S, Turro NJ, Ju J. 2003. Photocleavage of a 2-nitrobenzyl linker bridging a fluorophore to the 5' end of DNA. *Proc Natl Acad Sci USA* 100:409–413.
- Balogh K, Patocs A, Majnik J, Racz K, Hunyady L. 2004. Genetic screening methods for the detection of mutations responsible for multiple endocrine neoplasia type 1. *Mol Genet Metab* 83:74–81.
- Bonilla FA, Geha RS. 2006. 2. Update on primary immunodeficiency diseases. *J Allergy Clin Immunol* 117:S435–S441.
- Buckley RH. 2006. Primary immunodeficiency or not? Making the correct diagnosis. *J Allergy Clin Immunol* 117:756–758.
- Chee M, Yang R, Hubbell E, Berno A, Huang XC, Stern D, Winkler J, Lockhart DJ, Morris MS, Fodor SP. 1996. Accessing genetic information with high-density DNA arrays. *Science* 274:610–614.
- Conley ME, Mathias D, Treadaway J, Minegishi Y, Rohrer J. 1998. Mutations in Btk in patients with presumed X-linked agammaglobulinemia. *Am J Hum Genet* 62:1034–1043.
- Conley ME, Notarangelo LD, Etzioni A. 1999. Diagnostic criteria for primary immunodeficiencies. Representing PAGID (Pan-American Group for Immunodeficiency) and ESID (European Society for Immunodeficiencies). *Clin Immunol* 93:190–197.
- Cotton RGH. 1997. Mutation detection. New York: Oxford University Press. p 32–127.
- Cowie S, Drmanac S, Swanson D, Delgrosso K, Huang S, du Sart D, Drmanac R, Surrey S, Fortina P. 2004. Identification of APC gene mutations in colorectal cancer using universal microarray-based combinatorial sequencing-by-hybridization. *Hum Mutat* 24:261–271.
- Cross AR, Heyworth PG, Rae J, Curnutte JT. 1995. A variant X-linked chronic granulomatous disease patient (X91+) with partially functional cytochrome b. *J Biol Chem* 270:8194–8200.
- Derry JM, Kerns JA, Weinberg KI, Ochs HD, Volpini V, Estivill X, Walker AP, Francke U. 1995. WASP gene mutations in Wiskott-Aldrich syndrome and X-linked thrombocytopenia. *Hum Mol Genet* 4:1127–1135.
- Drmanac R, Strezoska S, Paunesku Z, Labat T, Zeremski I, Snoddy M, Funkhouser J, Koop WK, Hood BL. 1993. DNA sequence determination by hybridization: a strategy for efficient large-scale sequencing. *Science* 260:1649–1652.
- Ferrante A, Rowan-Kelly B, Beard LJ, Maxwell GM. 1986. An enzyme-linked immunosorbent assay for the quantitation of human IgG subclasses using monoclonal antibodies. *J Immunol Methods* 93:207–212.
- Ganguly A, Rock MJ, Prockop DJ. 1993. Conformation-sensitive gel electrophoresis for rapid detection of single-base differences in double-stranded PCR products and DNA fragments: evidence for solvent-induced bends in DNA heteroduplexes. *Proc Natl Acad Sci USA* 90:10325–10329.
- Ganguly A. 2002. An update on conformation sensitive gel electrophoresis. *Hum Mutat* 19:334–342.
- Garcia CA, Ahmadian A, Gharizadeh B, Lundeberg J, Ronaghi M, Nyren P. 2000. Mutation detection by pyrosequencing: sequencing of exons 5–8 of the p53 tumor suppressor gene. *Gene* 253:249–257.
- Hacia JG, Sun B, Hunt N, Edgemon K, Mosbrook D, Robbins C, Fodor SP, Tagle DA, Collins FS. 1998. Strategies for mutational analysis of the large multiexon ATM gene using high-density oligonucleotide arrays. *Genome Res* 8:1245–1258.
- Hayashi K, Yandell DW. 1993. How sensitive is PCR-SSCP? *Hum Mutat* 2:338–346.
- Henderson BG, Wenham PR, Ashby JP, Blundell G. 1997. Detecting familial defective apolipoprotein B-100: three molecular scanning methods compared. *Clin Chem* 43:1630–1634.

- Highsmith WE Jr, Jin Q, Nataraj AJ, O'Connor JM, Burland VD, Baubonis WR, Curtis FP, Kusakawa N, Garner MM. 1999. Use of a DNA toolbox for the characterization of mutation scanning methods. I: Construction of the toolbox and evaluation of heteroduplex analysis. *Electrophoresis* 20:1186–1194.
- Huber CG, Oefner PJ, Preuss E, Bonn GK. 1993. High-resolution liquid chromatography of DNA fragments on non-porous poly(styrene-divinylbenzene) particles. *Nucleic Acids Res* 21:1061–1066.
- Humphries SE, Gudnason V, Whittall R, Day IN. 1997. Single-strand conformation polymorphism analysis with high throughput modifications, and its use in mutation detection in familial hypercholesterolemia. International Federation of Clinical Chemistry Scientific Division: Committee on Molecular Biology Techniques. *Clin Chem* 43:427–435.
- Illoh OC. 2004. Current applications of flow cytometry in the diagnosis of primary immunodeficiency diseases. *Arch Pathol Lab Med* 128:23–31.
- Jirapongsananurak O, Malech HL, Kuhns DB, Niemela JE, Brown MR, Anderson-Cohen M, Fleisher TA. 2003. Diagnostic paradigm for evaluation of male patients with chronic granulomatous disease, based on the dihydrorhodamine 123 assay. *J Allergy Clin Immunol* 111:374–379.
- Karaman MW, Groshen S, Lee CC, Pike BL, Hacia JG. 2005. Comparisons of substitution, insertion and deletion probes for resequencing and mutational analysis using oligonucleotide microarrays. *Nucleic Acids Res* 33:e33.
- Kondoh T, Matsumoto T, Tsuji Y. 1997. Wiskott-Aldrich syndrome in two sisters. *Am J Med Genet* 74:218–219.
- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, Fitzhugh W. et al. 2001. Initial sequencing and analysis of the human genome. *Nature* 409:860–921.
- Law JC, Facher EA, Deka A. 1996. Nonradioactive single-strand conformation polymorphism analysis with application for mutation detection in a mixed population of cells. *Anal Biochem* 236:373–375.
- Lee HH, Lo WJ, Choo KB. 1992. Mutational analysis by a combined application of the multiple restriction fragment-single strand conformation polymorphism and the direct linear amplification DNA sequencing protocols. *Anal Biochem* 205:289–293.
- Liechti-Gallati S, Schneider V, Neeser D, Kraemer R. 1999. Two buffer PAGE system-based SSCP/HD analysis: a general protocol for rapid and sensitive mutation screening in cystic fibrosis and any other human genetic disease. *Eur J Hum Genet* 7:590–598.
- Li Z, Bai X, Ruparel H, Kim S, Turro NJ, Ju J. 2003. A photocleavable fluorescent nucleotide for DNA sequencing and analysis. *Proc Natl Acad Sci USA* 100:414–419.
- Liu Q, Feng J, Sommer SS. 1996. Bi-directional dideoxy fingerprinting (Bi-ddF): a rapid method for quantitative detection of mutations in genomic regions of 300–600 bp. *Hum Mol Genet* 5:107–114.
- Mandal MN, Heckenlively JR, Burch T, Chen L, Vasireddy V, Koenekoop RK, Sieving PA, Ayyagari R. 2005. Sequencing arrays for screening multiple genes associated with early-onset human retinal degenerations on a high-throughput platform. *Invest Ophthalmol Vis Sci* 46:3355–3362.
- Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437:376–380.
- Mockler TC, Chan S, Sundaresan A, Chen H, Jacobsen SE, Ecker JR. 2005. Applications of DNA tiling arrays for whole-genome analysis. *Genomics* 85:1–15.
- Mogensen J, Bahl A, Kubo T, Elanko N, Taylor R, McKenna WJ. 2003. Comparison of fluorescent SSCP and denaturing HPLC analysis with direct sequencing for mutation screening in hypertrophic cardiomyopathy. *J Med Genet* 40:e59.
- Morelli A, Falchetti A, Martinetti V, Becherini L, Mark M, Friedman E, Brandi ML. 2000. MEN1 gene mutation analysis in Italian patients with multiple endocrine neoplasia type 1. *Eur J Endocrinol* 142:131–137.
- Murphy KM, Geiger T, Hafez MJ, Eshleman JR, Griffin CA, Berg KD. 2003. A single nucleotide primer extension assay to detect the APC I1307K gene variant. *J Mol Diagn* 5:222–226.
- Nagamine CM, Chan K, Lau YF. 1989. A PCR artifact: generation of heteroduplexes. *Am J Hum Genet* 45:337–339.
- Nataraj AJ, Olivos-Glander I, Kusakawa N, Highsmith WE, Jr. 1999. Single-strand conformation polymorphism and heteroduplex analysis for gel-based mutation detection. *Electrophoresis* 20:1177–1185.
- Nollau P, Wagener C. 1997. Methods for detection of point mutations: performance and quality assessment. IFCC Scientific Division, Committee on Molecular Biology Techniques. *Clin Chem* 43:1114–1128.
- Notarangelo L, Casanova JL, Conley ME, Chapel H, Fischer A, Puck J, Roifman C, Seger R, Geha RS, International Union of Immunological Societies Primary Immunodeficiency Diseases Classification Committee. 2006. Primary immunodeficiency diseases: an update from the International Union of Immunological Societies Primary Immunodeficiency Diseases Classification Committee Meeting in Budapest, 2005. *J Allergy Clin Immunol* 117:883–896.
- O'Donovan MC, Oefner PJ, Roberts SC, Austin J, Hoogendoorn B, Guy C, Speight G, Upadhyaya M, Sommer SS, McGuffin P. 1998. Blind analysis of denaturing high-performance liquid chromatography as a tool for mutation detection. *Genomics* 52:44–49.
- Oefner PJ, Underhill PA. 1995. Comparative DNA sequencing by denaturing high-performance liquid chromatography (dHPLC). *Am J Hum Genet* 57:A266.
- Orita M, Iwahana H, Kanazawa H, Hayashi K, Sekiya T. 1989. Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc Natl Acad Sci USA* 86:2766–2770.
- Park JH, Kim IJ, Kang HC, Lee SH, Shin Y, Kim KH, Lim SB, Kang SB, Lee K, Kim SY, Lee MS, Lee MK, Moon SD, Park JG. 2003. Germline mutations of the MEN1 gene in Korean families with multiple endocrine neoplasia type 1 (MEN1) or MEN1-related disorders. *Clin Genet* 64:48–53.
- Patino PJ, Perez JE, Lopez JA, Condino-Neto A, Grumach AS, Botero JH, Curnutte JT, Garcia de Olarte D. 1999. Molecular analysis of chronic granulomatous disease caused by defects in gp91-phox. *Hum Mutat* 13:29–37.
- Piirilä H, Väliäho J, Vihinen M. 2006. Immunodeficiency mutation databases (IDbases). *Hum Mutat* 27:1200–1208.
- Ronaghi M. 2001. Pyrosequencing sheds light on DNA sequencing. *Genome Res* 11:3–11.
- Ronaghi M, Karamohamed S, Pettersson B, Uhlen M, Nyren P. 1996. Real-time DNA sequencing using detection of pyrophosphate release. *Anal Biochem* 242:84–89.
- Roos D, de Boer M, Kuribayashi F, Meischl C, Weening RS, Segal AW, Ahlin A, Nemet K, Hossle JP, Bernatowska-Matuszkiewicz E, Middleton-Price H. 1996. Mutations in the X-linked and autosomal recessive forms of chronic granulomatous disease. *Blood*. 87:1663–1681.
- Roos A, Dieltjes P, Vossen RH, Daha MR, de Knijff P. 2006. Detection of three single nucleotide polymorphisms in the gene encoding mannose-binding lectin in a single pyrosequencing reaction. *J Immunol Methods* 309:108–114.
- Rossetti S, Corra S, Biasi MO, Turco AE, Pignatti PF. 1995. Comparison of heteroduplex and single-strand conformation analyses, followed by ethidium fluorescence visualization, for the detection of mutations in four human genes. *Mol Cell Probes* 9:195–200.
- Royer-Pokora B, Kunkel LM, Monaco AP, Goff SC, Newburger PE, Baehner RL, Cole FS, Curnutte JT, Orkin SH. 1986. Cloning the gene for an inherited human disorder-chronic granulomatous disease on the basis of its chromosomal location. *Nature* 322:32–38.
- Ruparel H, Bi L, Li Z, Bai X, Kim DH, Turro NJ, Ju J. 2005. Design and synthesis of a 3'-O-allyl photocleavable fluorescent nucleotide as a reversible terminator for DNA sequencing by synthesis. *Proc Natl Acad Sci USA* 102:5932–5937.

- Sakharkar MK, Perumal BS, Sakharkar KR, Kanguene P. 2005. An analysis on gene architecture in human and mouse genomes. *In Silico Biol* 5:347–365.
- Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74:5463–5467.
- Schirinzì A, Drmanac S, Dallapiccola B, Huang S, Scott K, de Luca A, Swanson D, Drmanac R, Surrey S, Fortina P. 2006. Combinatorial sequencing-by-hybridization: analysis of the NF1 gene. *Genetic Testing* 10:8–17.
- Sheffield VC, Beck JS, Kwitek AE, Sandstrom DW, Stone EM. 1993. The sensitivity of single-strand conformation polymorphism analysis for the detection of single base substitutions. *Genomics* 16:325–332.
- Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AM, Wang MD, Zhang K, Mitra RD, Church GM. 2005. Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* 309:1728–1732.
- Stenson PD, Ball EV, Mort M, Phillips AD, Shiel JA, Thomas NS, Abeyasinghe S, Krawczak M, Cooper DN. 2003. Human Gene Mutation Database (HGMD): 2003 update. *Hum Mutat* 21:577–581.
- Sunnucks P, Wilson AC, Beheregaray LB, Zenger K, French J, Taylor AC. 2000. SSCP is not so difficult: the application and utility of single-stranded conformation polymorphism in evolutionary biology and molecular ecology. *Mol Ecol* 9:1699–1710.
- Syvänen AC. 1999. From gels to chips: “Minisequencing” primer extension for analysis of point mutations and single nucleotide polymorphisms. *Hum Mutat* 13:1–10.
- Tengs T, Lee JC, Guillermo Paez J, Zhao X, Laframboise T, Giannoukos G, Thomas RK. 2006. A transforming MET mutation discovered in non-small cell lung cancer using microarray-based resequencing. *Cancer Lett* 239:227–233.
- Tobler AR, Short S, Andersen MR, Paner TM, Briggs JC, Lambert SM, Wu PP, Wang Y, Spoonde AY, Koehler RT, Peyret N, Chen C, Broome AJ, Ridzon DA, Zhou H, Hoo BS, Hayashibara KC, Leong LN, Ma CN, Rosenblum BB, Day JP, Ziegler JS, De La Vega FM, Rhodes MD, Hennessy KM, Wenz HM. 2005. The SNPlex genotyping system: a flexible and scalable platform for SNP genotyping. *J Biomol Tech* 16:398–406.
- Turner MW, Hamvas RM. 2000. Mannose-binding lectin: structure, function, genetics and disease associations. *Rev Immunogenet* 2: 305–322.
- Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA. et al. 2001. The sequence of the human genome. *Science* 291:1304–1351.
- Vorechovsky I, Vihinen M, de Saint Basile G, Honsova S, Hammarstrom L, Muller S, Nilsson L, Fischer A, Smith CI. 1995. DNA-based mutation analysis of Bruton's tyrosine kinase gene in patients with X-linked agammaglobulinaemia. *Hum Mol Genet* 4:51–58.
- Warrington JA, Shah NA, Chen X, Janis M, Liu C, Kondapalli S, Reyes V, Savage MP, Zhang Z, Watts R, DeGuzman M, Berno A, Snyder J, Baid J. 2002. New developments in high-throughput resequencing and variation detection using high density microarrays. *Hum Mutat* 19: 402–409.
- Wilson DI, Burn J, Scambler P, Goodship J. 1993. DiGeorge syndrome: part of CATCH 22. *J Med Genet* 30:852–856.
- Wolach B, Ash S, Gavrieli R, Stark B, Yaniv I, Roos D. 2005. Acute lymphoblastic leukemia in a patient with chronic granulomatous disease and a novel mutation in CYBB: first report. *Am J Hematol*. 80: 50–54.
- Wong CW, Albert TJ, Vega VB, Norton JE, Cutler DJ, Richmond TA, Stanton LW, Liu ET, Miller LD. 2004. Tracking the evolution of the SARS coronavirus using high-throughput, high-density resequencing arrays. *Genome Res* 14:398–405.
- Xiao W, Oefner PJ. 2001. Denaturing high-performance liquid chromatography: a review. *Hum Mutat* 17:439–474.
- Zhang K, Zhu J, Shendure J, Porreca GJ, Aach JD, Mitra RD, Church GM. 2006. Long-range polony haplotyping of individual human chromosome molecules. *Nat Genet* 38:382–387.
- Zicha D, Dunn GA, Segal AW. 1997. Deficiency of p67phox, p47phox or gp91phox in chronic granulomatous disease does not impair leucocyte chemotaxis or motility. *Br J Haematol* 96:543–550.
- Zwick ME, McAfee F, Cutler DJ, Read TD, Ravel J, Bowman GR, Galloway DR, Mateczun A. 2005. Microarray-based resequencing of multiple *Bacillus anthracis* isolates. *Genome Biol* 6:R10.