

Biochips for Detection of DNA Mutations

Li-Jun Bi, Xian-En Zhang*, Ya-Feng Zhou and Jiao-Yu Deng

Joint Research Group for Analytical Pathogenic Microbiology, State Key Laboratory of Biomacromolecules and State Key Laboratory of Virology, Institute of Biophysics and Wuhan Institute of Virology, Chinese Academy of Sciences, Beijing 100101, China

Abstract: DNA mutation is the major cause of gene evolution. The variation of DNA sequence may result in genetic diseases, genotypes, single nucleotide polymorphism, and drug resistance of pathogens. Using biochip technologies, detection of mutations can be performed in parallel under identical conditions, hereby providing a solution to continuously increasing demands for large-scale genetic analysis. This article reviews the development of the biochips for mutation analysis. Features, limitations and applications of various principles are discussed.

Keywords: DNA mutations, Genetic diseases, Drug resistance, Biochip, Detection

INTRODUCTION

Gene mutations, including single-base substitutions and small insertions or deletions of bases are the major causes of genetic diseases, genotypes and single nucleotide polymorphism (SNP), as well as microbial drug resistance [1-4]. Detection of gene mutations helps to understand all these phenomena at molecular bases and thus, is demanded for clinical diagnosis and biological studies.

Conventional detection technique is DNA sequencing [5], which is accurate and widely accepted as the first choice among all the existing methods. However, the sequencing method is still technical, relatively expensive and time-consuming and thus hardly fulfills the clinical applications. There are many other approaches, among which the frequently cited ones include denaturing gradient gel electrophoresis (DGGE) [6], PCR-single strand conformation polymorphism (PCR-SSCP) [7], chemical or enzymatic cleavage of mismatch [8], allele-specific PCR [9], etc. All these approaches produce results with high fidelity. They are, however, still quite labor intensive and typically based on the use of polyacrylamide gel electrophoresis (PAGE), which significantly limits their suitability for automated applications. Fluorescence resonance energy transfer (FRET) detection is a newly developed method that avoids using gel electrophoresis. Two forms of this method, TaqMan genotyping and molecular beacons, have been already applied to detect mutations in real-time, but they are usually very expensive [10-12].

In recent years, more attention has been paid to the biochip technologies, which develop very fast and show their advantages in parallel and high-throughput analysis of DNA mutations. A biochip, sometimes called microarray, is actually a collection of miniaturized test sites arranged on a solid

substrate, which permit many tests to be performed simultaneously under the identical conditions in order to achieve higher throughput and speed [13-15]. The biochips have many applications [16]. Those used for mutation detection can be sorted as four categories: hybridization-based, PCR-based, oligonucleotide ligation assay (OLA)- and MutS-based biochips. This article will describe the principles and applications of these methods with discussion of their advantages and limitations.

HYBRIDIZATION-BASED BIOCHIPS FOR DETECTION OF DNA MUTATIONS

Sequencing by Hybridization and Detecting Mutation by Hybridization

DNA Hybridization is a process involving two single strands of DNA to form complementary double strands under appropriate condition. Based on this principle, the British biologist Edward M Southern has made two famous contributions. One is Southern blotting, the method that was named after its inventor and has already been widely used as an experimental protocol in the detection of existence of specific gene in the sample [17]. The other is called sequencing by hybridization (SBH), which relies on the fact that the hybridization strictly follows the rule of Watson-Crick pair [18]. Southern succeeded in binding a large number of different probes to a solid carrier surface such as glass. In addition, he developed methods with which these probes could be chemically produced directly on the surface, thereby laying the scientific and methodological groundwork for the production of DNA chips (<http://www.roche.de/presse/4-Background.pdf>). In the SBH experiment [19], an oligo array contains all possible sequences of length n , normally $n = 8\sim 10$ bases. This format is called universal oligonucleotide array. If there is any change in the sequence of the target DNA, failure hybridization may occur. Since the target DNA fragments are labeled with fluorescent dyes, the failure hybridization could be read out by either signing or signal silence. In other words, the base mutation of a target sequence could be detected out by a known probe, which forms the basic idea of mutation detection by hybridization.

*Address correspondence to this author at the Joint Research Group for Analytical Pathogenic Microbiology, State Key Laboratory of Biomacromolecules and State Key Laboratory of Virology, Institute of Biophysics and Wuhan Institute of Virology, Chinese Academy of Sciences, Beijing 100101, China; Tel: +86-(0)10-58881508; Fax: +86-(0)10-64888464; E-mail: zhangxe@most.cn

One of the early studies using oligonucleotide array for detection of known DNA sequence variations was reported by Saiki *et al.* in 1989 [20]. The analysis of DNA for the presence of particular mutations or polymorphisms could be readily accomplished by differential hybridization with sequence-specific oligonucleotide probes [21]. PCR has facilitated the use of these probes by greatly increasing the number of copies of target DNA in the sample prior to hybridization. In that format, the oligonucleotides were given homopolymer tails with terminal deoxyribonucleotidyltransferase, spotted onto a nylon membrane, and covalently bound by UV irradiation. Due to their long length, the tails were preferentially bound to the nylon, leaving the oligonucleotide probe free to hybridize. The target segment of the DNA sample to be tested was PCR-amplified with biotinylated primers and then hybridized to the membrane containing the immobilized oligonucleotides under stringent conditions. This technique had been applied to HLA-DQA genotyping and to the detection of Mediterranean beta-thalassemia mutations [20].

Oligonucleotide Microarrays and DNA Chips

In the late 1980s, a team of scientists led by Stephen P. A. Fodor enabled breakthroughs in understanding the human condition when they invented microarray technology. The theory behind their work was revolutionary, a notion that semiconductor manufacturing techniques could be united with advances in combinatorial chemistry to build vast amounts of biological data on a small glass chip. Fodor and his colleagues made the first scientific publication describing both microarray technology and combinatorial chemistry in 1991 [22]. It was a spatially addressable array of 1024 peptides. Soon after, they were able to construct the microarray containing over 1 million DNA sequences by combining photochemistry and photolithography with solid-phase DNA

synthesis chemistry. Later, as a demonstration study [23], scientists of Affymetrix made two types of light-generated DNA probe arrays to test a variety of mutations in the cystic fibrosis transmembrane conductance regulator (CFTR). One array, made up of 428 probes, was designed to scan through the length of CFTR exon 11 and identify differences from the wild-type reference sequence. The second type of array contained 1480 probes chosen to detect known deletions, insertions, or base substitution mutations. The validity of the probe arrays was established by hybridizing them with fluorescently labeled, controlled oligonucleotide targets. As result, ten unknown patient samples were genotyped and these genotype assignments were identical to those obtained by PCR product restriction fragment analysis. In 1996 the same year during which the joint Human Genome Program was proposed, Mirzabekov and his team in Moscow presented a SBH microchip [24] and its application to diagnostics for genetic diseases. A robot was constructed to manufacture sequencing "microchips". The microchip was an array of oligonucleotides immobilized into gel elements fixed on a glass plate. Hybridization of the microchip with fluorescently labeled DNA was monitored in real time simultaneously for all microchip elements with a two-wavelength fluorescence microscope equipped with a charge-coupled device (CCD) camera. The microchip had been used to detect beta-thalassemia mutations in patients by hybridizing PCR-amplified DNA with the microchips. The authors suggested the method for large-scale diagnostics and gene polymorphism studies.

General Procedure of Hybridization-based Biochips for DNA Mutation Detection

The experiment protocols for identification of DNA mutations by hybridization DNA chips include a number of successive steps (Fig. 1):

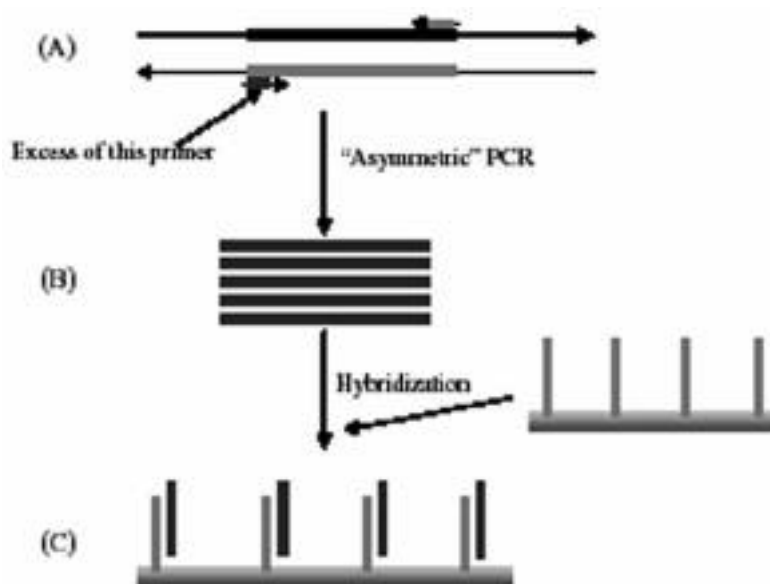


Fig. (1). Schematic representation of hybridization-based biochips for DNA mutation detection. (I) PCR amplification of clinical sample DNA's, (II) asymmetric PCR to yield a fluorescently labeled, predominantly single-stranded target DNA, and (III) Hybridization of the labeled product to the chip carrying immobilized oligonucleotides.

- Preparing oligonucleotide probe chip;
- PCR amplification of sample DNAs, during which synthesized DNA was labeled with fluorescent dyes. This step is necessary for a passive hybridization on chip, increasing the signal to detectable level;
- Asymmetric PCR to yield a fluorescently labeled, predominantly single-stranded target DNA. This is usually done in a separate test tube;
- Hybridization of the labeled products to the oligonucleotides immobilized on the chip, followed by successive washing;
- Signal measurement and data analysis.

To read and form the signal image from the chip, different designs were adopted and there are mainly two alternatives: gain-of-signal and loss-of-signal.

Gain-of-Signal Approach

Gain of hybridization signal analysis compares signals from probes complementary to mutant and wild-type sequences. Relative to their wild type counterparts, mutant targets should have increased affinity towards a corresponding mutation specific probe. This results in a "gain" of hybridization signal to this probe [25]. Relative 'gain' of signal by these probes indicates a sequence change similar to that of conventional dot blot analysis. The gain-of-signal approach allows for a partial scan of a DNA segment for all possible sequence variations. To interrogate both strands of an N bp long sequence for all single nucleotide substitutions, the array should consist of 8N probes (4N probes per strand). This total is derived from the fact that four probes complementary to each of the four possible target sequences for each nucleotide position on a given strand are represented in the array. To interrogate both target strands for all possible deletions of a specific length, 2N probes are needed, while $2(4^X)N$ probes are needed to scan for all X nt long insertions [26]. One early application of the method is the finding of extensive polymorphisms in HIV-1 clade B protease gene [27]. The sequences of 167 viral isolates from 102 patients had been determined. The gene was found to be extremely variable and 47.5% of the 99 amino acid positions varied. This level of diversity is greater than that previously known for all worldwide HIV-1 clades combined (40%). It is interesting that many of the amino acid changes that were known to contribute to drug resistance occurred as natural polymorphisms in isolates from patients who had never received protease inhibitors.

Loss-of Signal Approach

In the loss-of-signal approach, sequence variations are analyzed by quantitating the reduced signal from the test DNA at mutated positions. This can be done by either one-color labeling or two-color labeling [28]. In one-color labeling, ideally, a homozygous sequence change results in a complete loss of hybridization signal to perfectly match probes interrogating the region surrounding the sequence change as variation takes place at the two allelic sequences. Therefore, a 50% loss of signal intensity relative to the wild-type target would be found for heterozygous sequence varia-

tions. With this approach, an array designed to interrogate both target strands of N bp, for all possible sequence changes, minimally consist of 2N overlapping probes.

The two-color assay uses internal standards. The reference targets of known sequence are co-hybridized to the arrays along with the test target of interest. By labeling each target with a different fluorophore, a direct comparison of hybridization signals from these two targets can be made. This scheme is more or less similar to the competitive immunoassay where the target antigen competitively binds the immobilized antibody with the labeled known antigen, and the ratio of lost signal is the indication of the concentration of target antigen. The method has been used to analyze large DNA segments of hemizygous (mitochondrial DNA [29]) and heterozygous (BRCA1 [28]) sequence content.

Limitations of the Hybridization Chip-based Method

Despite many advantages, the hybridization chip-based mutation detection methods have several limitations. First, it requires high concentration of input DNA for efficiently rapid, passive hybridization. Amplification of sample DNA is necessary in separate PCR tubes, which adds extra steps for preparation, amplification and purification of DNA samples, bringing about cost and time consumption. Second, the oligonucleotide capture primers must be rigorously designed to maintain uniform stringency conditions for each hybridization reaction, which is rather difficult. Third, multiplex amplification often leads to large decreases in amplification efficiency, and requires extensive optimization of both amplification primer design and primer/reagent concentrations because of primer-primer interactions. Therefore, it is a challenge to integrate multiple amplifications with analysis of the amplified DNA on the same microchip. Finally, for the purpose of finding any mutations in a specific gene, it is too expensive to use the universal oligonucleotide chip. Taking loss of signal approach as an example, an array of 11,000 oligonucleotides would need to screen the 5.5 kb of BRCA1 for all possible sequence variations, when using one color labeling protocol. In case of two-color protocol, this number even increases more, to screen for all possible heterozygous germ-line mutations in the 9.17-kb coding region of the ATM gene, high-density arrays of > 90,000 oligonucleotide probes 25 nucleotides in length, were designed [30].

MINISEQUENCING-BASED APPROACH

Minisequencing is also called single nucleotide primer extension, in which extension of the primer by one base occurs only when the labeled nucleotide is complementary to the nucleotide of the target DNA adjacent to the 3'-end of the primer. As the allele-specific dye-labeled dNTP is linked to the primer extension in the presence of DNA polymerase and target DNA, failure of signal generation is the indication of mutation at the allele site. This principle became a chip format by Pastinen *et al.* [31]. Oligonucleotides are attached to the surface *via* 5'-end linkage to leave an exposed free 3'-OH group. The extension reaction occurs on the chip in the presence of all reaction components. The identity of the extended dideoxyribonucleotide is determined through fluorescent microscopy and used to assign the identity or identities of the target nucleotide extended from the 3' end of each

probe. All possible sequence changes of a 33-bp region of TP53 were analyzed by this approach. Kim *et al.* [32] reported an approach using solid phase single base extension for multiplex genotyping by mass spectrometry. The reaction products were purified by streptavidin coated solid phase magnetic beads, then released and analyzed by matrix-assisted laser desorption/ionization time-of-flight MS. Since only the pure extension DNA products are introduced to the MS for analysis, the resulting mass spectrum is free of non-extended primer peaks and their associated dimmers. Using the method, they simultaneously distinguished six nucleotide variations on synthetic DNA templates mimicking mutations in the p53 gene and two disease-associated SNAs in the human hereditary hemochromatosis gene.

DNA MUTATION DETECTION BY ON-CHIP PCR

On-chip PCR method is a dynamic process for DNA mutation detection, which undergoes two successive steps, all occurring on the chip. The first step is an asymmetric PCR, where one of a pair of PCR primers is immobilized on a chip matrix and the other is left free. Using sample DNA as template, single-stranded products with labeled primer at its 5' end accumulate. In the second step, the single-stranded asymmetric PCR products hybridize with the specific primers immobilized on the chip matrix, leading to the extension of the primers. Longer perfect duplexes formed during this process have significantly higher melting temperatures than the shorter duplexes between the single-stranded PCR product and immobilized primers. Therefore, when the temperature exceeds that of primer annealing, signals can be observed only in the wells in which the immobilized primer has been extended.

Huber *et al.* [33] developed such a method for detection of single base alterations in genomic DNA by solid-phase PCR on oligonucleotide array. They applied an experimental setup, which combines on-chip PCR of human genomic DNA with simultaneous nested solid phase amplification mediated by allele-specific oligonucleotide primers tethered to a glass slide. A gene-specific oligonucleotide tiling array contains covalently attached allele-specific primers which interrogate single nucleotide positions within a genomic region of interest. During a thermal cycling reaction, amplification products remain covalently bound to the solid support and can be visualized and analyzed by the incorporation of fluorescent dyes. Using the procedure, they unequivocally defined the presence of point mutations in the human tumor suppressor gene *p53* directly from a natural DNA source. This semi-multiplex solid phase amplification format allowed the rapid and correct identification of 20 nucleotide positions from minute amounts of human genomic DNA. Sergei [34] reported a method for parallel independent on-chip amplification and the following sequence variation analysis of multiple DNA regions directly using microchip with an array of nanoliter gel pads containing specific sets of tethered primers and applied it to simultaneously identify several abundant drug-resistant mutations in three genes of *Mycobacterium tuberculosis*. An allele-specific on-chip PCR method for identification of rifampin-resistant mycobacterium tuberculosis strains was developed by Mikhailovich V [35]. Using this method, the time to identify rifampin-resistant mycobacterium tuberculosis strains reduced to 1.5

h. In addition, it does not require any special probe preparation and can be applied to simultaneous analysis of several variable segments of the bacterial genome.

MUTATION DETECTION BY OLIGONUCLEOTIDE LIGATION ASSAY (OLA) ON BIOCHIP

In the late 1980s, it was found that, when two oligonucleotides were annealed immediately adjacent to each other on a complementary target DNA molecule, single nucleotide substitutions at the junction could be detected by T4 phage DNA ligase [36]. This was named as oligonucleotide ligation assay (OLA) by Nickerson *et al.* [37], who developed an automated, nonisotopic strategy for DNA diagnostics using amplification of target DNA segments and the discrimination of allelic sequence variants by a colorimetric OLA. A typical OLA chip uses three oligo primers: two immobilized allele-specific probes, one for the wild type allele and the other for the mutant allele, plus a labeled common probe. The forward primer has its 3' base place at the position of the variation and located immediately adjacent to the 5'-end of the common probe (or report probe). The reaction mixture contains the common probes, which hybridize with the target DNA immediately adjacent to the immobilized oligonucleotides and carry 5' phosphate to make the ligation reaction possible. After hybridization, ligation of both primers by DNA ligase will take place only when the 3'-end of the immobilized primer matches perfectly with the target sequence. The signal generated by the common probe reports the result of the detection. Nickerson *et al.* have applied the automated PCR/OLA procedure to the diagnosis of common genetic diseases, such as sickle cell anemia and cystic fibrosis, and to genetic linkage mapping of gene segments in the human T-cell receptor γ -chain locus. Barany *et al.* [38] later used a thermostable DNA ligase to make the procedure more practical. To enhance the positive fluorescence signal, the reaction was carried out at elevated temperature and all participating compounds underwent multiple cycles of annealing, ligation and melting. The ligase-mediated method was further modified and applied by several investigators to detect DNA mutations on large-scale. Chen *et al.* [39] combined the highly sensitive and specific PCR-OLA with FRET detection in one reaction in which the mutational status or genotype was determined without any purification or manipulation after the reaction was set up initially. FRET was observed when two fluorescent dyes were in close proximity and the donor dye's emission spectrum overlapped the acceptor dye's excitation spectrum. Gerry *et al.* [40] developed a new type of DNA microchip that combines PCR/OLA with "zip-code" hybridization. Zip-code sequences consist of 24 bases that are assembled from a set of 36 tetramers. Each tetramer differs from the others by at least two bases and is neither self-complementary, nor complementary to any other tetramer. Similarly, each zip code differs from the others by at least three tetramer units. The end result of this design is sequences that have comparable behavior in terms of the thermodynamics and kinetics of hybridization, while simultaneously maintaining distinct chemical identities that prevent cross-hybridization. Suitably designed allele-specific LDR (ligase detection reaction) primers become covalently ligated to adjacent fluorescently labeled primers if a mutation is present. The allele-specific OLA primers contain on

their 5'-ends "zip-code complements" that are used to direct OLA products to specific zip-code addresses attached covalently to a three-dimensional gel-matrix array. Since zip-codes have no homology to either the target sequence or to other sequences in the genome, false signals due to mismatch hybridizations are not detected. The zip-code sequences remain constant and their complements can be appended to any set of OLA primers, making the zip-code arrays universal.

OLA method can detect any nucleotide variation at the ligation junction using a single set of assay conditions. Other advantages include high specificity, speed and automation, as well as compatibility with PCR, making it suitable for genotyping on large-scale detection. However, the assay can only detect mutations at the ligation junction. False negative could happen if the mutations occur in the nearby sequence. Therefore the method needs to be improved further.

Short Oligonucleotide Ligation Assay on Chip (SOLAC)

It was known that T4 DNA ligase could ligate pentamers efficiently in solution. Pitchard and Southern [41] studied the effect of mismatches on ligation of short oligonucleotides, and found that even mismatches at the distal position from the ligation junction site were able to completely inhibit the ligation of octamers when *Thermus thermophilus* (Tth) DNA ligase was used. This suggests that shorter report probe may prove highly sensitive to the mismatch. Based on these

findings, we decided using pentamers, other than using longer ones, as the allelic (detecting) probe [42]. The approach is termed short oligonucleotide ligation assay on DNA chip (SOLAC), which can be carried out through two experiment schemes: loss-of-signal (SOLAC-LOS) and gain-of-signal (SOLAC-GOS).

In SOLAC-LOS, the common probe is immobilized on the chip and the allele specific probe is used as the detecting probe; both are perfectly complementary to the wild type target DNA. After hybridization of sample DNA with the immobilized common probe, T4 DNA ligase is applied to ligate the common probe and the detecting probe. Failure ligation occurs if there is any mismatch between the sample DNA and the detecting probe. This nick containing hybrid conjugate could not stand the denaturing and washing treatment, leading to the loss of signal, which indicates the presence of mutations in the target sample. Theoretically, with one pair of probes (one common probe and one pentamer), it can detect all mutations (substitutions, insertions and deletions) in a five-base region of target DNA.

In SOLAC-GOS, contrarily, the solid phase is the array of allele specific probes, which are designed complementary to all the known mutations of the target region of the sample DNA, while the common probes are detecting probes. After hybridization, ligation and washing, the gain-of-signal is the indication of the presence of mutations. Fig. 2 is the sche-

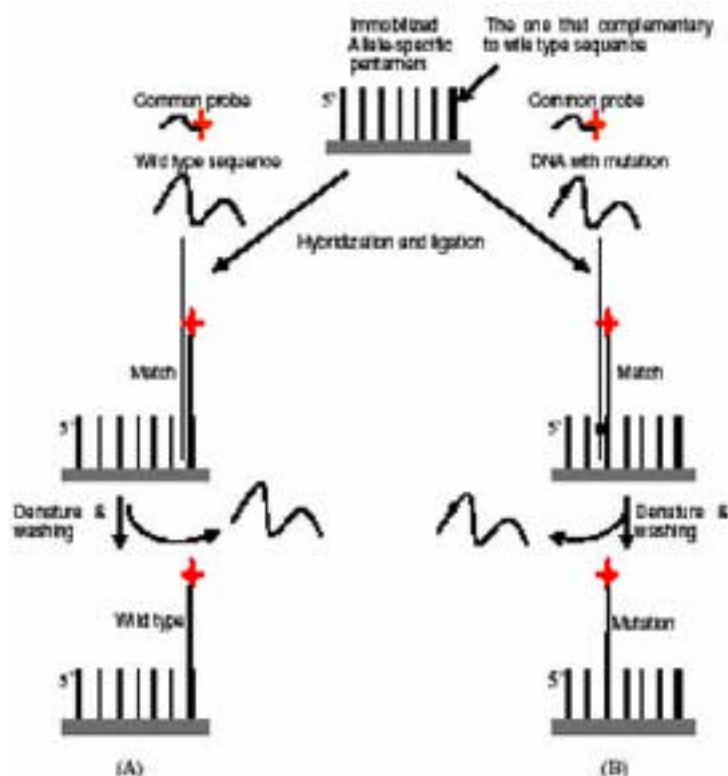


Fig. (2). Schematic diagram of SOLAC-GOS. The common probe is labeled with marker. The allele-specific probes are immobilized on the chip through DNA chemistry. Each allele-specific probe contains one single base variation at its 5' end except one (Bold line) that is complementary to the wild-type sequence. A) the sample DNA is a wild-type sequence; as a result, the signal is seen at the position of the immobilized allele probe that matches to the wild-type sequence; B) the sample DNA contains a single base mutation; the signal is seen at the position of the immobilized allele-specific probe that matches to the mutation.

matic of the approach. For a five-base region of a target DNA, basically it needs sixteen allele-specific pentamers and just one common probe to detect all possible mutations. Obviously, SOLAC-GOS needs more allelic probes than SOLAC-LOS does, but the strategy is more attractive because that the signal is directly linked to the mismatch, which not only is favorable to the users but can also avoid some false operations that do not generate signals.

It is figured out that ligation efficiency of T4 ligase is a critical point. The effects of various factors on the ligation reaction were investigated. In general, the closer the mismatch was to the ligation junction site it was taking G:T mismatch as an example, the easier to distinguish, or vice versa. The specificity of T4 DNA ligase was found to increase as the temperature increased. When ligations were carried out at room temperature (25°C), G:T mismatch at the fourth position of the pentamer was difficult to be discriminated. The background of the mismatch ligation was rather high. When the temperature of ligation was raised to 30°C, no detectable background signal was observed. The efficiency of mismatch discrimination was also influenced by the concentration of allele-specific pentamers, and this influence varied with the mismatch position. For instance, once the mismatch occurred at the first position, it could be distinguished effectively even if the concentration of the pentamer was as high as 10 µM. But the concentration of pentamer had to be decreased to 0.5 µM for successfully distinguishing the mismatch if it occurred at the fourth base of the 5' terminus.

In combination with the alkaline phosphatase reaction linked assay, these two schemes were used for the identification of the mutations in the *rpoB* gene of *Mycobacterium tuberculosis* from clinic isolates that revealed the rifampin resistant (Rifr) [43]. The results are almost identical to the results of DNA sequencing. However, false negative results must be encountered when the mutation sites are not in the target region. This can be overcome by designing more allelic probes on the chip, which is easy to realize with array technology.

Microfluidic Chips

This is a combined use of ligase detection reaction and microfluidic chip [44]. Low-density arrays were assembled into microfluidic channels to allow the detection of low-abundant mutations in K-ras gene fragments. Following spotting, the chip was assembled with a cover plate and the array accessed using microfluidics in order to enhance the kinetics associated with hybridization. The array was configured with zip code sequences (24-mers) that were complementary to sequences of the target. The hybridization targets were generated using an allele-specific ligase detection reaction, in which two primers (discriminating primer that carries the complement base to the mutation being interrogated and a common primer that flank the point mutation) were ligated, joined together only when the particular mutation to the appropriate site of the array, and the common primer carried on its 3' end a fluorescent dye (near-IR dye IRD-800). The coupling chemistry was optimized to maximize the loading level of normal sequences, and increase the stability of the linkage chemistry to permit re-interrogation of the array. The microfluidic addressing of the array reduced the hybridiza-

tion time from 3 h for a conventional array to less than 1 min. In addition, the coupling chemistry allowed reuse of the array > 12 times before noticing significant loss of hybridization signal. The array was used to detect a point mutation in a K-ras oncogene at a level of 1 mutant DNA in 10,000 wild-type sequences.

MutS-BASED BIOCHIPS FOR DNA MUTATION DETECTION

MutS Protein and DNA Mismatch Repair System

DNA mismatch repair (MMR) system exists in all organisms ranging from bacteria to the human and increases the fidelity of DNA replication by 100-1,000 times. *E. coli* MMR system contains at least three proteins: MutS, MutL and MutH, each of them has its own function but they interact with each other during the mismatch repair processing. MutS protein is a crucial component of the MMR system, which specifically recognizes and binds all possible single-base mismatches as well as 1-4 base insertion or deletion loops specifically with varying affinities independent of other proteins or cofactors [45-47]. MutS can be used as a recognition element for DNA mutations in either DNA chip or protein chip format.

MutS-based DNA Chip

Behrendorf HA *et al.* [48] initially used MutS as a recognition element of DNA mutations on the DNA chip. The method combines the features of electronically controlled DNA hybridization on open-format microarrays, with mutation detection by a fluorescence-labeled mismatch-binding protein. Electronic addressing of DNA strands to distinct test sites of the chip allows parallel analysis of several individuals and mutations in different exons of the p53 gene.

As a further approach, we constructed three different MutS fusion proteins, Trx-His₆-GFP-(Ser-Gly)₆-MutS (THGLM), Trx-His₆-(Ser-Gly)₆-Strep tagII-(Ser-Gly)₆-MutS (THLSLM) and Trx-His₆-(Ser-Gly)₆-MutS (THLM) as the mutation recognition elements (Fig. 3) [49]. The fusion protein THGLM is a ready-for-use reagent, which recognizes the mutation site and signals directly through its fusion partner, the green fluorescent protein (GFP). THLSLM contained Strep tagII that could specifically interact with streptavidin labeled with alkaline phosphatase, and THLM was subsequently modified with Cy3 dye. Thus, all these fusion proteins possess two functions: recognizing and binding mutations and generating signals. The bioactivities of these various fusion configurations were characterized, and comparison study concerning their usage build the reliable DNA chips was carried out. The common procedure of the detection using these fusions is schematically as represented follows,

- 1) Immobilization of 5'-disulfide DNA probes on the SH-modified glass support;
- 2) Hybridization of sample DNA to the immobilized DNA probes to form mismatched or complementary duplex DNA;
- 3) Binding of MutS fusion to the mismatched duplex DNA.

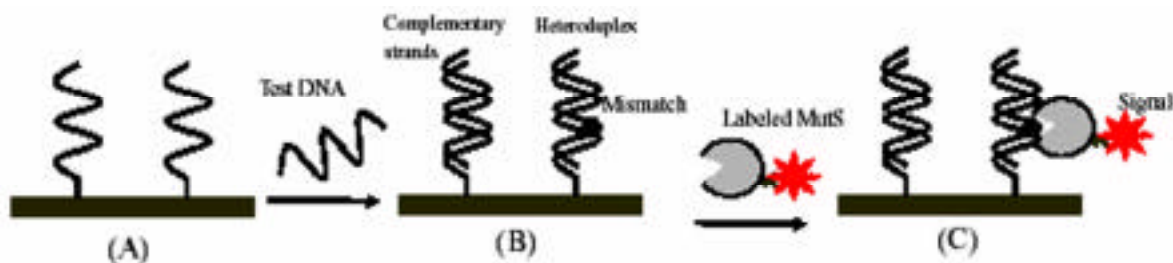


Fig. (3). Schematic representation of MutS-based DNA chip assay. (A) Reference strands (e.g. PCR products) are first immobilized to individual test sites of the array. (B) Test strands are hybridized to the reference duplex strands, thereby generating heteroduplex DNA or complementary DNA. (C) Labeled MutS binds preferentially to mismatched heteroduplex DNA.

After this common procedure, signal could be obtained through different means depending on each fusion format, i.e. GFP signaling, cy3 signaling or enzymatic signaling. The resulting DNA chips were used to detect the mismatched and unpaired mutations in the synthesized oligonucleotides, as well as the single base mutation in *rpoB* gene from *Mycobacterium tuberculosis*. Each MutS fusion has its own features. For instance, THGLM is a MutS-GFP fusion that is ready-for-use, and no subsequent labeling step is needed; THLSLM can be coupled to alkaline phosphatase directly through binding of Strep tagII to streptavidin, and the assay result can be seen directly without any machine; however, THLM with the label of cy3 gives the highest intensive of signal.

MutS-based Protein Chip

A MutS-based protein chip method for detection of single-base mismatches and unpaired bases of DNA, using a genetic fusion molecular system Trx-His₆-Linker peptide-Strep-tagII-Linker peptide-MutS (THLSLM) has already been established by Bi et al [50] as shown in Fig. 4. The THLSLM coding sequence was constructed by attaching *Strep-tag II* and *mutS* gene to pET32a (+) sequentially with insertion of a linker peptide coding sequence before and behind *Strep-tagII* gene, respectively. The backbone of the three fusions was built up with a package of Trx for solubilization, a His₆ for affinity purification, a (Ser-Gly)₆ for linkage and spacing, and MutS for recognizing and binding to the target DNA. THLSLM retained both mismatch recognition activity and streptavidin binding affinity. THLSLM was then immobilized on the chip matrix coated with the streptavidin through Strep-tag II-streptavidin binding reaction. The resulting protein chip was used to detect the mismatched

and unpaired mutations in the synthesized oligonucleotides, as well as a single-base mutation in *rpoB* gene from *Mycobacterium tuberculosis*, with high specificity. As an alternative to the MutS-based DNA chip, the MutS protein chip has two main advantages. Firstly, other than labeling MutS protein, which is rather difficult, contrarily only nucleotide acids were labeled, which is a facile job. Secondly, unlike the multi-oligonucleotide chip, the MutS-based protein chip is much easier to prepare because it is a unique biomolecule chip. This advantage was more obvious when an “anchor chain” model [51], a gene manipulation-based technique was introduced, where the orientation of all the protein molecules was well controlled and the homogeneity of the chip could be greatly improved [52].

Normally, DNA chip is an addressable array of many different probes, which is difficult to prepare but easy to use because one DNA sample directly reacts with the whole array. MutS protein chip functions contrarily, where the chip array is made of a unique element, i.e. MutS, which is easy to prepare, but its sample handling requires a robot facility, which is expensive.

SUMMARY

Methodology for mutation detection has been improved obviously during the last decade, not only due to the introduction of new principles, but also because of for the emergence of the chip technology, on which various principles are easy applied. One of the main advantages of these chip-based technologies is certainly high throughput, which is particularly obvious when using high density hybridization microarray. Each method has its own peculiarity, but they all share the features of fast, automate, easy to read out and ac-

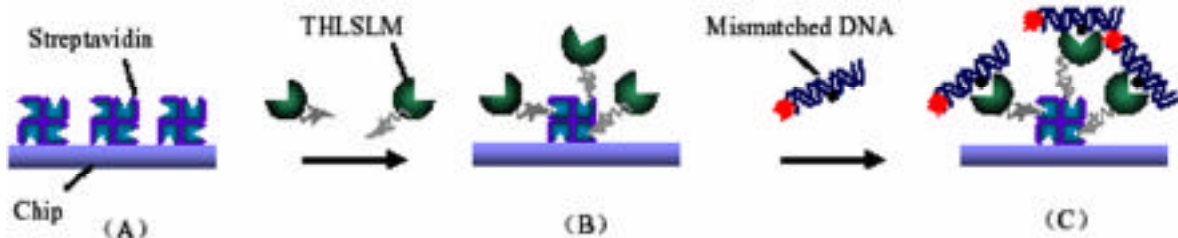


Fig. (4). Schematic diagram of MutS-based protein chip for gene mutation assay: (A) Streptavidin-modified chip; (B) THLSLM immobilized on the chip through the interaction of streptavidin/Strep-tag II; (C) Cy3-labeled mismatched heteroduplex DNA binding preferentially to THLSLM.

curate. The major challenges are still faced due to conventional DNA sequencing, owing to a drastic rise in the cost has been getting down dramatically along with the technology revolution, such as four-color fluorescent dye labeling and microfluidic chip capillary electrophoresis. Nevertheless, the biochip-based methods can be alternative means or, in many cases, the first choice for mutation detection.

ACKNOWLEDGEMENTS

We sincerely thank the National Science Foundation of China (30500097), Ministry of Science & Technology and the Chinese Academy of Science for their long-term supports to our studies on the methodology for mutagenesis analysis.

REFERENCES

- [1] Stewart, P.R.; Adhami, W.; Inglis, B.; Franklin, J.C. *J. Med. Microbiol.*, **1993**, *38*, 270.
- [2] Meggouh, F.; Benomar, A.; Rouger, H.; Tardieu, S.; Birouk, N.; Tassin, J.; Barhoumi, C.; Yahyaoui, M.; Chkili, T.; Brice, A.; LeGuern, E. *J. Med. Genet.*, **1998**, *35*, 251.
- [3] Zoller, B.; Dahlback, B. *Curr. Opin. Hematol.*, **1995**, *2*, 358.
- [4] Everett, L.A.; Glaser, B.; Beck, J.C.; Idol, J.R.; Buchs, A.; Heyman, M.; Adawi, F.; Hazani, E.; Nassir, E.; Baxevanis, A.D.; Sheffield, V.C.; Green, E.D. *Nat. Genet.*, **1997**, *17*, 411.
- [5] Losekoot, M.; Fodde, R.; Harteveld, C.L.; Heeren, H.; Giordano, P.C.; Bernini, L.F. *Br. J. Haematol.*, **1990**, *76*, 269.
- [6] Finkelstein, J.E.; Francomano, C.A.; Brusilow, S.W.; Traystman, M.D. *Genomics.*, **1990**, *7*, 167.
- [7] Dockhorn, D.B.; Dworniczak, B.; Brommelkamp, L.; Bulles, J.; Horst, J.; Bocker, W.W. *Nucleic Acids Res.*, **1991**, *19*, 2500.
- [8] Cotton, R.G.H.; Rodrigues, N.R. *Proc. Acad. Sci. U.S.A.*, **1988**, *85*, 4397.
- [9] Bottema, C.D.; Sommer, S.S. *Mutat. Res.*, **1993**, *288*, 93-102.
- [10] Wabuyele, M.B.; Farquar, H.; Stryjewski, W.; Hammer, R.P.; Soper, S.A.; Cheng, Y.W.; Barany, F. *J. Am. Chem. Soc.*, **2003**, *125*, 6937.
- [11] Frutos, A.G.; Pal, S.; Quesada, M.; Lahiri, J. *J. Am. Chem. Soc.*, **2002**, *124*, 2396.
- [12] Campsall, P.A.; Au, N.H.; Prendiville, J.S.; Speert, D.P.; Tan, R.; Thomas, E.E. *J. Clin. Microbiol.*, **2004**, *42*, 1409.
- [13] Geiss, G.K.; Bumgarner, R.E.; An, M.C.; Agy, M.B.; Wout, A.B.; Hammersmark, E.; Carter, V.S.; Upchurch, D.; Mullins, J.I.; Katze, M.G. *Virology*, **2000**, *266*, 8.
- [14] Zweiger, G. *Trends. Biotechnol.*, **1999**, *17*, 429.
- [15] Furushima, K.; Shimo-Onoda, K.; Maeda, S.; Nobukuni, T.; Ikari, K.; Koga, H.; Komiya, S.; Nakajima, T.; Harata, S.; Inoue, I. *J. Bone Miner. Res.*, **2002**, *17*, 128.
- [16] Jain, K.K. *Curr. Opin. Drug. Discov. Devel.*, **2004**, *7*, 285.
- [17] Southern, E.M. *J. Mol. Biol.*, **1975**, *98*, 503.
- [18] Southern, E.M. Patent Number: WO8910977, Publication date: 1989-11-16
- [19] Southern, E.M. *Trends Genet.* **1996**, *12*, 110.
- [20] Saiki, R.K.; Walsh, P.S.; Levenson, C.H.; Erlich, H.A. *Proc. Natl. Acad. Sci. U.S.A.*, **1989**, *86*, 6230.
- [21] Schwonbeck, S.; Krause-Griep, A.; Gajovic-Eichelmann, N.; Ehrentreich-Forster, E.; Meinel, W.; Glatt, H.; Bier, F.F. *Bioelectron.*, **2004**, *15*, 956.
- [22] Fodor, S.P.; Read, J.L.; Pirrung, M.C.; Stryer, L.; Lu, A.T.; Solas, D. *Science*, **1991**, *251*, 767.
- [23] Cronin, M.T.; Fucini, R.V.; Kim, S.M.; Masino, R.S.; Wespi, R.M.; Miyada, C.G. *Hum. Mutat.*, **1996**, *7*, 244.
- [24] Yershov, G.; Barsky, V.; Belgovskiy, A.; Kirillov, E.; Kreindlin, E.; Ivanov, I.; Parinov, S.; Guschin, D.; Drobishev, A.; Dubiley, S.; Mirzabekov, A. *Proc. Natl. Acad. Sci. U.S.A.*, **1996**, *93*, 4913.
- [25] Hacia, J.G.; Collins, F.S.; *J. Med. Genet.*, **1999**, *36*, 730.
- [26] Hacia, J.G. *Nature Genetics supplement*, **1999**, *21*, 42.
- [27] Kozal, M. *Nature Med.*, **1996**, *2*, 753.
- [28] Yim, S.C.; Park, H.G.; Chang, H.N.; Cho, D.Y. *Anal. Biochem.*, **2005**, *15*, 332.
- [29] Chee, M.; Yang, R.; Hubbell, E.; Berno, A.; Huang, X.C.; Stern, D.; Winkler, J.; Lockhart, D.J.; Morris, M.S.; Fodor, S.P., *Science*, **1996**, *274*, 610.
- [30] Hacia, J.G.; Sun, B.; Hunt, N.; Edgemon, K.; Mosbrook, D.; Robbins, C.; Fodor, S.P.A.; Tagle, D.A.; Collins, F.S. *Genome Research*, **1998**, *8*, 1245.
- [31] Pastinen, T.; Kurg, A.; Metspalu, A.; Peltonen, L.; Syvanen, A.C. *Genome Res.*, **1997**, *7*, 606.
- [32] Kim, S.; Edwards, J.R.; Deng, L.Y.; Chung, W.; Ju, J.Y. *Nucleic Acids Research*, **2002**, *30*, e85.
- [33] Huber, M.; Losert, D.; Hiller, R.; Harwanegg, C.; Mueller, M.W.; Schmidt, W.M. *Anal. Biochem.*, **2001**, *299*, 24.
- [34] Tillib, S.V.; Strizhkov, B.N.; Mirzabekov, A.D. *Anal. Biochem.*, **2001**, *292*, 155.
- [35] Mikhailovich, V.; Lapa, S.; Gryadunov, D.; Sobolev, A.; Strizhkov, B.; Chernyh, N.; Skotnikova, O.; Irtuganova, O.; Moroz, A.; Litvinov, V.; Vladimirovskii, M.; Perelman, M.; Chernousova, L.; Erokhin, V.; Zasedatelev, A.; Mirzabekov, A. *J. Clin. Microbiol.*, **2001**, *39*, 2531.
- [36] Landegren, U.; Kaiser, R.; Sanders, J.; Hood, L. *Science*, **1988**, *241*, 1077.
- [37] Nickerson, D.A.; Kaiser, R.; Lappin, S.; Stewart, J.; Hood, L.; Landegren, U. *Proc. Natl. Acad. Sci. U.S.A.*, **1990**, *87*, 8923
- [38] Barany, F. *Proc. Natl. Acad. Sci. U.S.A.*, **1991**, *88*, 189.
- [39] Chen, X.; Livak, K.J.; Kwok, P.Y. *Genome Res.*, **1998**, *8*, 549.
- [40] Gerry, N.P.; Witowski, N.E.; Day, J.; Hammer, R.P.; Barany, G.; Barany, F. *Nucleic Acids Res.*, **1997**, *25*, 3403.
- [41] Pritchard, C.E.; Southern, E.M. *Nucleic Acids Res.*, **1997**, *25*, 3403.
- [42] Deng, J.Y.; Zhang, X.E.; Mang, Y.; Zhang, Z.P.; Zhou, Y.F.; Liu, Q.; Lu, H.B.; Fu, Z.J. *Biosens. Bioelectron.*, **2004**, *19*, 1277.
- [43] Deng, J.Y.; Zhang, X.E.; Lu, H.B.; Liu, Q.; Zhang, Z.P.; Zhou, Y.F.; Xie, W.H.; Fu, Z.J. *J. Clin. Microbiol.*, **2004**, *42*, 4850.
- [44] Wang, Y.; Vaidya, B.; Farquar, H.D.; Stryjewski, W.; Hammer, R.P.; McCarley, R.L.; Soper, S.A.; Cheng, Y.W.; Barany, F. *Anal. Chem.*, **2003**, *75*, 1130.
- [45] Lu, A.L.; Clark, S.; Modrich, P. *Proc. Natl. Acad. Sci. U.S.A.*, **1983**, *80*, 4639.
- [46] Lieb, M. *J. Bacteriol.*, **1987**, *169*, 5241.
- [47] Jiricny, J.; Su, S.S.; Wood, S.G.; Modrich, P. *Nucleic Acids Res.*, **1988**, *25*, 7843.
- [48] Behrens, H.A.; Pignot, M.; Windhab, N.; Kappel, A. *Nucleic Acids Res.*, **2002**, *30*, e64.
- [49] Bi, L.J.; Zhou, Y.F.; Zhang, X.E.; Deng, J.Y.; Wen, J.K.; Zhang, Z.P. *Bioelectron.*, **2005**, *21*, 135.
- [50] Bi, L.J.; Zhou, Y.F.; Zhang, X.E.; Deng, J.Y.; Zhang, Z.P.; Xie, B.; Zhang, C.G. *Anal. Chem.*, **2003**, *75*, 4113.
- [51] Shao, W.H.; Zhang, X.E.; Liu, H.; Zhang, Z.P.; Cass, A.E.; *Bioconjug. Chem.*, **2000**, *11*, 822.
- [52] Shi, J.X.; Zhang, X.E.; Xie, W.H.; Zhou, Y.F.; Zhang, Z.P.; Deng, J.Y.; Zhang, Z.L.; Pang, D.W.; Zhang, C.G.; *Anal. Chem.*, **2004**, *76*, 632.