

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
12 September 2008 (12.09.2008)

PCT

(10) International Publication Number
WO 2008/106987 A1

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/EE2007/000003

(22) International Filing Date: 6 March 2007 (06.03.2007)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant (for all designated States except US): **EESTI BIOKESKUS** [EE/EE]; Riia 23B, EE51010 Tartu (EE).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **KRJUTSKOV, Kaarel** [EE/EE]; Riia 23, EE51010 Tartu (EE). **METSPALU, Andres** [EE/EE]; Riia 23, EE51010 Tartu (EE).

(74) Agents: **SARAP, Margus** et al.; Patent Agency Käosaar & Co, Tähe 94, EE50107 Tartu (EE).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

(54) Title: A METHOD TO DETERMINE SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS IN NUCLEIC ACID SEQUENCE

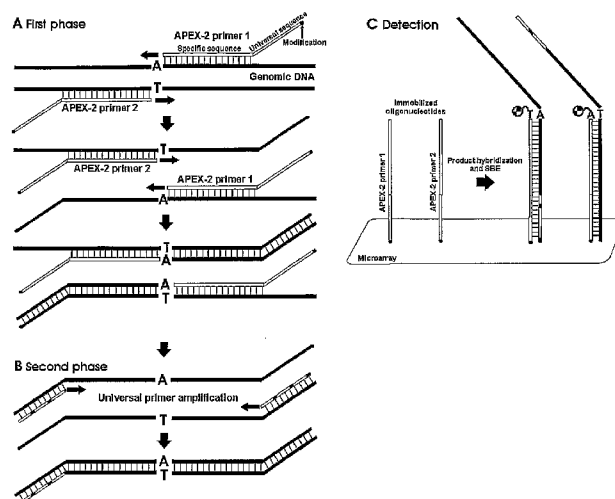


Fig. 5

(57) Abstract: A genotyping method and a prepared oligomicroarray as device are to determine single nucleotide polymorphism and mutations are provided. The method uses two specific APEX-2 primers per each studied nucleotide or mutations to be determined. The same primers are used in amplification phase (primer extension and PCR with universal primer) and in the single base extension phase on an array. All studied nucleotide-containing sequences can be genotyped and amplified in one reaction tube and visualized on a microarray.

Description

A method to determine single nucleotide polymorphisms and mutations in nucleic acid sequence

Technical Field

[0001] This invention relates generally to genotyping technologies and mutation detection. More specifically this invention relates to methods to determine single nucleotide polymorphisms and mutations in genome. This invention is suitable for nucleic acid primary structure analyzes, for epigenetic effects and in molecular diagnostics.

Background Art

[0002] Genotyping technologies enable to construct high-resolution genetic maps for genomes of human and other organisms, resulting in LD blocks or r^2 bins dependent of the method used. These LD blocks (haplotypes) can be used in association studies to localize genes responsible for a particular phenotype (e.g. disease). In other words, the aim of genotyping is to find correlations between genome organization (haplotype) and a particular phenotype.

[0003] It is becoming increasingly evident, that hundred thousands single nucleotide polymorphisms (SNP-s) have to be analyzed simultaneously to get a complete picture of haploblocks' (or r^2 bin) arrangement. Thus, new methods are needed to accomplish this expansive and expensive task. Below is an overview of well-known genotyping methods that despite of their high throughput technology have also their limiting factors and therefore new cost-effective, flexible and focused genotyping methods are needed.

[0004] Genotyping and mutation detection is increasingly more substantial in clinical practice. Modern and reliable methods are alternative for RFLP, allele-specific PCR or sequencing in future. Flexible genotyping and mutation detection analyzes systems help to diagnose tens of candidate genetic markers in one reaction in short time as a routine procedure.

[0005] 1. Detection based hybridization methods such as Affymetrix *GeneChip* (Figure 1.) technology enables to analyze half a million SNP-s
MATSUZAKI, H., et al. Parallel genotyping of over 10,000 SNPs using a

one-primer assay on a high-density oligonucleotide array. *Genome Res.* 2004, no.14, p.414-425. Studied loci in genomic DNA are amplified without PCR with specific primers. An alternative way to reduce complexity of the genome is to use the restriction enzyme cut of genomic DNA, ligation of the fragments produced in this way to a universal linker and amplification with universal primers in PCR. Amplified regions are fragmented, labeled and hybridized to complementary oligonucleotides synthesized on microarray. In case of a perfect match a duplex is formed and a signal from the perfect match pairings, which is higher than signals from the mismatched pairing, is detected. *GeneChip* microarray uses 25 base pairs (bp) long oligonucleotides containing a central nucleotide corresponding to the studied SNP. Hence, on both sides of a SNP, there are regions with 12 bp perfect match and in the middle, at the position of the SNP, a perfect match or a mismatch can occur.

- [0006] The key step of the method is DNA restriction and addition of universal linker sequences to the ends of all created fragments. Subsequently 250-2000 bp long fragments, containing the SNP(s) of interest are amplified with primers that anneal to the universal linker sequences (universal primer). With this approach a vast number of genomic sequences can be amplified with minimal cost. The limiting factor in this method is the number of restriction sites across the genome. Thus, SNP-s not located in the synthesized fragments (250-2000 bp), cannot be detected. Hence, this method in principle doesn't cover the whole genome and therefore the Illumina 300K array is as informative as Affymetrix 500K array in CEU population BARRETT, J. C., et al. Evaluating coverage of genome-wide association studies. *Nat Genet.* June 2006, vol.38, no.6, p.659-62.
- [0007] 2. Affymetrix *Molecular Inversion Probe* (MIP) method uses MIP molecules, which are special "padlock" probes NILSSON, M., et al. Padlock probes: circularizing oligonucleotides for localized DNA detection.. *Science.* 1994, no.265, p.2085-2088. for genotyping. MIP molecule is a linear oligonucleotide that contains specific regions, universal sequences, restriction sites and a *Tag (index)* sequence (16-22 bp). MIP hybridizes directly around the genetic marker/SNP of interest (Figure 2).

- [0008] MIP method uses 1500 “padlock” probe sets that hybridize to genomic DNA in parallel HARDENBOL, P., et al. Multiplexed genotyping with sequence-tagged molecular inversion probes. *Nat. Biotechnol.* 2003, no.21, p.673–678. In case of a perfect match binding genomic homology regions are ligated by creating a circular molecule. After the first restriction all molecules are amplified with universal primers. Amplicons are restricted again to ensure short fragments for hybridization on microarray. Generated short fragments are labeled and through *Tag* sequence hybridized to *cTag* (*complementary strand for index*) on array. After the formation of *Tag-cTag* duplex a signal is detected.
- [0009] Despite the complexity of experimental procedures, it is possible to amplify up to 10 000 SNP-containing sequences in one reaction. Thereby polymorphisms can theoretically be detected in any genomic region of interest HARDENBOL, P., et al. Highly multiplexed molecular inversion probe genotyping: Over 10,000 targeted SNPs genotyped in a single tube assay. *Genome Res.* 2005, no.15, p.269–275. There are two MIP probes for each allele, thus the method uses four probes (70 to 100 bp), a universal primer and a *cTag* sequence on array for a SNP detection in both strand.
- [0010] 3. Illumina *GoldenGate* genotyping platform outstands primarily with an original solution of the gene chip, but the molecular approach amplifying genomic regions is similar to MIP probes. Genomic DNA fragments are attached to specific particles, followed by hybridization with specific probe molecules (Figure 3).
- [0011] Probe molecules are supplied with three different universal sequences and a *Tag* sequence, situated between the specific region and universal primer GUNDERSON, K. L., et al. A genome-wide scalable SNP genotyping assay using microarray technology. *Nat. Genet.* 2005, no.37, p.549–554. Specific primers are hybridized to genomic DNA and designed to be allele specific, meaning that in order to identify a SNP, an oligonucleotide must be synthesized in a way that its 3' end binds to the SNP under study.
- [0012] Hence, to determine a SNP two probe molecules, each supplied with a different universal primer sequence, are needed. In case of a 3' perfect

base-pairing at the 3'-end one primer is elongated by primer extension reaction up to the other oligonucleotide, followed thereafter by ligation and generation of a linear molecule. The formed molecule includes two universal primer-binding sites and a *Tag* sequence.

- [0013] The formation of a linear molecule enables the PCR amplification with universal primers, which in turn are supplied with two different fluorescence labels to detect homo- or heterozygosity at the studied position. *GoldenGate* method uses hybridization-based signal detection through formation of *Tag-cTag* complex.
- [0014] 4. The *Infinium 1* method of Illumina (Figure 4) enables similarly to Affymetrix *GeneChip* technology a genome-wide analysis and to test up to half a million SNPs on a gene chip so far. The method doesn't use PCR to amplify the studied loci. Instead genomic DNA is amplified using the WGA method (*Whole Genome Amplification*) GUNDERSON, K. L., et al. A genome-wide scalable SNP genotyping assay using microarray technology. *Nat. Genet.* 2005, no.37, p.549–554.
- [0015] Amplified genomic DNA is fragmented and hybridized to oligonucleotides (75 bp) on the gene chip. After hybridization from 16 to 18 hours unbound or mismatched fragments are removed during a specific wash step.
- [0016] The studied genotype on the gene chip is determined by using allele-specific oligonucleotides (*Infinium 1*), where the 3'- terminal nucleotide is complementary to the SNP and which are primer extended away from the marker. In case of a perfect match primer sequence is elongated by DNA polymerase with fluorescence-labeled desoxynucleotides, which in case of the switch in the chain give a fluorescence signal. In order to detect a SNP two oligonucleotides are used on a gene chip.
- [0017] *Infinium 2* platform incorporates two-color single base extension to detect a single nucleotide polymorphism with *100K BeadChip*, using only one oligonucleotide per SNP STEEMERS, F. J., et al. Whole-genome genotyping with the single-base extension assay. *Nat Methods.* Januar 2006, vol.3, no.1, p.31-3.

[0018] Compared to *GeneChip* technology *Infinium 1 and 2* aren't defined by the restriction sites in genomic DNA and therefore has a good potential to solve the genotyping of all SNPs.

Disclosure of Invention

[0019] The method according to this description is a genotyping and mutation analysis method that enables a robust, flexible, cost-effective SNP and mutation analysis accompanied with high fidelity. The method is hereinafter called APEX-2 (Arrayed Primer EXtension using Universal primer). APEX-2 is very well suited for a focused study with 15 to 1500 SNPs or mutations. In principle all SNP-containing sequences can be amplified in one reaction tube and further genotyped on a microarray by single base extension.

Brief Description of Drawings

[0020] Fig. 1 schematically shows the known method of Affymetrix *GeneChip*.

[0021] Fig. 2 schematically shows the principle of Affymetrix *MIP method*.

[0022] Fig. 3 schematically shows that principle of *Illumina Bead Chip* method.

[0023] Fig. 4 schematically shows the principle of *Illumina Infinum 1 and 2* method.

[0024] Fig. 5 schematically shows the principle of APEX-2 and three main phases (preferred embodiment).

[0025] Fig. 6 schematically shows the principle of APEX-2, second embodiment.

[0026] Fig. 7 schematically shows an example of the APEX-2 primer structure. In this example the universal sequence has sequence gatcaggcgtctgtcgtgctc (SEQ ID NO:1); the first APEX-2 primer has sequence gattgagctgctgctttctctcctt (SEQ ID NO:2) and the second APEX-2 primer has sequence ccctgcctcacacctgatagcac (SEQ ID NO:3)

[0027] Fig. 8 schematically shows binding of the specifically designed primers to genomic DNA and their location with respect to the studied SNP or mutation in the APEX-2 system.

[0028] Fig. 9 schematically shows detection of mutations (insertion and deletion) in the APEX-2 system.

[0029] Fig. 10 illustrates visualization of the proof of principle for APEX-2 genotyping method using 621 genetic loci

Mode(s) for Carrying Out the Invention

[0030] APEX-2 method, according to preferred embodiment, enables template amplification through high levels of multiplexing PCR in two phases: primer extension (First phase) and PCR with universal primers (Second phase) (Figure 5). That means, in the first step a specific primer extension product (APEX-2 pre-product) will be obtained where studied nucleotide (e.g. SNP or mutation) has been analyzed, in the second step we just increase the amount of this product using universal primer in order to have enough molecules to be detected in signal detector. Only new information in this amplicon is the studied nucleotide itself. In detail, in the first phase, the method uses two specific oligonucleotide primers for determination of one studied nucleotide. These primers are called APEX-2 primers. First APEX-2 primer binds to the coding DNA strand and the second APEX-2 primer to its complementary DNA strand. As a result of primer extension in the first phase, the APEX-2 pre-product is formed. The Pre-product contains two APEX-2 primer sequences, complementary sequences for APEX-2 primers and studied nucleotide pair in the middle of the two APEX-2 primers (see Fig. 5). In this case only one new non primer studied nucleotide will be amplified. The second phase (universal primer PCR) is essential for amplification of the pre-products. The universal primer PCR amplification step is only for completely synthesized pre-products. Universal primer is identical to the universal part of APEX-2 primers and thus can bind to synthesized complementary sequences of APEX-2 primers.

[0031] According to another preferred embodiment (Figure 6) the method uses two specific APEX-2 primers for detection of two studied nucleotides in close proximity to each other. In this case more than one studied nucleotides (2-25) are amplified in the first phase primer extension depending on the distance between the studied nucleotides. The principles of second- and third phase are same for the embodiment described above and this embodiment.

[0032] Each APEX-2 primer consists of two specific sequence regions: 3' complementary sequence for template nucleic acid (genomic DNA, cDNA,

RNA), and a universal sequence (for example SEQ ID NO:1) for large-scale amplification and an optional modification in the 5' end (attached via linker, for example Amine-C6) (Figure 7) to enable APEX-2 primer immobilization on solid surface as microarray (glass, plastic, beads, and paper). Therefore, determination of one studied nucleotide or determination of two studied nucleotides locating close to each other requires two APEX-2 primers and one universal sequence in amplification and in detection on microarray.

- [0033] The APEX-2 primers according to this invention are designed in the way that the 3' end of the primer terminates at the nucleotide before the studied nucleotide. As according to one preferred embodiment (Figure 8), only one nucleotide from target nucleic acid is being amplified after the first and the second phase on APEX-2 reaction. According to another preferred embodiment (Figure 6) the APEX-2 primers are designed so that APEX-2 primer 1 for coding DNA strand terminates at the nucleotide before the studied *allele 2* position and APEX-2 primer 2 for complementary DNA strand terminates at the nucleotide before the studied *allele 1* position.
- [0034] In the third phase (detection) the studied nucleotides are detected through single base extension on an microarray where the primers have same design as the primers of the first phase. In other words detection on microarray uses the same APEX-2 oligonucleotides (attached to solid phase) as used in APEX-2 amplification (in liquid phase).
- [0035] Primer extension in the first APEX-2 phase and single-base extension in detection phase enable to detect di, -tri or quatro allelic mutations or SNPs. APEX-2 primer ends one nucleotide before the studied position and all four nucleotides (A, T, C, G) which are complementary to template nucleic acid are possible in primer extension. In the detection step, APEX-2 uses four labeled nucleotides (A, U, C, G) as in published APEX method KURG, A., et al. Arrayed primer extension: solid-phase four-color DNA resequencing and mutation detection technology. *Genet Test.* 2000, vol.4, no.1, p.1-7.
- [0036] APEX-2 method is suitable for SNP- as well as for single nucleotide mutation (i.e. one nucleotide substitutions, insertions, deletions) analysis.

APEX-2 method also enables detection of 2-25 nucleotide long deletions or insertions by using APEX-2 primers which are designed in a way that the primer 3' ends terminates one nucleotide before the studied region and the first nucleotide of the insertion sequence or one nucleotide substitution is detected (Figure 9A). In case of a deleted region or single nucleotide deletion, APEX-2 primer 3' ends terminates one nucleotide before the studied deletion and the first nucleotide of the opposite APEX-2 primer is detected (deletion) (Figure 9B). As comparison a wild type sequence is used, where no deletion has taken place. In this case the first nucleotides of both ends of the deletion region is detected (Fig. 9A).

[0037] APEX-2 method differs from Affymetrix MIP probe and Illumina *GoldenGate* techniques as follows:

- a) ligation reaction is not needed;
- b) highly multiplex PCR (in principle 15 to 1500-plex) is used;
- c) APEX-2 primers are used for PCR and for single base extension on the microarray;
- d) index sequences (*tag* and *ctag*) are not needed.

[0038] Another difference between the method according to this invention and any prior art disclosures is the primer-target architecture – the APEX-2 primer ends one nucleotide before the studied nucleotide in first and second phase amplification and also in detection phase on microarray (single base extension). Thus, no allele-specific primers are needed for studied nucleotide detection and only two APEX-2 primer sequences are essential for amplification and detection of both the alleles of the nucleic acid.

[0039] APEX-2 primers used in the first phase of APEX-2 and detection step have a modification at their 5' end (attached via linker, for example Amine-C6), which enables to spot the same primers (or primer) on a microarray and to detect the appropriate studied nucleotide by single base extension with fluorescently labeled ddNTP-s.

EXAMPLE 1

[0040] The first example describes amplification of the SNPs or mutations under study (Figure 10).

[0041] To amplify all SNPs in highly multiplex PCR reaction, genomic DNA was diluted in TE buffer (10 mM Tris (pH 8.0), 1 mM EDTA) to 100 ng/μl, denaturized for 5 minutes at 98°C and cooled down to room temperature just before PCR. Multiplex PCR was carried out in 20 μl volume that contained PCR buffer (60 mM Tris-HCl (pH 8.3), 60 mM KCl, 15 mM (NH₄)₂SO₄) (Fermentas), 0.2 mM each desoxynucleotide (G, C, A, T), 5 mM MgCl₂, 2 U TrueStart DNA polymerase (Fermentas), 30 nM each specific primer (Metabion) and 200 ng genomic DNA. Twenty-four cycles of PCR were run on GeneAmp PCR System 9700 (Applied Biosystems) thermocycler: initial denaturation and TrueStart DNA polymerase activation 5 minutes at 98°C, denaturation at 95°C for 30 sec, annealing at 56°C for 20 sec, extension at 72°C for 20 sec and final extension for 5 minutes at 72°C. Time between annealing and extension step is 6.5 minutes (ensured by 3% ramp speed). First, PCR products created using specific primers are the templates for universal primer amplification in phase two. Universal primer amplification is carried on in 150 μl volume that includes 20 μl of first phase PCR product. Mixture (130 μl) contains PCR buffer (80 mM Tris (pH 9.5) 20 mM (NH₄)₂SO₄, 0.2% w/v Tween-20) (Solis Biodyne), 3.5 mM each nucleotide, 3.5 mM MgCl₂, 15 HotFire DNA polymerase (Solis Biodyne) and 40 μM universal primer (gatcaggcgtctgtcgtgctc) (SEQ ID NO:1). Initial enzyme activation and denaturation last for 15 minutes at 95°C, denaturation at 95°C for 30 sec, annealing at 54°C for 30 sec, extension at 72°C for 5 sec and after twenty cycles final extension for five minutes at 72°C. As a check, 1 μl of PCR product is visualized on 2.5% TBE agarose gel to confirm the size range of amplicons (Figure 9). PCR products from universal primer amplification are purified using the MinElute Purification kit (Qiagen) under modified protocol. 600 μl binding buffer is mixed with 150 μl of PCR product. 5 μl 5.4 M (pH 5.0) of sodium-acetate is added to guarantee the optimum pH for the spin column. Products are eluted with 23 μl EB buffer (Qiagen).

EXAMPLE 2

[0042] This second example describes the single base extension on microarray (Figure 10, detection and analyze)

[0043] Alkaline phosphatase (SAP) treatment is essential before primer extension with dideoxynucleotides to eliminate active desoxynucleotides in this reaction. 3 µl of 10X ThermoSequenase™ (Amersham Bioscience) reaction buffer and 0.5 U SAP (Fermentas) is added to purify the PCR product. Mixture is incubated for 15 minutes at 37°C and SAP is inactivated during heating for 10 minutes at 95°C. This incubation is also as template denaturation phase before primer extension on array. After 10 minutes of denaturation at 95°C 1.25 µM dideoxynucleotide mixture (Cy3-ddATP; Cy5-ddGTP; Texas Red®-ddCTP and Fluoresceine-12-ddUTP (Perkin Elmer Life Science)), 5 U of ThermoSequenase™ DNA polymerase (Amersham Biosciences) and 3 mM MgCl₂ in the final volume of 30 µl are added. Mixture is applied to the pre-warmed arrays on a heat-plate and the arrays are covered with LifterSlip™ ((22 x 25 mm) Erie Science Company). The hybridization and APEX reaction is performed at 58°C for 20 minutes and terminated by washing at 95°C for 1 min in Milli-Q water, followed by washing for 3 min in 0.3% Alconox® solution (Alconox). Alconox is removed by washing the arrays two times for 1 min with 95°C Milli-Q water. To reduce bleaching, 12 µl SlowFade® antifade reagent (Invitrogen) is applied to the slide. The arrays are scanned with the Genorama® imaging system (Asper Biotech) at 20 µm resolution.

References

[0044]

- MATSUZAKI, H., et al. Genotyping over 100,000 SNPs on a pair of oligonucleotide arrays. *Nat. Methods*. 2004, no.1, p.109–111.
- MATSUZAKI, H., et al. Parallel genotyping of over 10,000 SNPs using a one-primer assay on a high-density oligonucleotide array. *Genome Res*. 2004, no.14, p.414-425.
- BARRETT, J. C., et al. Evaluating coverage of genome-wide association studies. *Nat Genet*. June 2006, vol.38, no.6, p.659-62.
- NILSSON, M., et al. Padlock probes: circularizing oligonucleotides for localized DNA detection.. *Science*. 1994, no.265, p.2085–2088.
- HARDENBOL, P., et al. Multiplexed genotyping with sequence-tagged molecular inversion probes. *Nat. Biotechnol*. 2003, no.21, p.673–678.

- **HARDENBOL, P., et al.** Highly multiplexed molecular inversion probe genotyping: Over 10,000 targeted SNPs genotyped in a single tube assay. *Genome Res.* 2005, no.15, p.269–275.
- **GUNDERSON, K. L., et al.** A genome-wide scalable SNP genotyping assay using microarray technology. *Nat. Genet.* 2005, no.37, p.549–554.
- **STEEMERS, F. J., et al.** Whole-genome genotyping with the single-base extension assay. *Nat Methods.* Januar 2006, vol.3, no.1, p.31-3.
- **KURG, A., et al.** Arrayed primer extension: solid-phase four-color DNA resequencing and mutation detection technology. *Genet Test.* 2000, vol.4, no.1, p.1-7.

Claims

1. A method to determine single nucleotide polymorphism (SNP) in a target nucleic acid sequence, said method comprising the steps of:
 - a) providing a template nucleic acid, consisting of a first and a second strand and having one studied nucleotide pair;
 - b) designing a first and a second APEX-2 oligonucleotide primer;

said first APEX-2 primer consisting of a first specific oligonucleotide sequence in its 3'-end and a universal sequence in its 5'end, said universal sequence further having a modification in its 5'end , and said first specific oligonucleotide sequence being complementary to nucleotides before the unknown nucleotide in the first strand of the template nucleotide acid;

said second APEX-2 primer consisting of a second specific oligonucleotide sequence in its 3'-end and a universal sequence in its 5'end, said universal sequence optionally having a modification in its 5'end and, said second specific oligonucleotide sequence being complementary to nucleotides before the unknown nucleotide in the second strand of the template nucleotide acid;
 - c) amplifying the unknown nucleotide of the first and the second strand of the template nucleotide acid by running a multiplex PCR primer extension reaction with the first and the second APEX-2 primers as templates, thereby producing a first amplification product, said first amplification product consisting of a first and a second strand, said first strand consisting of sequence of the first APEX-2 primer and a complementary sequence to the second APEX-2 primer and the unknown nucleotide between them, and the second strand consisting of complementary sequence to the APEX-2 first primer and the sequence of the second APEX-2 primer and the unknown nucleotide between them;
 - d) amplifying the first amplification product in a PCR reaction with universal primer as template, thereby producing a second amplification product;
 - e) providing a microarray of probes, where probes are identical to the first and the second APEX-2 primers of step b) and where the primers are immobilized on the array by attaching them to a solid surface from the modified 5'end;
 - f) letting the second amplification products anneal with immobilized primers and providing labeled or modified terminator nucleotides for single base extension reaction; and

- g) identifying terminating nucleotide that has been added to the probes, said terminating nucleotides corresponding to the unknown nucleotides of the template nucleic acid.
2. The method according to claim 1, wherein the template nucleic acid has a first and a second pair of unknown nucleotides in proximity of 2 to 25 bp to each other and wherein the first specific oligonucleotide sequence is complementary to nucleotides before unknown nucleotide of the first nucleotide pair on the first nucleic acid strand and the second specific oligonucleotide sequence is complementary to nucleotides before unknown nucleotide of the second nucleotide pair of the second nucleic acid strand.
 3. The method according to claim 1 or 2, wherein the universal sequence of both the first and the second APEX-2 primers have a modification in its 5' end.
 4. The method according to claim 1 or 2, wherein only two cycles of amplification are used to produce the first amplification product in step c).
 5. The method according to claim 1 or 2, wherein the modifications in the 5' end of the first specific primer and/or the second specific primer are attached with a linker.
 6. The method according to claim 5, wherein the linker is Amine-C6.
 7. The method according to claim 1 or 2, wherein the template nucleic acid is selected from a group consisting of genomic DNA, cDNA and RNA.
 8. The method according to claim 1 or 2, wherein the solid surface is made of a material selected from the group consisting of glass, plastic, beads, and paper.
 9. The method according to claim 1 or 2, wherein the studied nucleotide pair is di,-tri or quatro allelic mutation or SNP.
 10. The method according to claim 1 or 2, wherein the unknown nucleotide pair is an insertion and the identifying terminating nucleotide of step g) corresponds to first nucleotide of insertion sequence.
 11. The method according to claim 1 or 2, wherein the unknown nucleotide pair is a deletion and the identifying terminating nucleotide of step g) corresponds to first nucleotide of opposite APEX-2 primer.
 12. The method according to claim 1 or 2, wherein the universal sequence is according to SEQ ID NO:1.

1/10

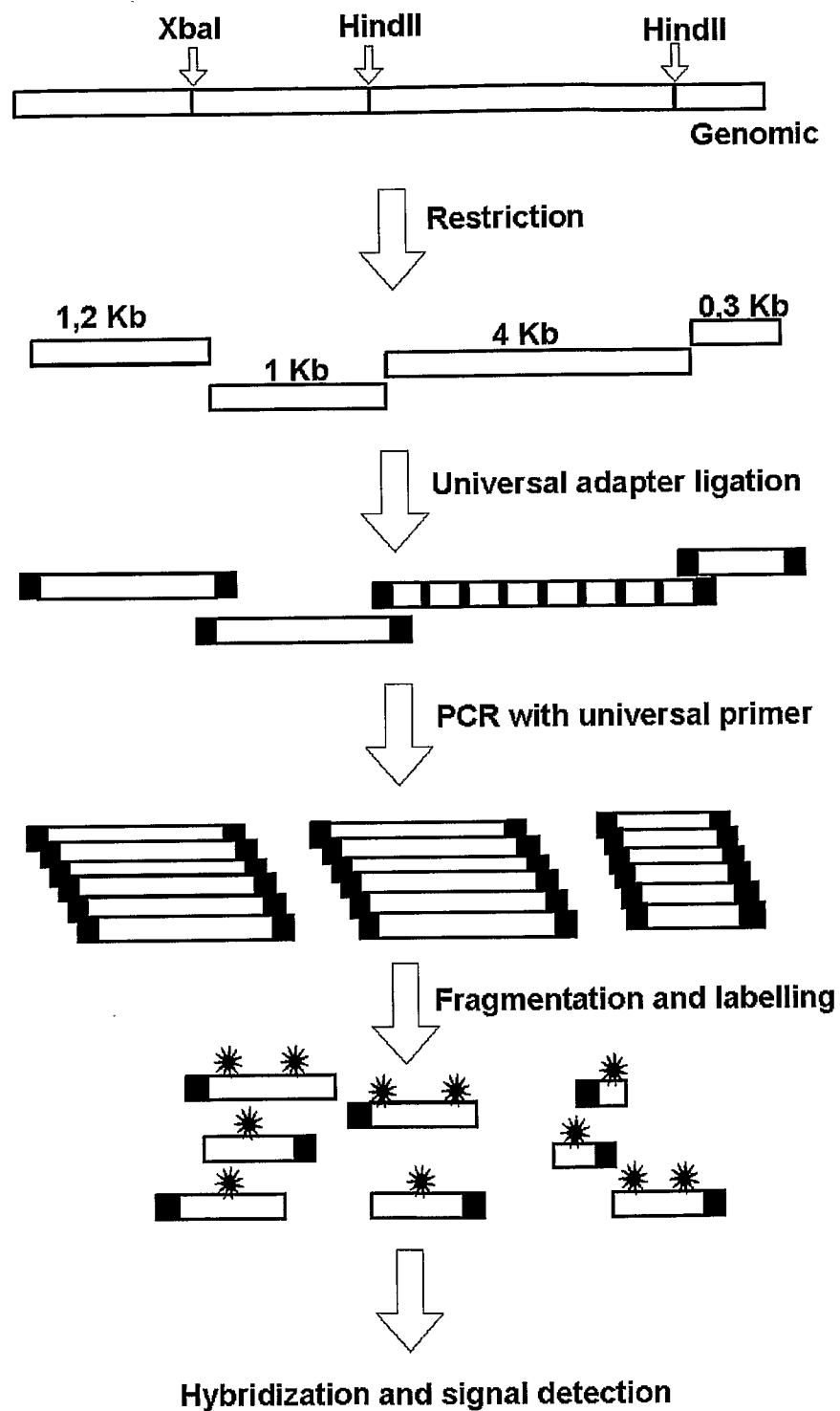


Fig. 1

2/10

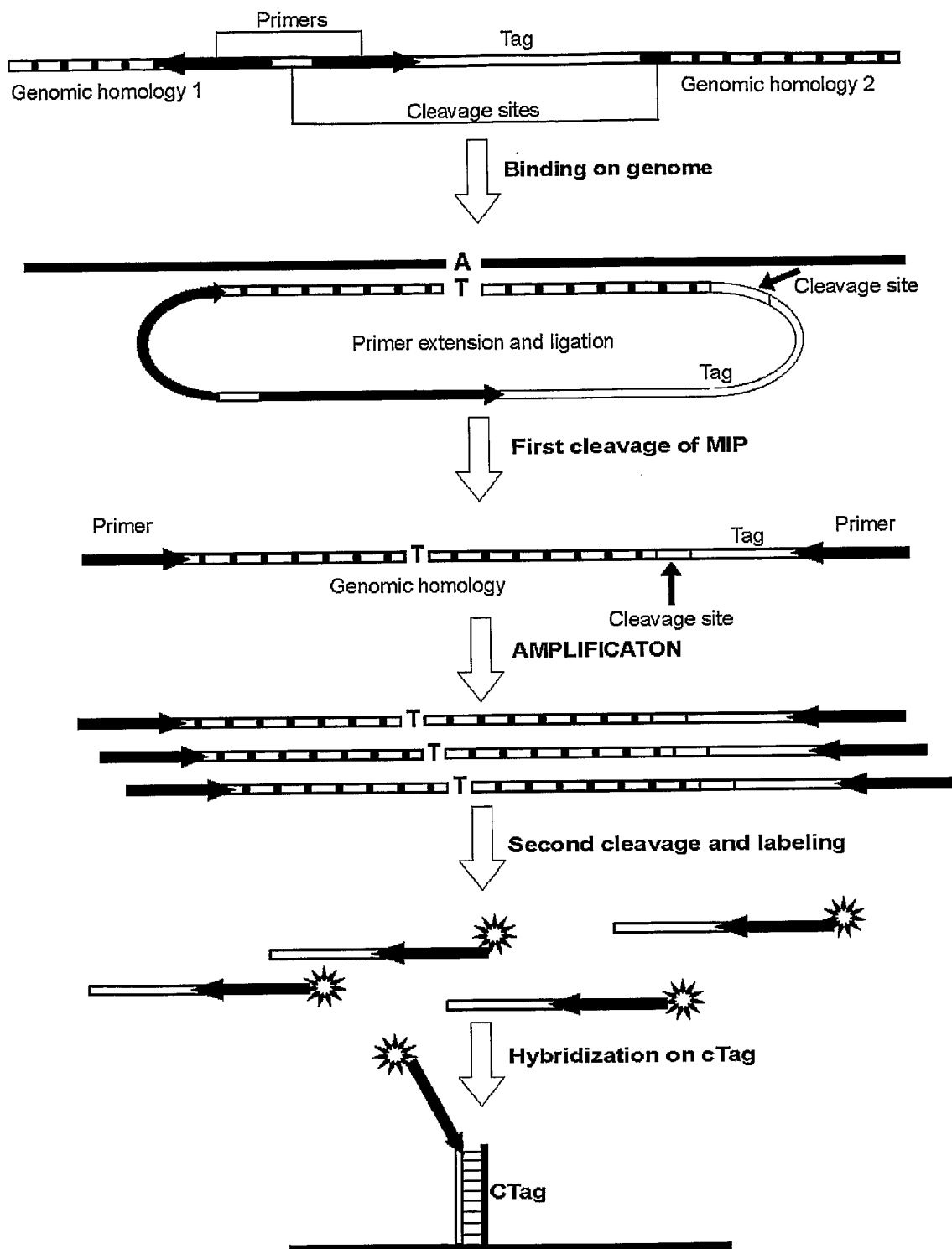


Fig. 2

3/10

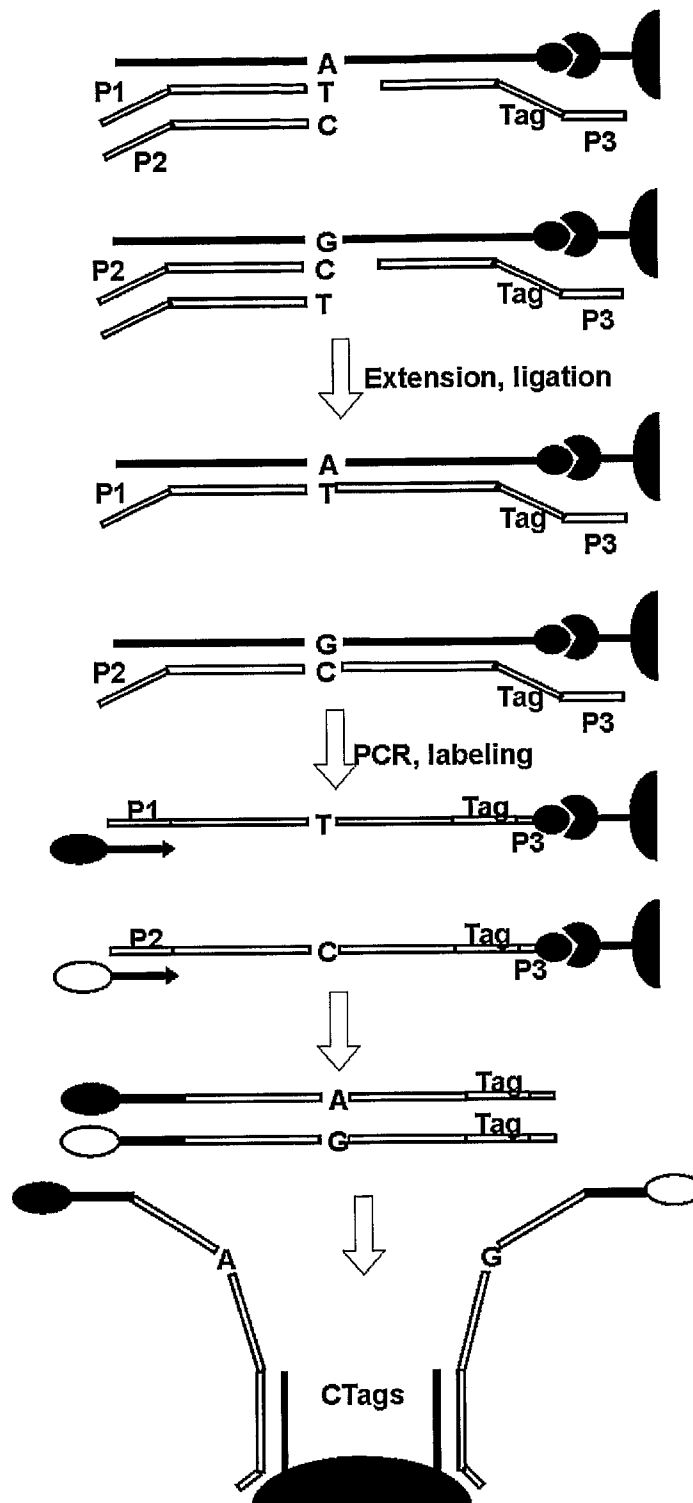
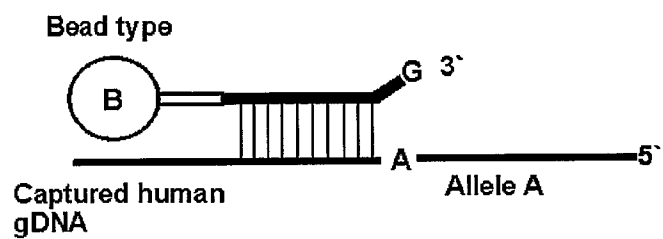
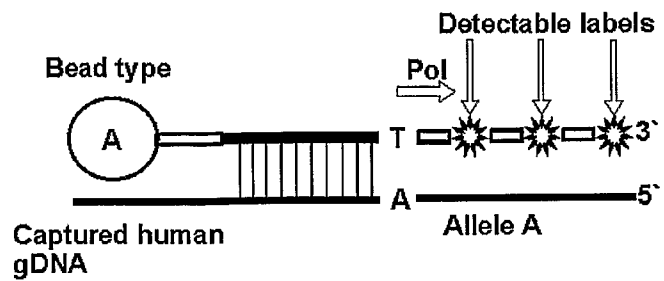


Fig. 3

4/10

Infinium I
Allele-Specific Primer Extension



Infinium II
Single Base Extension

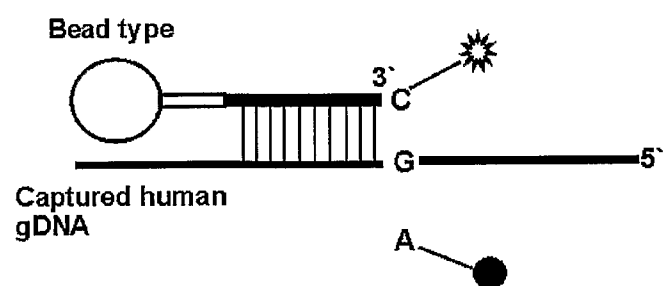


Fig. 4

5/10

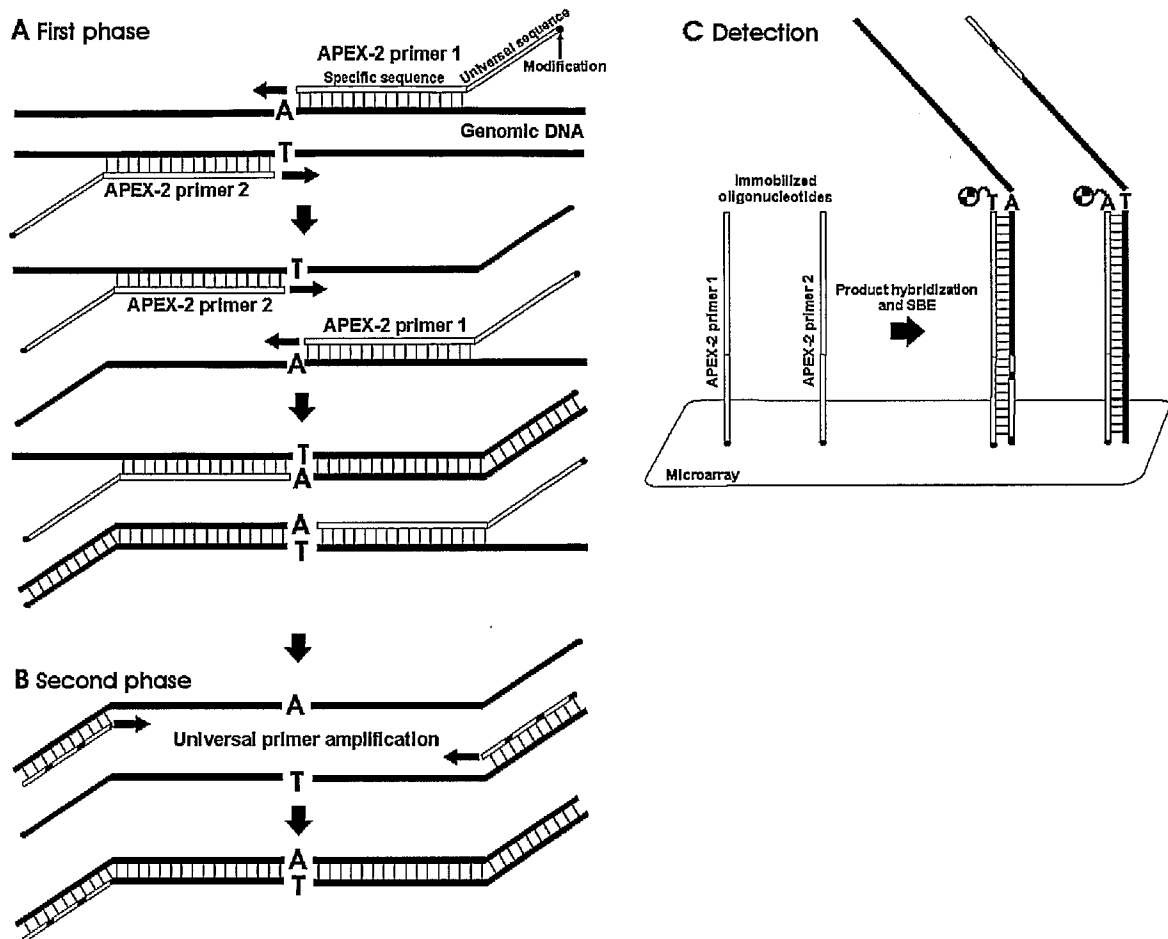
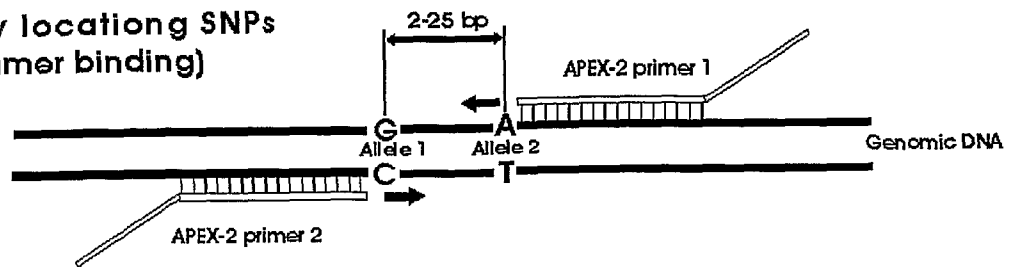
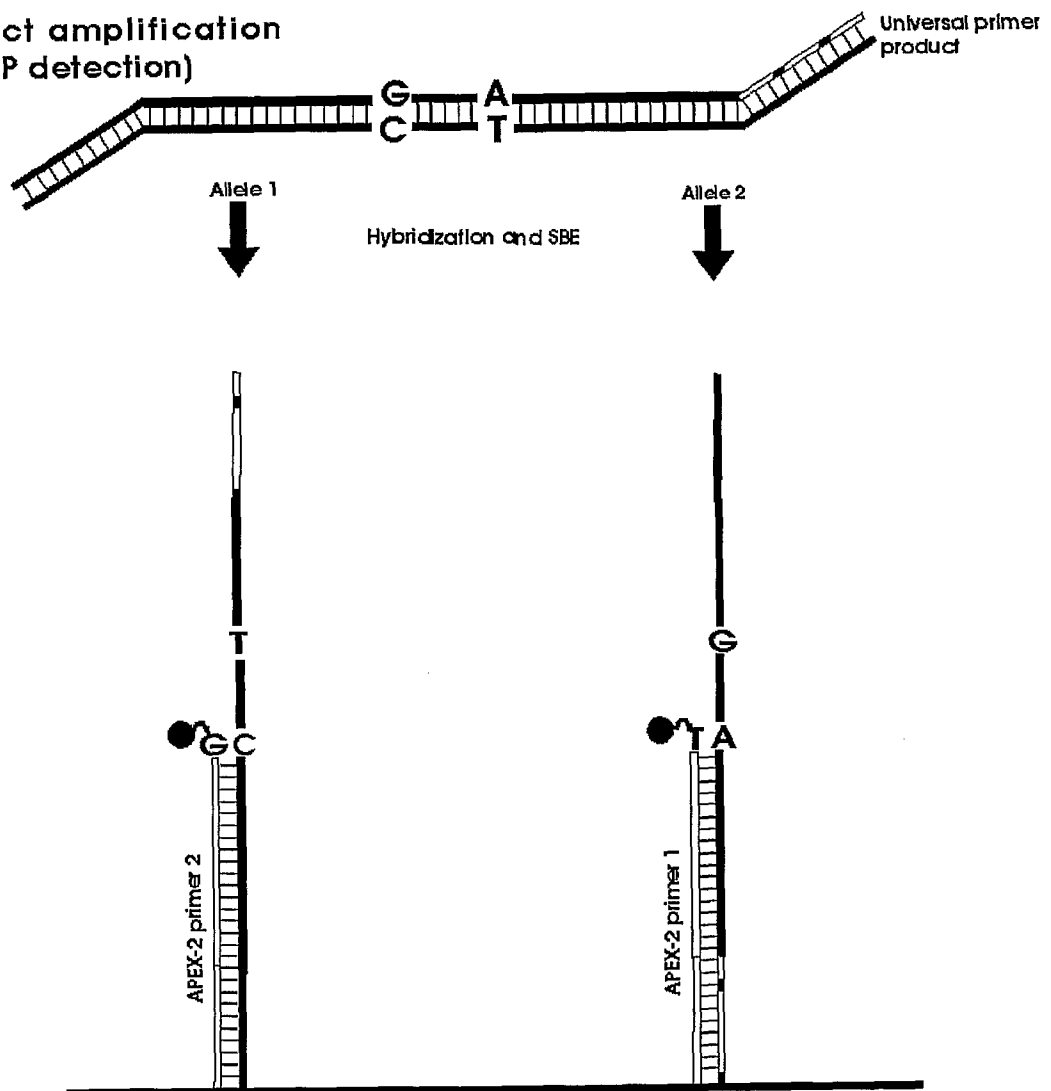


Fig. 5

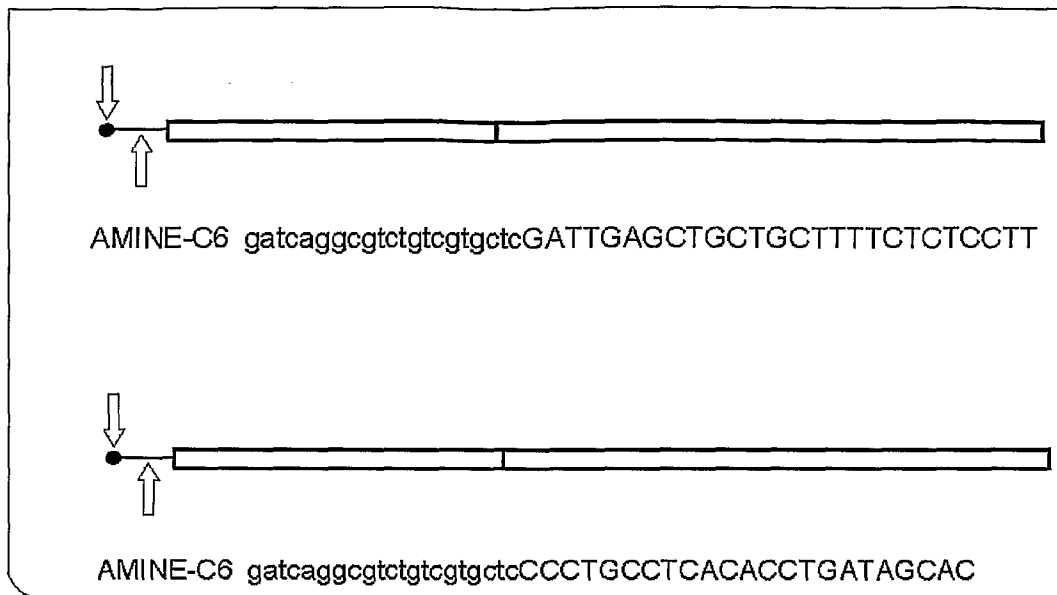
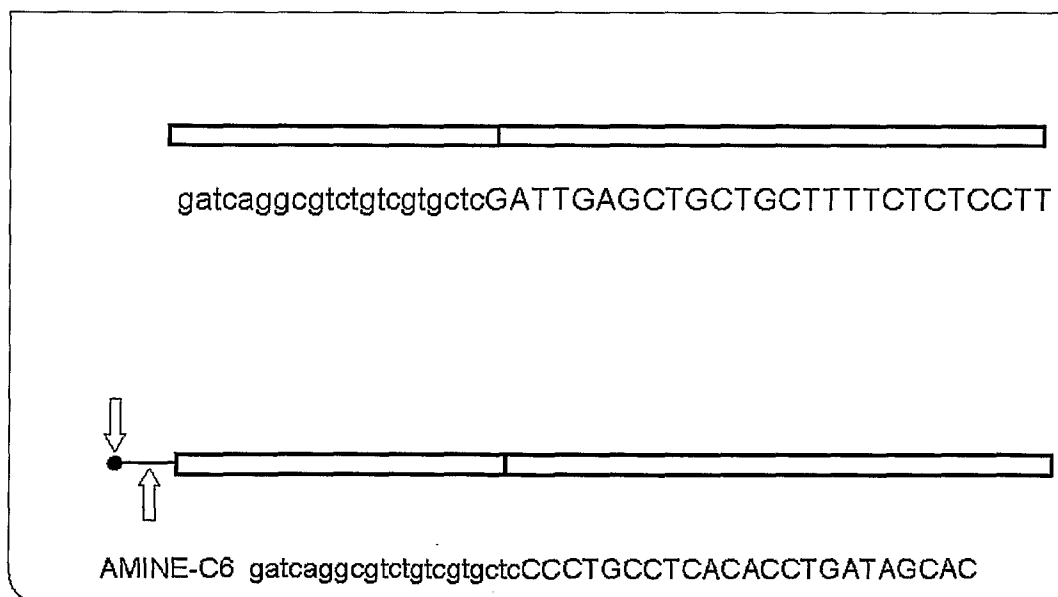
6/10

A (Nearly locating SNPs and primer binding)

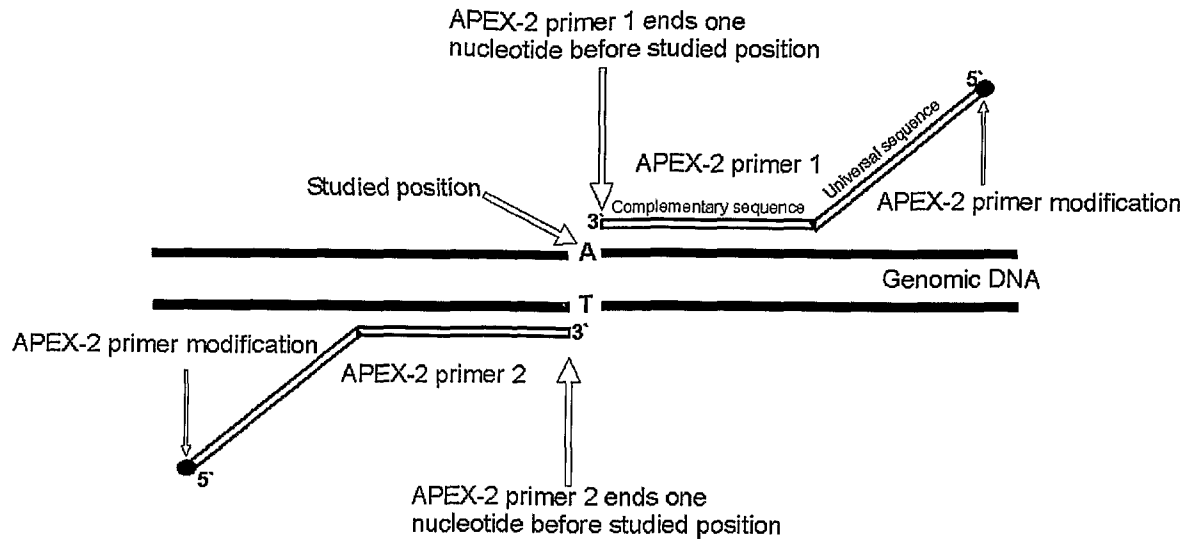
↓ APEX-2 amplification

B (Product amplification and SNP detection)**Fig. 6**

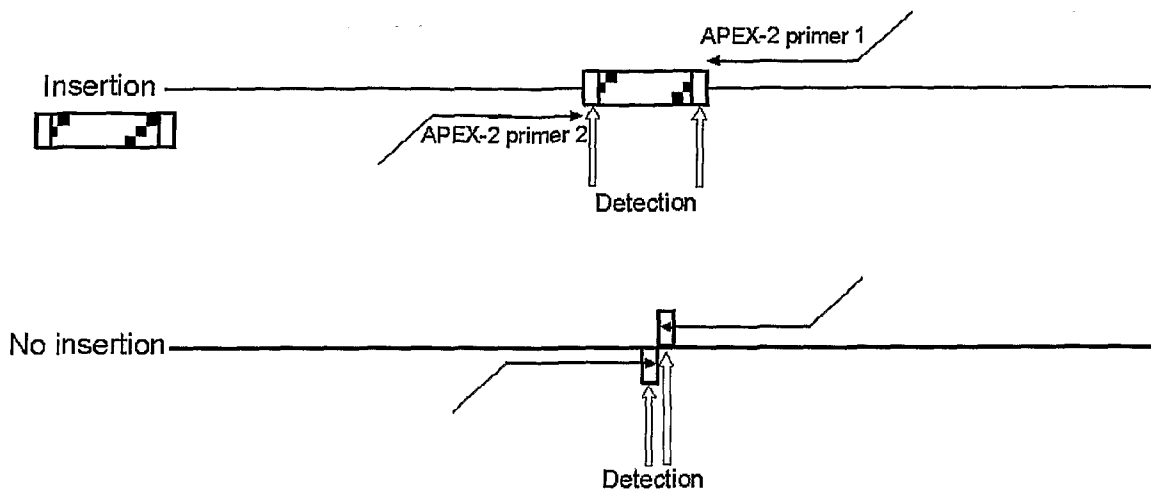
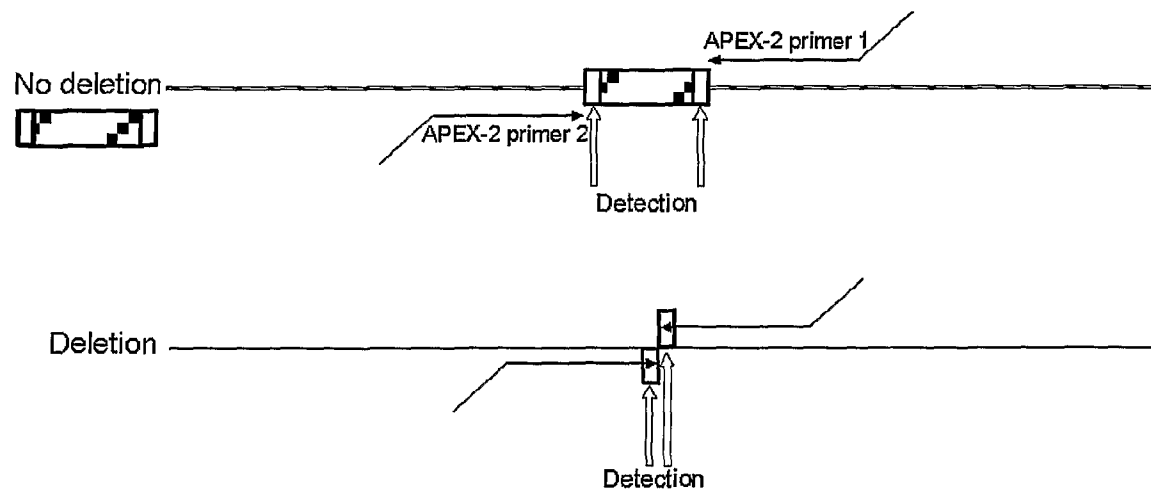
7/10

APEX-2 primer pair with two modified oligonucleotides for immobilization**APEX-2 primer pair with one modified oligonucleotide for immobilization****Fig. 7**

8/10

**Fig. 8**

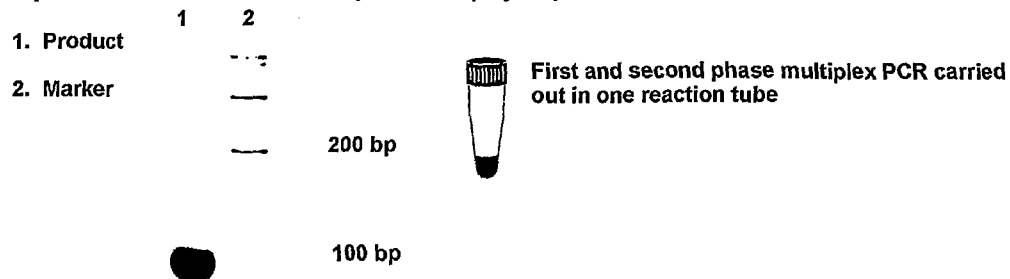
9/10

A Insertions**B Deletions****Fig. 9**

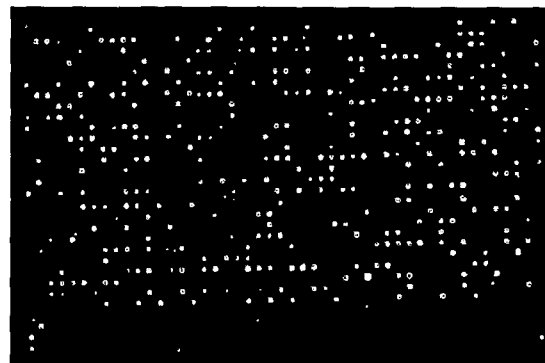
10/10

First phase APEX-2 primers anneal to template DNA and first round product is generated

Second phase PCR with universal primers amplify all products from the first round. Product control on gel.



Third phase APEX-2 amplification products are purified and hybridized to microarray. Single base extension adds one labeled terminatrnucleotide to immobilized primers and fluorescence detected.



621 genetic markers on microarray

Detection of signals

Analyze Software converts detected array signals to database and compare all samples

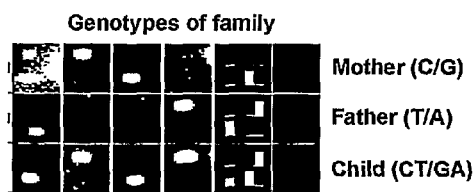


Fig. 10

INTERNATIONAL SEARCH REPORT

International application No
PCT/EE2007/000003

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EP0-Internal, BIOSIS, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SHUMAKER J M ET AL: "APEX disease gene resequencing: Mutations in exon 7 of the p53 tumor suppressor gene" BIOORGANIC AND MEDICINAL CHEMISTRY 2001 UNITED KINGDOM, vol. 9, no. 9, 2001, pages 2269-2278, XP002444656 ISSN: 0968-0896 abstract	1-12
A	US 6 448 010 B1 (ZHAO XIAODONG [US]) 10 September 2002 (2002-09-10) abstract	1-12

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *&* document member of the same patent family

Date of the actual completion of the international search

31 July 2007

Date of mailing of the international search report

09/08/2007

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Costa Roldán, Nuria

INTERNATIONAL SEARCH REPORT

International application No

PCT/EE2007/000003

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KURG A ET AL: "ARRAYED PRIMER EXTENSION: SOLID-PHASE FOUR-COLOR DNA RESEQUENCING AND MUTATION DETECTION TECHNOLOGY" GENETIC TESTING, LARCHMONT, NY, US, vol. 4, no. 1, 21 March 2000 (2000-03-21), pages 1-7, XP009003695 ISSN: 1090-6576 the whole document	1-12
A	TEBBUTT SCOTT J ET AL: "Microarray genotyping resource to determine population stratification in genetic association studies of complex disease" BIOTECHNIQUES, vol. 37, no. 6, December 2004 (2004-12), pages 977-985, XP001536623 ISSN: 0736-6205	1-12
A	TEBBUTT SCOTT J ET AL: "SNP Chart: an integrated platform for visualization and interpretation of microarray genotyping data" BIOINFORMATICS (OXFORD), vol. 21, no. 1, 1 January 2005 (2005-01-01), pages 124-127, XP002444657 ISSN: 1367-4803 abstract	1-12
A	PODDER MOHUA ET AL: "Dynamic variable selection in SNP genotype autocalling from APEX microarray data." BMC BIOINFORMATICS 2006, vol. 7, 2006, page 521, XP002444658 ISSN: 1471-2105 abstract	1-12

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EE2007/000003

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 6448010	B1	10-09-2002	NONE