(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 November 2010 (04.11.2010)

PCT

(10) International Publication Number
WO 2010/125351 A1

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

(21) International Application Number:
PCT/GB2010/000861

(22) International Filing Date:
29 April 2010 (29.04.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0907412.1 29 April 2009 (29.04.2009) GB

(71) Applicant *(for all designated States except US)*: **THE ANTHONY NOLAN TRUST** [GB/GB]; The Royal Free Hospital, Pond Street, Hampstead, London NW3 2QG (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **MADRIGAL, Jose, Alejandro** [MX/GB]; The Anthony Nolan Research Institute, The Royal Free Hospital, Pond Street, Hampstead, London NW3 2QG (GB). **MARSH, Steven, George, Edward** [GB/GB]; The Anthony Nolan Research Institute, The Royal Free Hospital, Pond Street, Hampstead, London NW3 2QG (GB). **MAYOR, Neema, Philippa** [GB/GB]; The Anthony Nolan Research Institute, The Royal Free Hospital, Pond Street, Hampstead, London NW3 2QG (GB). **ROBERTS, Christopher, Henry** [GB/GB]; The Anthony Nolan Research Institute, The Royal Free Hospital, Pond Street, Hampstead, London NW3 2QG (GB).

(74) Agent: **CAMPBELL, Patrick, John, Henry**; J.A. Kemp & Co., 14 South Square, Gray's Inn, London WC1R 5JJ (GB).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, 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, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, 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, LR, 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, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: SIMULTANEOUS SEQUENCING OF MULTIPLE NUCLEIC ACID SEQUENCES

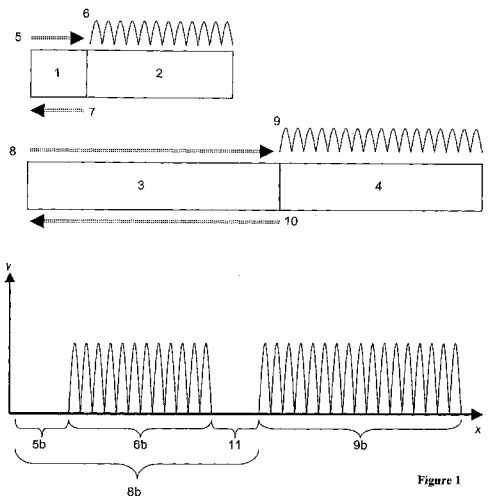


Figure 1

(57) Abstract: The invention provides a method and a kit for the simultaneous sequencing of multiple nucleic acid sequences. The method is called "Short Template Amplicon and Multiplex Megaprimer Enabled Relay" (STAMMER) sequencing. The method provides a platform that is superior to pre-existing methods for simultaneous sequencing, including the bidirectional sequencing of double- stranded nucleic acid templates. In its broadest aspect, the invention provides a method for simultaneously sequencing a plurality of nucleic acid target sites, comprising (a) annealing each of a plurality of different-sized primers to the 5 '- end of each of said plurality of nucleic acid target sites; (b) performing a sequencing reaction by simultaneously extending said primers in the presence of labelled terminators and thereby obtaining a pool of extended primers in a single reaction vessel; and (c) determining the sequences of the nucleic acid target sites by analysing said pool of extended primers without separating the pool prior to analysis; wherein the largest extension product of each primer is smaller in length than the next largest, non-extended primer, and wherein at least one of the primers is a megaprimer comprising double-stranded nucleic acid of at least 100 base pairs in length.

WO 2010/125351 A1

Published:

— *with international search report (Art. 21(3))*

SIMULTANEOUS SEQUENCING OF MULTIPLE NUCLEIC ACID SEQUENCES

FIELD OF THE INVENTION

The invention relates to the simultaneous sequencing of multiple nucleic acid sequences.

BACKGROUND OF THE INVENTION

Simultaneous sequencing is the process of performing nucleic acid sequencing reactions in such a way that the sequences of a plurality of different templates are detected in multiplex, requiring the use of only a single reaction vessel and sequencing capillary. The term 'unidirectional sequencing' refers herein to the sequencing of a single strand of a polynucleotide template and may be performed on either the sense or anti-sense strand of a template that is either single or double-stranded. The term 'bidirectional sequencing' refers meanwhile to the sequencing of complementary sense and anti-sense strands of a double-stranded nucleic acid template.

The 'Sanger' method of chain-terminator sequencing[1] is the industry standard for nucleic acid sequence analysis, and a well-optimised sequencing reaction is able to produce around 1500 bases of continuous sequence data. Many sequencing requirements cannot take full advantage of this potential because the sequence of interest may be much shorter than the maximum possible and because it is often necessary to sequence multiple regions of the genome that are not in linkage disequilibrium. When this is the case, it has traditionally been necessary to perform numerous separate sequencing reactions, leading to a hyperbolic increase in the per sample cost of sequencing and a proportional decrease in sample throughput.

There have been several previous attempts to perform 'simultaneous sequencing' wherein the sequencing of several regions of interest is performed in a single sequencing reaction and then detected 'on-line' in a single capillary of the DNA analysis platform. These methods have been limited in their application by their various dependencies upon the use of non-classical fluorochromes[2,3], upon the physical segregation of the various sets of sequencing products before electrophoresis and detection in separate capillaries[4] or upon the use of modified sequence analysis platforms[5].

Superior to these approaches was 'SimulSeq', a technique that allowed simultaneous sequence detection from several template Polymerase Chain Reaction (PCR) products but that required the use of only 'off-the-shelf' sequencing reagents and analysis platforms[6]. This technique (referring to WO/2003/056030) was strictly limited to the simultaneous detection of only a very small (~ 150 bp) total number of bases; allowing for say three multiplexed templates of 50 bp or of five multiplexed templates of 30 bp, or ten multiplexed templates of 15 bp and so on. SimulSeq could also be used to perform simultaneous bi-directional sequencing, where both strands of an individual double-stranded template were detected in multiplex; but this was similarly limited to templates of length shorter than 150 bp. The limitations of SimulSeq are due to its reliance on commercially synthesised oligonucleotide primers and the fact that these cannot be prepared at lengths exceeding the 150 bp – 200 bp limit of current oligonucleic acid synthesis techniques. Until now, no simultaneous sequencing or bidirectional sequencing method has allowed a truly flexible approach that can be applied to any group of templates where there would be a practical benefit to performing this type of sequencing assay and where the combined lengths may stretch into the hundreds or early thousands of bases.

Automated 'Sanger' sequencing reactions are enabled by the extension of oligonucleotide primers[7] and these reactions produce nested, chain-terminated[1] and fluorescently labelled[8] sequence fragments of every possible length greater than that of the oligonucleotide primer and lesser than or equal to the length of the

template nucleic acid. The earliest fluorescence signal detected by the DNA analysis system during electrophoresis is from those fragments where the first primer extension step incorporates a chain terminator base. Increasing the length of the oligonucleotide primer can delay the onset of this first signal because the rate of the electrophoretic migration of the sequence fragments is proportional to their masses. Molecular typing assays based on single nucleotide extensions have taken advantage of this behaviour to great effect by using oligonucleotides of different sizes to differentially delay the electrophoresis and detection of sometimes several dozen single nucleotide primer extension products relating to known polymorphisms in a multiplex reaction[9-11].

SUMMARY OF THE INVENTION

The invention provides a method and a kit for the simultaneous sequencing of multiple nucleic acid sequences. The method is called "Short Template Amplicon and Multiplex Megaprimer Enabled Relay" (STAMMER) sequencing. The method provides a platform that is superior to pre-existing methods for simultaneous sequencing, including the bidirectional sequencing of double-stranded nucleic acid templates.

In its broadest aspect, the invention provides a method for simultaneously sequencing a plurality of nucleic acid target sites, comprising (a) annealing each of a plurality of different-sized primers to the 5′-end of each of said plurality of nucleic acid target sites; (b) performing a sequencing reaction by simultaneously extending said primers in the presence of labelled terminators and thereby obtaining a pool of extended primers in a single reaction vessel; and (c) determining the sequences of the nucleic acid target sites by analysing said pool of extended primers without separating the pool prior to analysis; wherein the largest extension product of each primer is smaller in length than the next largest, non-extended primer, and wherein at least one of the primers is a megaprimer comprising double-stranded nucleic acid of at least 100 base pairs in length.

The invention also provides a kit for performing the method of the invention, wherein the kit comprises (a) at least one megaprimer comprising double-stranded nucleic acid of at least 100 base pairs in length and (b) labelled chain terminators.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows how the electrophoretic migration of labelled nucleic acid sequencing fragments can be slowed in order to delay the onset of detectable signal until after a previous sequence has fully migrated past the optical detector. (1) is the 130 bp Megaprimer Target Region of a Short Template Amplicon of 260 bp of which (2) is the region of interest. (3) is the 276 bp Megaprimer Target Region of a second Short Template Amplicon of 410 bp of which (4) is the region of interest. (5) is the forward strand of a 130 bp megaprimer specific to (1) and (5b) is a silent period during which no sequence detection occurs during electrophoresis. The duration of (5b) is proportional to the molecular weight of the megaprimer (5). (6) shows how megaprimer (5) is extended from the 3′ end during the sequencing reaction and (6b) is the detectable sequence relating to this. (7) is the reverse strand of megaprimer (5). Unlike (5), this strand cannot be extended during the sequencing reaction, as there is no downstream template. (8) shows the forward strand of a 276 bp megaprimer specific to (3) and (8b) is the silent period of electrophoresis, during which no sequencing products relating to the extension of the megaprimer (8) are detected. (9) shows how megaprimer (8) is extended from the 3′ end during the sequencing reaction and (9b) is the detectable sequence relating to this. (10) is the reverse strand of megaprimer (8). Unlike (5) and (8), this strand cannot be extended during the sequencing reaction as there is no downstream template. (11) is a period of silence wherein all products of (6) have passed the detector but where the products of (9) have yet to be detectable. The lower part of Figure 1 shows how the electrophoresis chromatogram would be expected to look in the results of a STAMMER sequencing reaction based upon the model in the upper part of Figure 1.

Figure 2A is the electrophoresis chromatogram of the short template amplicon (1-2) sequence derived from the forward strand (5) of the dsDNA megaprimer (5 and 7) and corresponds to the sequence described in Figure 1 as (6b). Figure **2B** is the electrophoresis chromatogram of the short template amplicon (3-4) sequence derived from the forward strand (8) of the dsDNA megaprimer (8 and 10) and corresponds to the sequence described in Figure 1 as (9b). Figure **2C** is the electrophoresis chromatogram of a STAMMER multiplex reaction in which both of the sequencing reactions described in Figures 2A and 2B were carried out in multiplex according to the model described in Figure 1. Figure **2D** shows a portion of the fully analysed and base-called nucleotide sequence derived from the electrophoresis chromatogram pictured in 2C.

Sequence 6b is the NOD2/CARD15 exon 11 short template amplicon and sequence 9b is the NOD2/CARD15 exon 4 short template amplicon. They were prepared according to the primers described in table 1 and amplification and sequencing parameters were identical to those used with regards to the experiments described in figures 3 and 4 (see Example 1). Exon 11 (4A) and Exon 4 (4B) were detectable in the periods equivalent to size ranges 131-260 bp and 277-410 bp respectively. The quality data for the sequence shown in 9c is labelled Q and is prepared using the Phred algorithm according to Ewing, B. & Green, P., Base-calling of automated sequencer traces using Phred. II. Error probabilities. *Genome Res* **8**, 186-94 (1998) and Ewing, B., Hillier, L., Wendl, M.C. & Green, P. Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome Res* **8**, 175-85 (1998).

Figure 3 shows the STAMMER sequencing principle as applied to the NOD2/CARD15 gene. **Figure 3A** shows the assay design using three pairs of Short Template Amplicons (STA), each of which is composed of a region of interest (ROI) and a megaprimer target region (MTR) which is complementary to one of the three corresponding megaprimers (MGP). The regions of interest are spaced some 18 kbp apart in total. **Figure 3B** shows how i) priming of the MTR and extension of the megaprimer forward strand produces ii) nested and chain

terminated sequence fragments complementary to the short template amplicon and how iii) there is no potential for extension of the megaprimer reverse strand due to a lack of a suitable extensible template. **Figure 3C** shows how the detection of fluorescence signals from the three sets of nested megaprimer extension fragments would be predicted to occur on-line, in the manner of a relay.

Figure 4. In individual sequencing reactions, megaprimers were used to delay the detection of fluorescence signals relating to the three Short Template Amplicons of the NOD2/CARD15 gene. Signals relating to Exon 11 (4A) were detectable in the scanning time period equivalent to fragments in the size range of 131-260 bp. Exon 4 (4B) and Exon 8 (4C) were detectable in the periods equivalent to size ranges 277-410 bp and 438-546 bp respectively. When the STAMMER multiplex relay sequencing reaction was performed using all three Short Template Amplicons (4D), sequences were detectable in relay and were interspersed with clear intervals of silence in which no fluorescently labelled fragments were detected. Each sequence was delimited by strong adenosine peaks that were artefacts of the non-template dependent addition of ddATP residues to the megaprimer and short template amplicon homopolymers. Guides to fluorescence intensity scale have been digitally removed from these images for reasons of clarity.

Figure 5 shows a schematic of STAMMER applied to the bidirectional sequencing of a single short template amplicon. **Figure 5A** shows how the anti-sense strand of a megaprimer (MGP-AS) and a sense oligonucleotide primer (OP) respectively hybridise to the sense and anti-sense strands of a short template amplicon. The two primers extend during the cycle sequencing reaction in the direction of the solid arrows. The sense strand of the megaprimer (MGP-S) does not extend due to a lack of suitable template. A 3′-propyl-oligonucleotide 'roadblock' oligonucleotide (RB) hybridises upstream of the megaprimer sense strand. **Figure 5B** shows how a multiplex bidirectional STAMMER sequencing reaction can be performed according to this model. The megaprimer enabled sequence covers the region of interest and some irrelevant flanking sequences

whilst the oligonucleotide primer enabled sequence covers the complementary strand over the region of interest. **Figure 5C** is the base called electropherographic data output of a complete STAMMER bi-directional sequencing reaction. Sequence quality is demonstrated by the topographical detail where vertical bars above the dashed horizontal line labelled Q20, represent base calls assigned with more than 99% accuracy according to Ewing *et al.* Guides to the actual scale of fluorescence intensity have been digitally removed from these images for reasons of clarity.

Figure 6 shows the effects of a roadblock on the sequencing reaction. The raw chromatographic data in **6A** shows the effect of a roadblock oligonucleotide on the progression of the otherwise unimpeded sequencing reaction shown in **6B**. The arrow marks the position where the sequencing enzyme is halted by the obstacle of the 5′ end of the roadblock molecule.

DETAILED DESCRIPTION OF THE INVENTION

STAMMER differs from the SimulSeq method partly in that it is based upon the use of one or more double-stranded polynucleotide 'megaprimers', which are generally PCR products and are used in the enabling of the sequencing reaction in place of classical oligonucleotide primers. Megaprimers of any length amenable to PCR can be constructed, thus relieving the constraints on template length imposed on 'SimulSeq' by the limitations of commercial oligonucleic acid synthesis technology. Megaprimers of the invention are at least 100 bp in length, optionally at least 120 bp, at least 140 bp, at least 160 bp, at least 180 bp, at least 200 bp, at least 220 bp, at least 240 bp, at least 260 bp, at least 280 bp, or at least 300 bp in length. For use with conventional sequence analysis platforms, the maximum length of a megaprimer is about 1500 bp but the length of the largest megaprimer is typically less than this; for example the length of the largest megaprimer is typically up to 1000 bp, 700 bp or 500 bp. The lengths of the megaprimers can be tailored so that the electrophoretic migration rate (and therefore the detection of primer extension products relating to the megaprimer

enabled sequences) may be delayed by the equivalent of tens, hundreds or potentially thousands of bases (see Figures 1 and 2). This allows an opportunity for one or more complete oligonucleotide and/or megaprimer enabled sequence(s) of any given length shorter than the subsequent megaprimer to migrate past the point of signal detection before the onset of detectable signal from the subsequent megaprimer enabled sequence(s).

Megaprimers can also be prepared by a number of non-PCR methods, such as restriction digestion with blunt end cutting restriction endonucleases, followed by purification of the desired fragment. Similarly, restriction endonuclease digestion with enzymes producing sticky ends, followed by degradation of the sticky ends to blunt ends using appropriate enzymes (generally possessing exonuclease activity), could be used.

Megaprimers could also be prepared using non-PCR based methods of DNA amplification. These include, though are not exclusively limited to, isothermal DNA amplification methods including rolling circle amplification (RCA), real time strand displacement amplification (SDA) and multiple displacement amplification (MDA). These methods, and RCA in particular, might be used in combination with restriction endonuclease digestion.

The STAMMER sequencing technique is dependent not only upon the ability to delay the onset of sequence detection but also upon limiting the lengths of template that cause the individual primer/megaprimer extension reactions to terminate once the region of interest has been passed; thereby providing a 'silence' into which subsequent megaprimer enabled sequences may initiate. This is achieved by using site-specific quenchers of the sequencing reaction (i.e. roadblock primers) or Short Template Amplicons as templates; the latter are PCR products comprising only the informative region of interest and a 'megaprimer target region', an expanse of sequence flanking the region of interest that is equal in length and complementary to the enabling megaprimer. Each strand of the double-stranded megaprimer is capable of hybridising to its complementary strand

in the megaprimer target region, but there is no downstream template for the extension of the second megaprimer strand and the detectable sequence therefore represents only megaprimer extension across the region of interest. Megaprimers may enable the sequencing of template strands in either the sense or anti-sense context.

Short Template Amplicon and Multiplex Megaprimer Enabled Relay (STAMMER) sequencing allows simultaneous 'relay' sequencing of practically any combination of nucleic acid templates. Previous systems such as SimulSeq were limited to very short combined template lengths and were only practically applicable in molecular typing assays or in the bidirectional sequencing of very short regions of interest[6]. The STAMMER approach to simultaneous sequencing is applicable to a far wider range of templates than has been previously possible and we believe that it could impact massively upon the costs and throughput of many routine and diagnostic re-sequencing applications. The technique will be of particular interest to those laboratories (especially in geographically isolated facilities, in developing and restructuring nations) where access to funds, labour and reagents are limiting and to those where sample numbers are very high and significant amounts of nucleic acid re-sequencing are necessary. We believe that the potential to combine the STAMMER method with the use of 'roadblock' oligonucleotides provides a superior and flexible means to perform relay sequencing on nucleic acid templates of almost any type and, most crucially, of almost any length. STAMMER represents a true sequencing assay rather than a means to molecular typing and, as it requires only 'off-the-shelf' sequencing reagents, analysis platforms and software, it is very easily implemented in laboratories already performing nucleic acid sequencing. We have shown that STAMMER sequencing can be used to perform not only full sequencing and consequently molecular typing of known polymorphic positions but also the bidirectional sequencing of individual double-stranded templates. The use of vector specific STAMMER sequencing systems utilising multiple 'roadblock' oligonucleotides will allow simultaneous bidirectional sequencing of circular templates; thus facilitating double-stranded plasmid screening at half the cost and

twice the throughput of normal sequencing. It is also possible to design STAMMER sequencing systems whereby several Short Template Amplicons are bidirectionally sequenced in a single assay.

The 'roadblock' molecule of the invention is a single-stranded nucleic acid that provides a physical barrier to the progression of a nucleic acid polymerase. The 3'-end of a roadblock molecule may be modified to prevent the extension of this end. This modification may involve chemical alteration of the 3'-end, such as the attachment of an alkyl group (e.g. C1 to C18 alkyl, preferably C2 to C6 alkyl, e.g. propyl). Different functional groups such as amino, thiol, or similar, or any group to which one might conjugate a terminating molecule, could also be attached to the 3'-end of the roadblock to prevent its extension. In addition, the last base at the 3'-end of the roadblock molecule might be a dideoxy base, or might otherwise lack potential to be extended as a result of similar modification of the nucleotide triphosphate structure itself or of the use of a non-classical or non-naturally occurring nucleotide. Further, the 3'-end of the roadblock may be abasic.

We recommend that the maximum combined length of all sequences in the STAMMER sequencing multiplex does not exceed 1500 bp as this allows operation of the sequence analysis platform within the parameters of conventional sequencing. Assuming that only such conventional sequence analysis platforms are available, we estimate that the maximum number of primers (to include megaprimers) that could be used in a single STAMMER sequencing experiment (and hence the maximum number of individual targets that could be sequenced) to be about 75. This estimate of 75 is based on the assumptions that a) the smallest primer used is 20 bp in length, b) 10 bp of sequence are obtained for each target, and c) a 10 bp 'silence' is used between each primer and the fully extended product of the next smallest primer. However, smaller numbers of primers are more typical; for example the number of primers may be up to 2, 3, 5, 7, 10, 15, 20, 25 or 40. For example, 20 primers could be used to provide 20 sequences of 75 bases in relay. An advantage of the invention is that STAMMER is the only technique that could utilise a future increase in the maximum size of sequencing

product resolved by a sequencing capillary (currently about 1500 bp). In principle, the total number of bases sequenced and the total number of primers used in a STAMMER multiplex is unlimited, provided that the technology exists to resolve the full range of extension products derived from the multiplex.

Preferably, the molar ratio of each megaprimer to target site is 1:20 to 20:1, more preferably 1:10 to 10:1, more preferably 1:5 to 10:1, more preferably 1:3 to 3:1, most preferably about 5:1.

A further embodiment of the invention involves the removal of the non-priming strand of at least one megaprimer prior to the sequencing reaction. The benefit of doing this is that one might, by removing the non-priming strand, allow the sequencing reaction to behave more like a conventional sequencing reaction in which a very high molar ratio between the oligonucleotide primer and amplicon is used. This increases the signal strength (which may be coupled to an increase in the sequence quality).

In a particular embodiment, a single strand of the megaprimer is removed prior to performing the STAMMER reaction. The simplest way of removing a single strand of a megaprimer(s) would be the application of linear-after-the-exponential PCR (LATE-PCR) according to Sanchez, J.A., Pierce, K.E., Rice, J.E. & Wangh, L.J. Linear-after-the-exponential (LATE)-PCR: an advanced method of asymmetric PCR and its uses in quantitative real-time analysis. *Proc Natl Acad Sci U S A* **101**, 1933-8 (2004). This process of asymmetric PCR uses an excess of forward strand primer to ensure that in the later stages of PCR, the reverse primer is entirely used up and a phase of linear amplification takes place, leading to an excess of single stranded products, which may be used as primers. This may be carried out in a single combined LATE-PCR amplification in one vessel. The short template amplicons, dideoxynucleotides and/or roadblock oligonucleotide(s) required to perform the STAMMER reaction could then be added after a defined number of cycles of such a linear amplification.

Another approach might be to selectively deplete a double stranded PCR product of the second strand. This might be carried out using a biotinylated reverse primer. The PCR product that is biotinylated in this way would be bound to a streptavidin paramagnetic bead or solid substrate, then subjected to conditions sufficient to denature the double helix and release the forward strand into solution which may be separated by magnetism or otherwise by simple elution. The use of paramagnetic methods might allow one to perform the second strand depletion in the same reaction vessel used for the STAMMER sequencing reaction. This might require the application of chemical denaturants, followed by the magnetic separation of the bead-bound reverse strands. This would be followed immediately by the sequencing reaction.

Single stranded megaprimers might also be constructed using two or more synthetic oligonucleotides representing a continuous sequence, which had been specifically bound to a third, possibly immobilised or bead bound scaffolding oligonucleotide. The action of DNA ligase enzyme could then be used to ligate the two synthetic oligonucleotides end to end, producing a much longer megaprimer, which could then be eluted from the scaffold.

The pool of extended primers produced by the sequencing reaction of the invention can be analysed by separating them according to their size. This separation can be achieved by electrophoresis, such as capillary electrophoresis, or by using a non-capillary based electrophoresis sequencing platform (such as a slab gel ABI 377 machine). The separation of extended primer products from a plurality of sequencing reactions can be performed in a single capillary. The chain terminators of the sequencing reactions, dideoxynucleotides (ddATP, ddGTP, ddCTP, or ddTTP), are labelled so that each extended primer product can be detected. For instance, each of the four dideoxynucleotide chain terminators can be labelled with a different fluorescent dye. Non-fluorescent chemical dyes (chromophores), substrates of enzymatic reactions, proteins or radioisotopes, or any other group to which one can attach any of the same, may also be used as labels, provided that one different label is used for each base.

The nucleic acids of the invention, including sequencing reaction templates, oligonucleotide primers, and megaprimers, can comprise deoxyribonucleic acid, ribonucleic acid, or a copolymer of deoxyribonucleic acid and ribonucleic acid. The nucleic acids of the invention can be natural (i.e. sourced from any organism) or synthetic, and can be made *in vitro* or *in vivo*, enzymatically or non-enzymatically. The template nucleic acid can comprise non-natural nucleotide analogues such as deoxyinosine or 7-deaza-2-deoxyguanosine. These analogues destabilize DNA duplexes and could allow a primer annealing and extension reaction to occur in a double-stranded sample without completely separating the strands.

The sequencing reaction of the invention may be carried out with a number of different nucleic acid polymerase enzymes, such as DNA polymerase, RNA polymerase, reverse transcriptase or RNA replicase, depending on the composition of the template nucleic acid.

The invention also provides a kit comprising components to carry out the methods of the invention. The kit may comprise at least one megaprimer comprising double-stranded nucleic acid of at least 100 bp in length, optionally at least 120 bp, at least 140 bp, at least 160 bp, at least 180 bp, at least 200 bp, at least 220 bp, at least 240 bp, at least 260 bp, at least 280 bp, or at least 300 bp in length. For use with conventional sequence analysis platforms, the maximum length of a megaprimer is about 1500 bp. The kit may comprise labelled chain terminators, optionally labelled with different fluorescent dyes. The kit may comprise non-terminating dNTPs or NTPs, an enzyme or enzymes for nucleic acid polymerisation, appropriate buffer solutions, purified water and/or instructions for carrying out any one of the methods of the invention.

Methods

DNA samples were prepared from whole blood using a standard salting out method[15]. PCR was performed in 30 μl reactions consisting of 100 ng of the genomic DNA template; 1 X Phusion™ HF Buffer (Finnzymes, Espoo, Finland);

0.8 mM dNTPs (Bioline, London, UK); each oligonucleotide primer (Sigma-Genosys, Cambridge, UK) at 500 μM in combination according to table 1a (NOD2/CARD15 simultaneous sequencing) and 1b (bidirectional STAMMER sequencing); MgCl2 (Bioline, London, UK) at concentrations according to table 1 and finally 0.4 U Phusion High-Fidelity DNA Polymerase (Finnzymes, Espoo, Finland). PCR cycling was carried out in a G-STORM GSX thermal cycler (GRI, Braintree, UK) and conditions were 34 cycles of (95°C: 0'10", 63°C: 0'10", 72°C: 1'00'') for the NOD2/CARD15 experiments. PCR conditions for the bidirectional experiments were 30 cycles of (95°C: 0'10", 67°C: 0'10", 72°C: 1'00''). The products were visualized on a 2% agarose gel then purified using the Wizard SV gel and PCR cleanup system (Promega, WI, USA).

Short Template Amplicons and Megaprimers were roughly quantified using a Nanodrop ND1000 Spectrophotometer (Wilmington, DE, USA) according to the Beer-Lambert equation using measurements of absorbance at 260nm and 280nm with an assumed average extinction coefficient of 50 ng-cm μl-1.

STAMMER sequencing reactions consisted of 2.5 μl Sequencing Buffer (ABI, Foster City, Ca, USA) and roughly 10 ng of each megaprimer and short template amplicon, made up to 28.25 μl with H2O. The final molar ratios of the NOD2/CARD15 megaprimers to their respective short template amplicon were around 2.22 (Exon 4), 1.88 (Exon 8) and 2.98 (Exon 11); the bidirectional sequencing megaprimer was used at a 4:1 ratio against the short template amplicon. Synthesis of the 'roadblock' oligonucleotide included the addition of a chain-terminating 3' propyl group (Sigma-Genosys, Cambridge, UK).

The 'roadblock' oligonucleotide was included in the bidirectional sequencing reaction at 500 nM whilst the oligonucleotide primer was included at just 16 nM.

In a G-STORM GSX thermal cycler (GRI, Braintree, UK), cycle sequencing reactions were subjected to 95°C for 1'00'' and were then incubated at 85°C. Hot-start was carried out by the manual addition to each reaction of 1.75 μl of a

mixture consisting of 1 µl H2O and 1 µl BigDye Terminators v3.0 Dye Mix (ABI, Foster City, Ca, USA). Cycle sequencing was then performed with 25 cycles of 98°C: 0'10'' and 60°C: 0'30''.

Sequences were detected on a 3730XL 96 capillary DNA analyser (ABI, Foster City, Ca, USA). EDTA was added to a final concentration of 6 mM and samples were then precipitated with 70 µl absolute propan-2-ol before recovery by centrifugation. Pellets were washed in 100 µl 70% propan-2-ol and were finally resuspended in 10 µl Hi-Di Formamide (ABI, Foster City, Ca., USA). Sample injection was for fifteen seconds at 1.5 kV. Electrophoresis was carried out at 60°C for 80 minutes. Sequences were analysed using Sequencing analysis v5.2 and SeqScape v2.5. Basecalling and assignment of Phred (Q20) quality values[16,17] was performed using the 'KB' basecaller (ABI, Foster City, Ca, USA).

Table 1

A) NOD2/CARD15		[$MgCl_2$]	Forward primer Sequence	Reverse Primer Sequence
Short Template Amplicons	Exon 4	1.6 mM	5' gTgCCAAAggTgTCgTgCCA	5' ggATggAgTggAAgTgCTTg
	Exon 8	2.2 mM	5' gCAggAgATAAAgAgCTAgCTTg	5' CTgCAggATAgACTCTgAAgC
	Exon 11	1.5 mM	5' gggAATCTCAgACATgAgCA	5' AgAgCTAAAACAggCCTgCC
Megaprimers	Exon 4	1.5 mM	5' gTgCCAAAggTgTCgTgCCA	5' CCAggAAggCTgCTgTgATC
	Exon 8	2.6 mM	5' gCAggAgATAAAgAgCTAgCTTg	5' CTgAAAAggCCAAAAgAgTCAAC
	Exon 11	1.8 mM	5' gggAATCTCAgACATgAgCA	5' TACAATggTgAgCTTCAAgg

B) Bidirectional STAMMER assay	[$MgCl_2$]	Forward primer sequence	Reverse Primer Sequence
Short Template Amplicon	1.0 mM	5' gTAgAAggATgAgTgCAAATCAgACA	5' CTGGAACATTCAGAGGGAAGG
Megaprimer	1.0 mM	5' TCCgTgCACATgACACTTCCACTTACT	5' CTggAACATTCAgAgggAAgg
		Reverse Primer Sequence	
Oligonucleotide primer	N/A	5' ggTAgAAggATgAgTgCAAATCAgACATTCTTCTCAggA	
Roadblock oligonucleotide	N/A	5' GTggAACAAgAgCTTTgCggTAAgAgAgAACgTAgTT-C3 spacer	

EXAMPLES

Example 1 - STAMMER sequencing of three unlinked Short Template Amplicons

We selected three regions of interest from Exons 11, 4 and 8 of the NOD2/CARD15 gene as polymorphisms within these regions had been shown to be informative with regards to disease associations[12].

Using PCR, we generated the three corresponding Short Template Amplicons whose lengths were 260 bp, 410 bp and 546 bp respectively. The megaprimers targeting each of the Short Template Amplicons for Exons 11, 4 and 8 were PCR products of respectively 130 bp, 276 bp and 437 bp (Figure 3). To demonstrate that double-stranded megaprimers could efficiently prime the sequencing reactions and suitably delay their detection, we primed each short template amplicon with its corresponding megaprimer in a separate reaction. We maintained relatively even molar ratios between the two moieties in order to maximise the number of productive megaprimer to template hybridisations during the cycle sequencing reaction. The data produced demonstrated that the detection of the fluorescence signals could be delayed to an extent that was proportional to the size of the megaprimer (Figure 4A-C). Signals relating to Exon 11 (Figure 4A) were detectable in the scanning time period equivalent to fragments in the size range of 131-260 bp. Exon 4 (Figure 4B) and Exon 8 (Figure 4C) were detectable in the periods equivalent to size ranges of 277-410 bp and 438-546 bp respectively.

Having demonstrated that megaprimers could prime sequencing reactions in the correct way, we went on to prepare a complete multiplex STAMMER sequencing reaction by mixing equal quantities of the three Short Template Amplicons with roughly equimolar amounts of the three megaprimers. All three sequencing reactions were then carried out simultaneously in a single reaction vessel. When the STAMMER multiplex relay sequencing reaction was performed using all three Short Template Amplicons, sequences were detectable in relay and were

interspersed with clear intervals of silence in which no fluorescently labelled fragments were detected (Figure 4D). The products of this reaction, when electrophoresed in a single capillary of the sequence analysis platform, produced three discrete sequences 'on-line' in the ranges of 131-260 bp (Exon 11), 277-410 bp (Exon 4) and 438-546 bp (Exon 8). The combined length of 370 simultaneously sequenced bases far exceeds that possible with, and is therefore superior to, other simultaneous sequencing methods.

The start and end of the megaprimer enabled sequences were marked by strong peaks relating to the tendency of the DNA polymerase enzyme to add a single and non template dependent adenosine residue to the 3′ termini of the megaprimer and Short Template Amplicon homoduplexes. These artifactual peaks are somewhat useful in providing a means to delimit and thereby distinguish the start and end of each sequence in the multiplex; but megaprimer strands so altered at their 3′ ends are incapable of efficiently and specifically hybridising to the megaprimer target region and this impacts greatly upon the strength and quality of the derived sequences. Such addition also depletes the adenosine content of the reaction and may lead to truncation of sequences. To eliminate the addition of non-template dependent 3′ adenosine residues during the STAMMER sequencing method, we made use of a proofreading enzyme. To reduce the extent to which such bases were incorporated during the initial ramp to denaturation temperature at the start of the cycle sequencing reaction, we applied a 'hot-start' wherein the sequencing enzyme was added to the reaction vessel only after the first round of denaturation and annealing had taken place[13]. In an ideal implementation of the STAMMER sequencing method, we would prevent non-template dependent A addition through utilisation of a sequencing reaction mixture modified to include a proof-reading enzymic activity or where this was not possible (i.e. during bidirectional sequencing) a compensating quantity of deoxy- and dideoxy-adenosine trinucleotides.

Another approach to reducing or eliminating the 'A' peaks caused by non-template dependent addition of adenosine to the 3′-end of the megaprimer and

amplicon homoduplexes during the STAMMER technique is to use a non-proofreading enzyme during PCR preparation of the megaprimer(s) or Short Template Amplicons. This ensures that most megaprimer/Short Template Amplicon products are already adenylated at the 3′ end and so little or no ddATP is added during STAMMER sequencing and the 'A' peaks are thus reduced. With time the megaprimers may lose the terminal A and therefore the peaks may return, because the ssDNA end is less stable than dsDNA.

Through incubation of the megaprimer with DNA polymerase and dATP in a suitable buffer, the A bases can be restored prior to STAMMER sequencing. It would be preferable to use megaprimers where the addition of the extra A has no effect on the binding footprint of the 3′ end of the megaprimer (i.e. where the base immediately downstream of the megaprimer target region is a thymine residue); in this way specificity is not lost. The addition of the 3′ A residue should, however, interfere with the megaprimer forward strand's ability to anneal to its own reverse strand, thereby reducing self-annealing that would otherwise limit the productivity of the reaction.

The addition of 3′ terminal adenosine residues to the short template amplicon can be eliminated by the application of exonuclease 1 and shrimp alkaline phosphatase (EXO-SAP) subsequent to the PCR amplification. Exonuclease removes the single stranded portion of the molecule (along with excess amplification primers) and the alkaline phosphatase dephosphorylates it and so prevents any form of polymerase/endonuclease from acting upon it. The megaprimer cannot be treated with Exo-SAP, as it is essential that the megaprimer can be subject to polymerase activity.

Example 2 - Bidirectional sequencing of a single short template amplicon using the STAMMER sequencing method

Many applications of sequencing technology such as plasmid screening and sequence-based typing[14] require the sequencing of both the anti-sense and sense

strands of a double-stranded nucleic acid template region of interest. This has historically meant that two sequencing reactions have been necessarily performed for the detection of bidirectional sequence in any one region of interest. As a further proof of the application of the STAMMER sequencing principle, we demonstrate how the basic method described previously can be adapted to perform the sequencing of both strands of a single short template amplicon in a single relay sequencing reaction. The bidirectional STAMMER approach (Figure 5A) depends on the fact that the DNA polymerase enzyme used in the sequencing reaction lacks a 5′-3′ exonuclease activity and thus the primer extension reaction can be halted by the obstacle of the 5′ end of a "roadblock", a hybridised nucleic acid strand in the path of the enzyme.

We selected a single 170 bp region of interest in KIR2DL4, a highly complex and GC rich gene (Robinson, J., Waller, M.J., Stoehr, P. & Marsh, S.G.E. IPD--the Immuno Polymorphism Database. *Nucleic Acids Res* 33, D523-6 (2005)) that had previously been shown to be a troublesome sequencing template requiring bidirectional coverage to reconstruct continuous high quality sequence (Roberts, C.H., Madrigal, J.A. & Marsh, S.G. Cloning and sequencing alleles of the KIR2DL4 gene from genomic DNA samples. *Tissue Antigens* 69 Suppl 1, 88-91 (2007)). An oligonucleotide was used in the manner of a conventional reaction to prime the sequencing of the antisense strand of a 475 bp short template amplicon constructed according to the primers described in table 1b. A 231 bp megaprimer simultaneously enabled the sequencing of a 224 bp region of the sense strand from the same short template amplicon. Under normal circumstances, the oligonucleotide-enabled component of the sequencing reaction would progress across the entire short template amplicon and would still be detectable when the megaprimer-enabled component later became detectable. To prevent this, it was necessary to halt the oligonucleotide-enabled reaction before the end of the template was reached. A previous method of bidirectional sequencing introduced abasic regions to the template in order to halt the sequencing reaction in such a way[6], but this was not tenable in the STAMMER system. Instead, we simply placed a second oligonucleotide, a "roadblock", on the sense strand 179 bp

downstream of 3′ end of the sequence enabling oligonucleotide primer and 2bp upstream of the 5′ end of the megaprimer target region (Figure 5A). The 3′ end of this roadblock was modified with a chain-terminating propyl group (C3 spacer) and therefore could not itself act as a sequence enabling primer. The 5′ end meanwhile presented an obstacle aimed at halting the progression of the 5′ to 3′ endonuclease activity of the polymerase enzyme and thus limited the size of the oligonucleotide primer enabled sequencing products. The use of the roadblock oligonucleotide at a high molar ratio with respect to both the short template amplicon and the oligonucleotide primer ensured that essentially all free template sense strands were hybridised to a roadblock molecule capable of halting the oligonucleotide primer extension reaction.

We went on to perform cycle sequencing reactions incorporating the Short Template Amplicon, its corresponding megaprimer and both the oligonucleotide primer and the roadblock oligonucleotide. Figure 5C is the base called electrophoretic data output of a complete STAMMER Bidirectional sequencing reaction. Analysis of the products from this reaction showed that we were able to detect two discrete sequences in relay. Further scrutiny of these sequences confirmed that we had sequenced in relay both strands of a double-stranded template region of the KIR2DL4 gene and that the two sequences were complementary to one another across a 167 bp expanse of the region of interest. The average Phred quality value across the bidirectional sequence was 39.4; most nucleotide positions were identified by high quality base-calls (Q≥20) in both directions and all positions had such in at least one direction.

REFERENCES

1. Sanger, F., Nicklen, S. & Coulson, A.R. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 74, 5463-7 (1977).

2. Wiemann, S. *et al.* Simultaneous on-line DNA sequencing on both strands with two fluorescent dyes. Anal Biochem 224, 117-21 (1995).

3. Wiemann, S. *et al.* "Doublex" fluorescent DNA sequencing: two independent sequences obtained simultaneously in one reaction with internal labeling and unlabeled primers. Anal Biochem 234, 166-74 (1996).

4. van den Boom, D., Jurinke, C., Ruppert, A. & Koster, H. Forward and reverse DNA sequencing in a single reaction. Anal Biochem 256, 127-9 (1998).

5. Yager, T.D. *et al.* High performance DNA sequencing, and the detection of mutations and polymorphisms, on the Clipper sequencer. Electrophoresis 20, 1280-300 (1999).

6. Murphy, K.M. & Eshleman, J.R. Simultaneous sequencing of multiple polymerase chain reaction products and combined polymerase chain reaction with cycle sequencing in single reactions. Am J Pathol 161, 27-33 (2002).

7. Padmanabhan, R. & Wu, R. Nucleotide sequence analysis of DNA. IX. Use of oligonucleotides of defined sequence as primers in DNA sequence analysis. Biochem Biophys Res Commun 48, 1295-302 (1972).

8. Smith, L.M. *et al.* Fluorescence detection in automated DNA sequence analysis. Nature 321, 674-9 (1986).

9. Kuppuswamy, M.N. *et al.* Single nucleotide primer extension to detect genetic diseases: experimental application to hemophilia B (factor IX) and cystic fibrosis genes. Proc Natl Acad Sci U S A 88, 1143-7 (1991).

10. Nyren, P., Pettersson, B. & Uhlen, M. Solid phase DNA minisequencing by an enzymatic luminometric inorganic pyrophosphate detection assay. Anal Biochem 208, 171-5 (1993).

11. Gilbert, M.T. *et al.* Multiplex PCR with minisequencing as an effective high-throughput SNP typing method for formalin-fixed tissue. Electrophoresis 28, 2361-7 (2007).

12. Mayor, N.P. *et al.* Single nucleotide polymorphisms in the NOD2/CARD15 gene are associated with an increased risk of relapse and death for patients with acute leukemia after hematopoietic stem-cell transplantation with unrelated donors. J Clin Oncol 25, 4262-9 (2007).

13. Chou, Q., Russell, M., Birch, D.E., Raymond, J. & Bloch, W. Prevention of pre-PCR mis-priming and primer dimerization improves low-copy-number amplifications. Nucleic Acids Res 20, 1717-23 (1992).

14. Santamaria, P. *et al.* HLA class II "typing": direct sequencing of DRB, DQB, and DQA genes. Hum Immunol 33, 69-81 (1992).

15. Miller, S.A., Dykes, D.D. & Polesky, H.F. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16, 1215 (1988).

16. Ewing, B. & Green, P. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res 8, 186-94 (1998).

17. Ewing, B., Hillier, L., Wendl, M.C. & Green, P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 8, 175-85 (1998).

CLAIMS

1. A method for simultaneously sequencing a plurality of nucleic acid target sites, comprising:

(a) annealing each of a plurality of different-sized primers to the 5′-end of each of said plurality of nucleic acid target sites;

(b) performing a sequencing reaction by simultaneously extending said primers in the presence of labelled chain terminators and thereby obtaining a pool of extended primers in a single reaction vessel; and

(c) determining the sequences of the nucleic acid target sites by analysing said pool of extended primers without separating the pool prior to analysis;

wherein the largest extension product of each primer is smaller in length than the next largest, non-extended primer, and

wherein at least one of the primers is a megaprimer comprising double-stranded nucleic acid of at least 100 base pairs in length.

2. The method of claim 1 wherein the length of extension product of at least one primer is limited by the length of template available for that primer.

3. The method of claim 1, wherein the length of extension product of at least one primer is limited by annealing a roadblock nucleic acid molecule to a site downstream of the primer.

4. The method of claim 3, wherein the roadblock nucleic acid molecule comprises a modification of its 3′-end that prevents extension.

5. The method of claim 4, wherein said modification is attachment of an alkyl group.

6. The method of any one of the preceding claims, further comprising removing the non-priming strand of at least one megaprimer prior to the sequencing reaction.

7. The method of any one of the preceding claims, wherein the pool of extended primers is analysed by separating them according to their size.

8. The method of claim 7, wherein the separation is by electrophoresis.

9. The method of claim 8, wherein the electrophoresis is performed in a single capillary or in a single lane of a slab gel.

10. The method of any one of the preceding claims, wherein at least one megaprimer is produced by:

(a) PCR;

(b) restriction digestion; or

(c) non-PCR DNA amplification, optionally rolling circle amplification (RCA), multiple displacement amplification (MDA), real time strand displacement amplification (SDA), or other isothermal amplification.

11. The method of any one of the preceding claims, wherein a plurality of non-complementary target sites are sequenced unidirectionally.

12. The method of claim 11, wherein a polymerase possessing 3′ to 5′ endonuclease proofreading activity is used to carry out the sequencing reaction.

13. The method of any one of claims 1 to 10, which comprises bidirectional sequencing of at least one duplex nucleic acid molecule.

14. The method of claim 13, wherein said duplex nucleic acid molecule is linear or circular.

15. A kit for performing a method of any one of claims 1 to 14, wherein the kit comprises (a) at least one megaprimer comprising double-stranded nucleic acid of at least 100 base pairs in length and (b) labelled chain terminators.

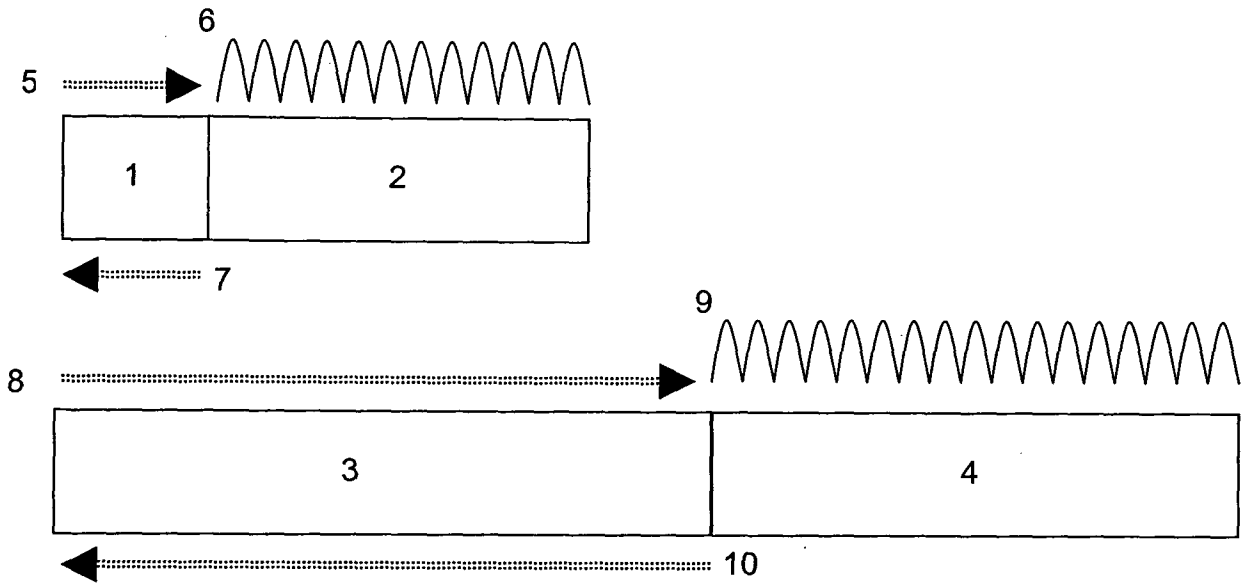
16. The kit of claim 15, wherein the kit comprises at least one roadblock nucleic acid molecule comprising a modification at the 3′-end that prevents extension.

17. The kit of claim 16, wherein the modification is attachment of an alkyl group.

18. The kit of claim 15, 16 or 17, wherein the kit comprises a DNA polymerase.

19. The kit of any one of claims 15 to 18, wherein the kit comprises instructions for performing a method of any one of claims 1 to 14.

Figure 1



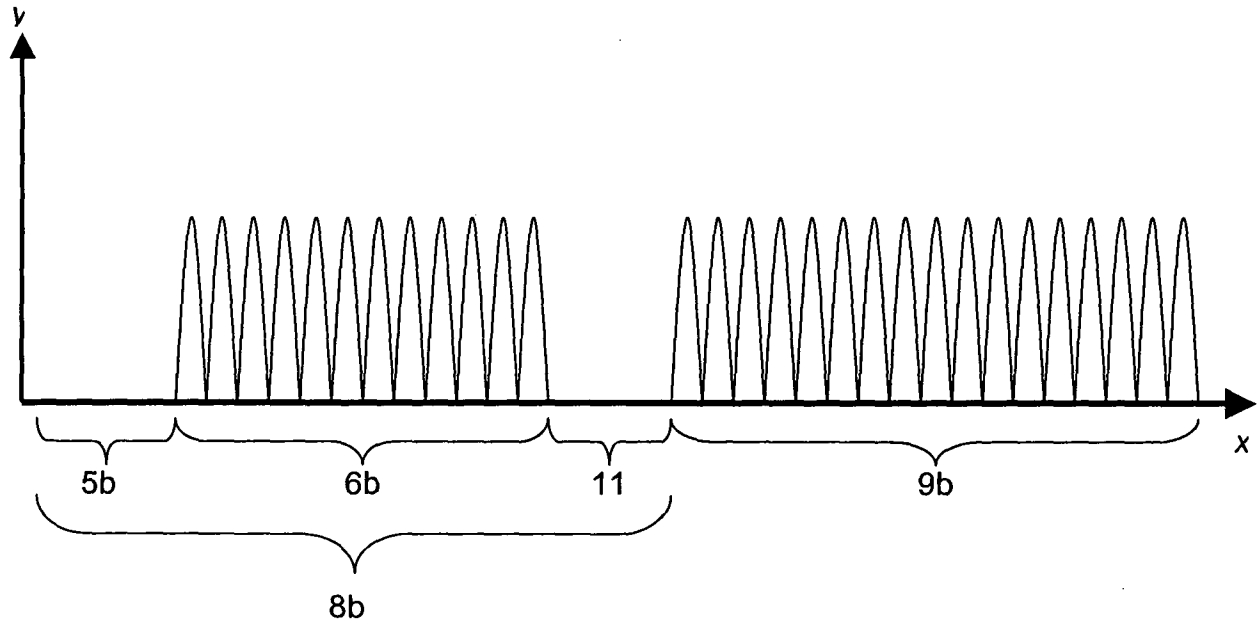


Figure 2

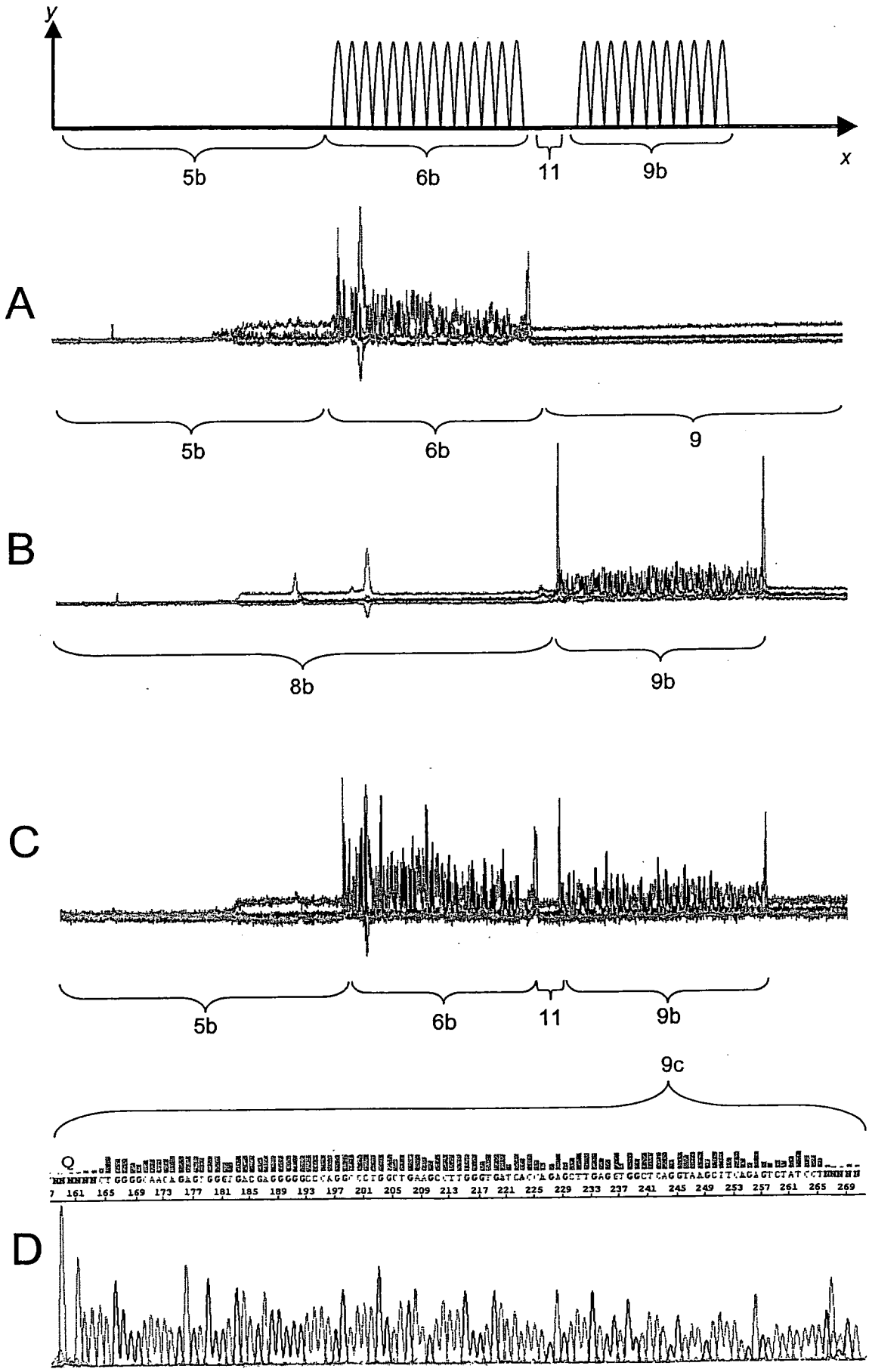


Figure 3

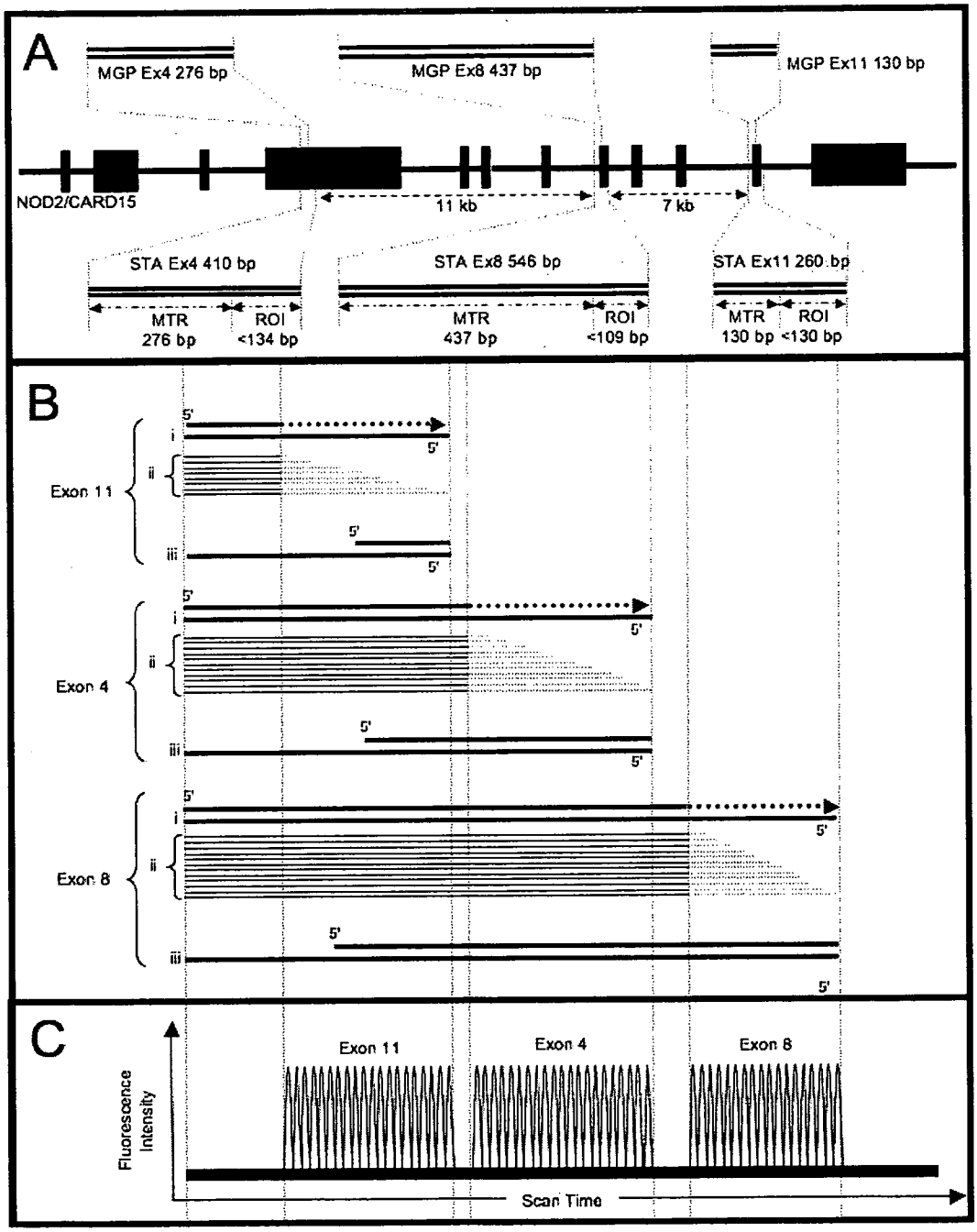

A
MGP Ex4 276 bp
MGP Ex8 437 bp
MGP Ex11 130 bp
NOD2/CARD15
11 kb
7 kb
STA Ex4 410 bp
STA Ex8 546 bp
STA Ex11 260 bp
MTR 276 bp
ROI <134 bp
MTR 437 bp
ROI <109 bp
MTR 130 bp
ROI <130 bp
B
Exon 11
Exon 4
Exon 8
i
ii
iii
5'
C
Exon 11
Exon 4
Exon 8
Fluorescence Intensity
Scan Time

Figure 4

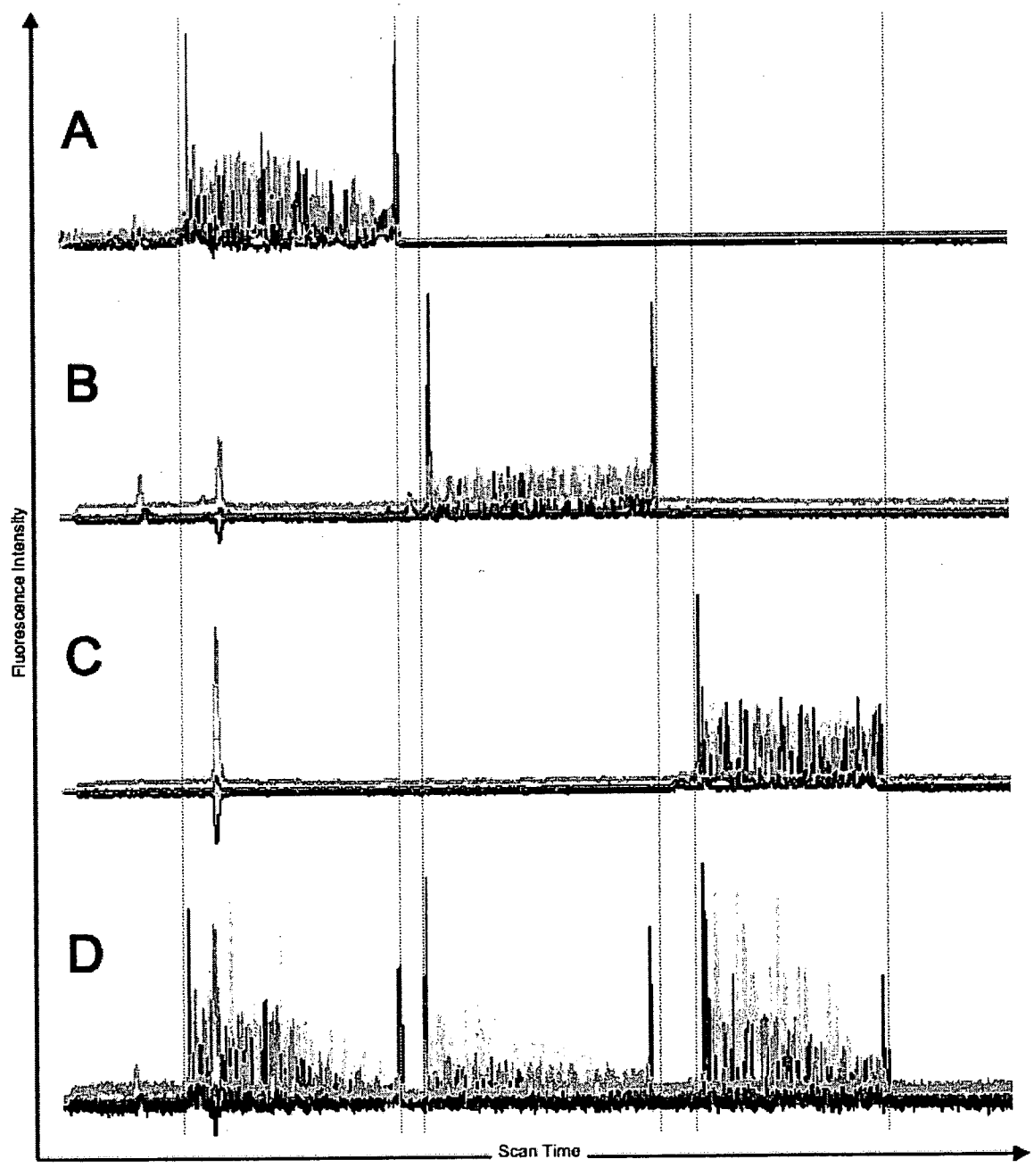


Figure 5

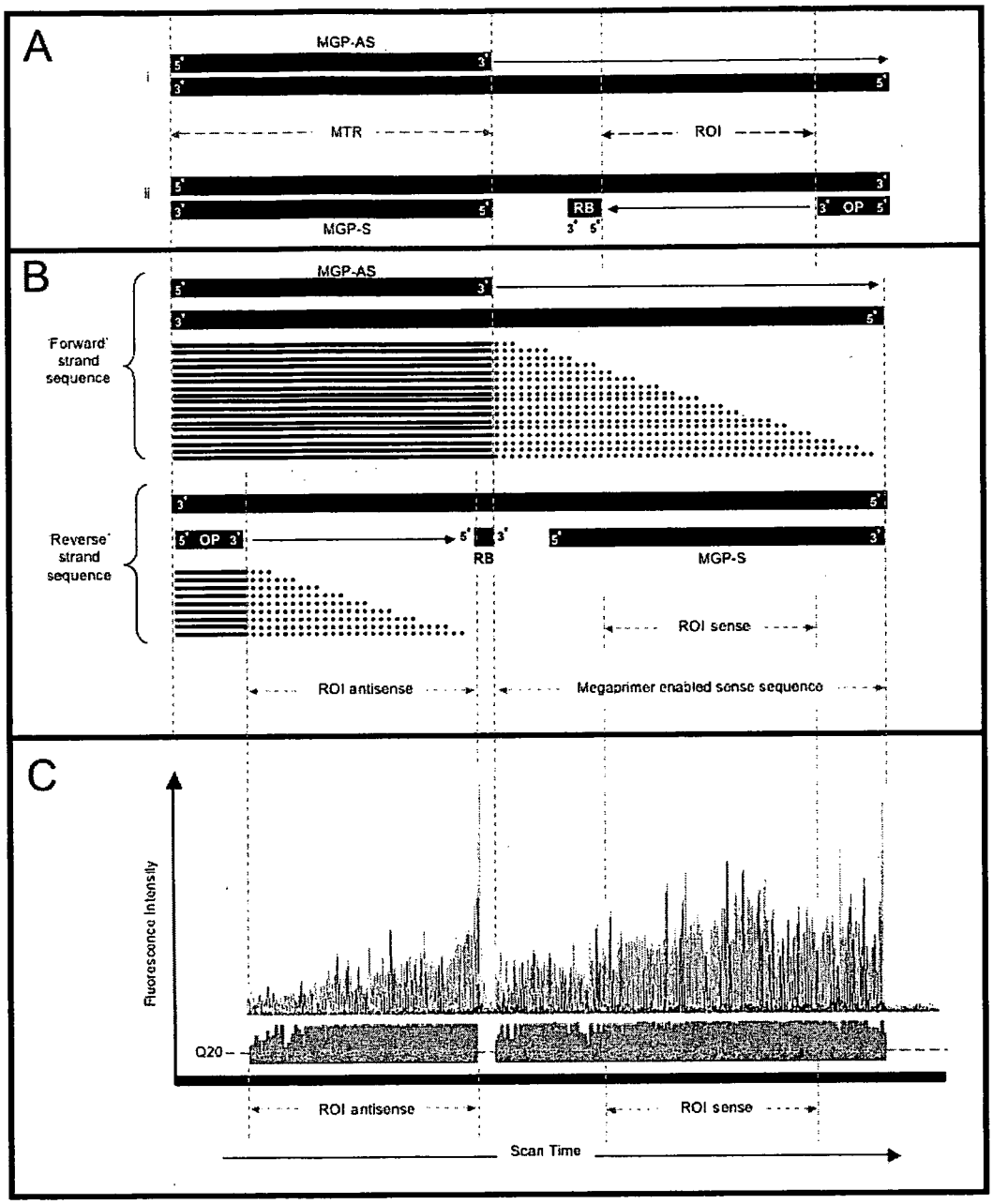
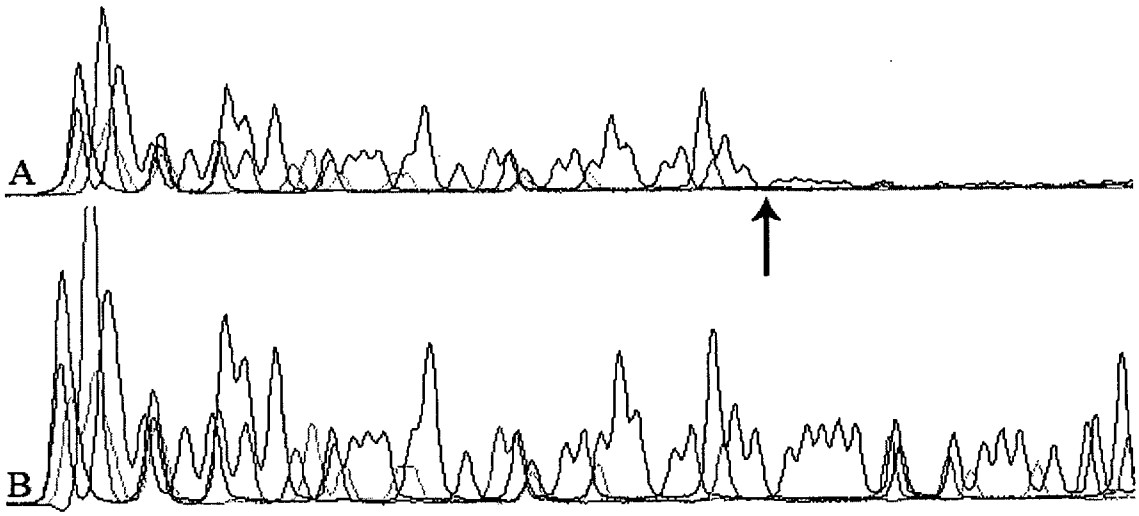

A
MGP-AS
MTR
ROI
MGP-S
RB
OP
B
MGP-AS
'Forward' strand sequence
'Reverse' strand sequence
OP
RB
MGP-S
ROI sense
ROI antisense
Megaprimer enabled sense sequence
C
Fluorescence Intensity
Q20
ROI antisense
ROI sense
Scan Time

Figure 6



INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2010/000861

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12Q1/68
ADD.

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)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	ROBERTS CHRISSY H ET AL: "Short template amplicon and multiplex megaprimer-enabled relay (STAMMER) sequencing, a simultaneous approach to higher throughput sequence-based typing of polymorphic genes." IMMUNOGENETICS APR 2010 LNKD-PUBMED:20204613, vol. 62, no. 4, April 2010 (2010-04), pages 253-260, XP002590288 ISSN: 1432-1211 the whole document ----- -/--	1-19

[X] Further documents are listed in the continuation of Box C.

[X] 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	Date of mailing of the international search report
15 July 2010	04/08/2010

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Quirin, Katharina

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2010/000861

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 00/20628 A1 (BIO ID DIAGNOSTIC INC [CA]; VINAYAGAMOORTHY THURAIAYAH [CA]) 13 April 2000 (2000-04-13) the whole document pages 8-9 page 7; figures 1-3	1-19
X	WO 03/056030 A2 (UNIV JOHNS HOPKINS [US]; ESHLEMAN JAMES R [US]; MURPHY KATHLEEN M [US]) 10 July 2003 (2003-07-10) cited in the application the whole document pages 16-17; figure 1	1-19
A	US 5 871 929 A (BARNES WAYNE M [US]) 16 February 1999 (1999-02-16) columns 12-13 the whole document	1-15
A	SARKAR G ET AL: "THE MEGAPRIMER METHOD OF SITE-DIRECTED MUTAGENESIS" 1 January 1990 (1990-01-01), BIOTECHNIQUES, NATICK, MA, US, PAGE(S) 404 - 407 , XP008009292 the whole document	1-19
X	WO 02/090505 A2 (UNIV VIRGINIA COMMONWEALTH [US]; CHEN XIANGNING [US]) 14 November 2002 (2002-11-14) the whole document pages 4-6; figures 1,2 pages 9-17	1-19

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/GB2010/000861

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0020628 A1	13-04-2000	AT 266738 T	15-05-2004
		AU 765504 B2	18-09-2003
		AU 5963899 A	26-04-2000
		CA 2344741 A1	13-04-2000
		CN 1329673 A	02-01-2002
		DE 69917303 D1	17-06-2004
		DE 69917303 T2	16-06-2005
		EP 1117839 A1	25-07-2001
		ES 2221479 T3	16-12-2004
		ID 29579 A	06-09-2001
		MX PA01004309 A	06-06-2003
		NZ 511415 A	28-02-2003
		US 6197510 B1	06-03-2001
WO 03056030 A2	10-07-2003	AU 2002365157 A1	15-07-2003
		EP 1472335 A2	03-11-2004
US 5871929 A	16-02-1999	NONE	
WO 02090505 A2	14-11-2002	AU 2002305436 A1	18-11-2002