

Method for Preparing Polynucleotides for Analysis

Field of the Invention

This invention relates to methods for modifying polynucleotides to allow analysis of the polynucleotides to be carried out more readily.

5 Background to the Invention

Advances in the study of molecules have been led, in part, by improvement in technologies used to characterise the molecules or their biological reactions. In particular, the study of the nucleic acids DNA and RNA has benefited from developing technologies used for sequence analysis and the
10 study of hybridisation events.

WO-A-00/39333 describes a method for sequencing polynucleotides by converting the sequence of a target polynucleotide into a second polynucleotide having a defined sequence and positional information contained therein. The sequence information of the target is said to be "magnified" in the second
15 polynucleotide, allowing greater ease of distinguishing between the individual bases on the target molecule. This is achieved using "magnifying tags", which are predetermined units of nucleic acid sequence. Each of the bases adenine, cytosine, guanine and thymine on the target molecule is represented by an individual magnifying tag, converting the original target sequence into a
20 magnified sequence. Conventional techniques may then be used to determine the order of the magnifying tags, and thereby determine the specific sequence on the target polynucleotide.

In a preferred sequencing method, each magnifying tag comprises a label, e.g. a fluorescent label, which may then be identified and used to characterise
25 the magnifying tag.

WO-A-04/094664 describes an adaptation of the conversion method disclosed in WO-A-00/39333. In both methods, it is preferred that each magnifying tag comprises two units of distinct sequence which can be used as a binary system, with one unit representing "0" and the other representing "1".
30 Each base on the target is characterised by a combination of the two units, for example adenine may be represented by "0" + "0", cytosine by "0" + "1", guanine by "1" + "0" and thymine by "1" + "1".

One difficulty with the prior art methods is that the eventual read-out step is often hindered by the need to discriminate between the different magnifying tags or units. It is therefore desirable to identify improvements which permit discrimination to occur.

5 Summary of the Invention

The present invention provides a method for analysing polynucleotides, preferably those polynucleotides which have been formed with distinct units of polynucleotide sequence each representing a particular characteristic. The method utilises a concatemer of the target polynucleotide, i.e. repeating the
10 sequence of the target polynucleotide, and then forming a further polynucleotide on this, the further polynucleotide being hybridised at specific portions of the target, such that hybridised or non-hybridised sequences can be identified and the order of hybridisation (or non-hybridisation) reveals the identity and/or order of the units of the target polynucleotide. The intention is, preferably, to identify
15 sequentially one unit of each target polynucleotide on the concatemer. In this way, the units to be identified are more separated than if the units of the original target polynucleotide were to be sequenced. Increasing the separation allows the eventual read-out technology to discriminate between the units, thereby improving the efficiency of the eventual sequencing/identification step.
20 Alternatively the method may be carried out so that repeated target polynucleotides are separated by additional nucleic acid sequences which act to space apart the targets, so that analysis can be performed.

According to a first aspect of the present invention, a method for analysing a target polynucleotide having distinct units of nucleic acid sequence, comprises:
25 (i) forming a first polynucleotide which is a concatemer having multiple repeating target polynucleotide sequences;
(ii) hybridising a portion of the first polynucleotide to a second polynucleotide, such that the portion hybridised, or the portion not hybridised, corresponds to at least a sequence unit on the target, and determining the
30 sequence unit on the target.

In a second aspect of the invention, the concatemer is formed directly on to a second polynucleotide having a predetermined order of defined first and second nucleic acid sequence units. The second polynucleotide is designed to

hybridise to specific sequence units on the target. A specific interaction between the concatemer and the second polynucleotide occurs such that there are hybridised portions that can be interrogated to reveal the identity of the sequence unit.

5 According to a second aspect of the invention, a method for the conversion of a target polynucleotide having distinct units of nucleic acid sequence into a polynucleotide having nucleic acid sequences separating one or more of the distinct nucleic acid sequence units, comprises:

10 (i) forming a second polynucleotide having defined first and second sequence units in a predetermined order;

 (ii) forming on the second polynucleotide, a first polynucleotide made up of a series of the target polynucleotides, wherein the order of the first and second sequence units permits the specific interaction between the first and second polynucleotides, such that distinct units of nucleic acid sequence on the
15 first polynucleotide can be distinguished from other distinct units by the extent of interaction; and optionally

 (iii) determining the sequence and/or order of the distinct sequence units on the basis of the interaction to thereby determine one or more specific sequence units on the target polynucleotide.

20 According to a third aspect of the present invention, a method for the formation of a double-stranded polynucleotide, comprises:

 (i) forming a first polynucleotide;
 (ii) carrying out a rolling circle amplification reaction from a primer molecule attached to the first polynucleotide, wherein the amplification reaction
25 utilises a circular polynucleotide molecule having a sequence that is complementary to repeating units of the first polynucleotide.

Brief Description of the Drawings

The invention is described with reference to the accompanying drawings, wherein:

30 Figure 1 (a) illustrates the use of different nucleic acid sequences to represent either a "0" or "1" characteristic, and the sequence on the second polynucleotide which will hybridise to the sequence, or not;

 Figure 1 (b) is a graphic illustration of roll back PCT;

Figure 2 illustrates the masking or unmasking of a target polynucleotide;

Figure 3 illustrates the use of padlock probes in the interrogation of a nucleic acid sequence which is "unmasked";

Figure 4 illustrates the design of a target polynucleotide and the second
5 polynucleotide;

Figure 5 illustrates the formation of a target polynucleotide having defined units of nucleic acid sequence;

Figure 6 illustrates the concatemerisation of a target with additional intervening nucleic acid sequence;

10 Figure 7 (a) illustrates the use of an array of second polynucleotides used to capture and characterise the target polynucleotide;

Figure 7 (b) illustrates the use of padlock probes to interrogate the target polynucleotide;

Figure 8 illustrates the hybridisation between the concatemerised target
15 polynucleotides and the second polynucleotide and the presence of mismatched (unmasked) sequence;

Figure 9 illustrates the use of padlock probes to hybridise the unmasked regions of the concatemerised target polynucleotide;

20 Figures 10 to 12 each illustrate the use of restriction enzyme substrate sequences to interrogate the target polynucleotide; and

Figure 13 is an electrophoresis gel showing the products of restriction enzyme digestion of the hybrids shown in figures 10 to 12.

Description of the Invention

25 The term "polynucleotide" is well known in the art and is used to refer to a series of linked nucleic acid molecules, e.g. DNA or RNA. Nucleic acid mimics, e.g. PNA, LNA (locked nucleic acid) and 2'-O-methRNA are also within the scope of the invention.

30 The reference herein to the bases A, T(U), G and C, relate to the nucleotide bases adenine, thymine (uracil), guanine and cytosine, as will be appreciated in the art. Uracil replaces thymine when the polynucleotide is RNA, or it can be introduced into DNA using dUTP, again as well understood in the art.

The term "first polynucleotide" is used herein to refer to the concatemerised target polynucleotide, i.e. the first polynucleotide comprises

repeated target polynucleotide sequences. The target polynucleotides may be linked sequentially, or there may be additional nucleic acid sequences separating the target polynucleotides.

5 The term "second polynucleotide" is used herein to refer to a polynucleotide intended to hybridise to regions of the first polynucleotide. The second polynucleotide may also be referred to as a masking polynucleotide as it acts to prevent interrogation of those regions of the first polynucleotide to which it hybridises. The regions of the first polynucleotide that are not hybridised are said to be "unmasked".

10 The method of the present invention is used to convert a target polynucleotide having distinct units of nucleic acid sequence into a polynucleotide where the distinct units of nucleic acid sequence can be interrogated at intervals more spaced apart than that of the target. This has the benefit of separating the units to permit the ultimate read-out steps to be
15 performed with more accuracy and discrimination. The invention relies on the formation of a concatemer of the target polynucleotide which permits subsequent interrogation to be performed on selected units; the interrogated units are representative of the distinct units of the original target polynucleotide.

20 Having formed the concatemer, distinct units of the target polynucleotide can be interrogated in various ways to reveal the identity of one or more of the units or to identify the presence of a specific sequence present on the target.

The preferred way of interrogating the concatemer is to hybridise the concatemer with one or more second polynucleotides of defined sequence, such that the second polynucleotide acts to mask portions of the concatemer, leaving
25 the unmasked portions (sequence units) available for interrogation with, for example, a labelled polynucleotide specific for that portion. Identification of the labelled polynucleotide reveals the identity of the sequence units. Different sequence units on each of the concatemer's target polynucleotides can be identified in this way.

30 The method of the invention relies on the use of a target polynucleotide that is comprised of defined nucleic acid sequence "units", where each unit represents a specific characteristic of an earlier molecule. For example, each unit may represent a specific base on an original polynucleotide molecule of

unknown sequence. Each unit will preferably comprise 2 or more nucleotide bases, preferably from 2 to 50 bases, more preferably 5 to 20 bases and most preferably 5 to 10 bases, e.g. 6 bases. There are preferably at least two different bases contained in each unit. The design of the units is such that it will be possible to distinguish the different units during a "read-out" step, e.g. involving the incorporation of detectably labelled nucleotides in a polymerisation reaction, or on hybridisation of complementary oligonucleotides. Methods in which a molecule is converted into a target polynucleotide are well known in the art, for example as described in WO-A-00/39333 and WO-A-04/094664 the content of each being incorporated herein by reference. However, a unit may be a single base, if necessary.

In a preferred embodiment the target polynucleotide comprises a series of distinct units, the combination of which imparts sequence or other information. For example, two units may be used as a binary system, with one unit representing "0" and the other representing "1". The combination of different units may represent bases on an original polynucleotide that is being sequenced. This is disclosed in WO-A-04/094664.

The concatemer therefore comprises repeated units of nucleic acid sequence.

The target polynucleotide may be designed so that a part of each "unit" represents the characteristic of the original molecule under study, but that each unit also comprises nucleic acid sequence of known sequence. The known sequence can be different for each unit, so that, prior to analysis, at least a partial sequence of the target is known. This allows the second polynucleotide to be designed for hybridisation with the target. This is shown in more detail in figures 4 and 5, where there are two possible unit sequences for each unit position on the target; each unit may represent "0" or "1", depending on the sequence of the two nucleotides. However, the remaining eight nucleotides are the same for "0" or "1", for each position, but different to the next unit position. In this way, the eventual interrogation can be designed so that hybridisation with specific units can be performed leaving units at specific positions un-hybridised, thereby allowing interrogation of the un-hybridised units to take place. In one embodiment, the intention is to interrogate a different unit position on each target

polynucleotide of the concatemer, thereby allowing further separation of the units and improving the ability of the eventual read-out step to discriminate between the sequence units.

Alternatively, if the target sequence is not known, the target polynucleotide
5 may be ligated to a known sequence prior to concatemerisation, so that the concatemer comprises both target polynucleotide sequences and known sequences, so that hybridisation can occur between the concatemer and the second polynucleotide. In this embodiment, the known sequences should be of sufficient length to permit hybridisation with the second polynucleotide to occur.
10 For example, the known sequences should be more than 100 nucleotides, preferably more than 500 nucleotides. This provides separation between the hybridised sequences and the un-hybridised (target) sequences, which can then be interrogated.

The conditions necessary for carrying out the method of the invention,
15 including temperature, pH, buffer compositions etc., will be apparent to those skilled in the art.

The method of the invention follows three essential steps:

- i. concatemerisation of the target polynucleotide;
- ii hybridising the concatemer with a further (second) polynucleotide
20 of defined known sequence;
- ii identifying the units on the concatemer that are hybridised or, preferably, that are not hybridised.

The concatemer may be formed in any suitable way. In a preferred embodiment, the target polynucleotide is circularised and the polymerase
25 reaction is carried out to form the concatemer.

Circularisation

The target may be circularised in any convenient way. In one embodiment, the single-stranded target is hybridized to the 3' end of the second polynucleotide. Both the 5' and the 3' end of the target molecule will hybridize
30 to the second polynucleotide and will be ligated together forming a single-stranded circle. The efficiency of circle ligations is much better with increased complementarity and it is preferred to use at least 6 complementary nucleotides, preferably at least 9 complementary nucleotides for hybridisation to the second

polynucleotide. The ligase can be any available ligase, but is preferably T4 DNA ligase, E.coli DNA ligase or Taq DNA ligase.

In an alternative method, a support oligonucleotide can be used to hybridise to the target and to ligate to the second polynucleotide. In one embodiment of this, the hybrid forms a partially double-stranded molecule with an overhang complementary to the second polynucleotide's 3' end. The support oligonucleotide can then be ligated to the second polynucleotide at the 3' end. The 5' end of the target is also complementary to the second polynucleotide and so the target will hybridise to the second polynucleotide bringing the two ends of the target into position for a ligase to join the two ends of the target, forming a circle. The support oligonucleotide acts to help retain the now circularised target at the second polynucleotide, ready for concatemerisation.

In an alternative embodiment, the support oligonucleotide is used as a splint oligonucleotide. The splint oligonucleotide is complementary to the 3' and 5' regions of the target and brings these two ends into proximity, allowing a ligase reaction to occur. The hybrid of the now circularised target and splint oligo is then brought into contact with the second polynucleotide and the target hybridises at the end of the second polynucleotide in a way designed to bring the splint oligonucleotide into proximity to the second polynucleotide. Ligation of the splint oligo to the second polynucleotide can then occur, which aids future polymerase amplification by providing a sufficient number of bases to which the circular target can hybridise to prior to a further round of amplification.

The support oligonucleotide will be of a size sufficient to aid hybridisation and circularisation with the target.

In a third alternative method, a splint oligo is used. The splint oligo is a short single-stranded molecule complementary to the 3' and 5' ends of the target. Hence, the splint oligo will bring the termini of the target into position for a ligase to join the two ends, forming a circle. The splint can then be removed by using exonucleases e.g. exonuclease I and exonuclease III, to digest the splint oligo.

Concatemerisation

The target polynucleotide can be concatemerised using a polymerase reaction. In one embodiment, the circularised target polynucleotide acts as a template for a polymerase reaction. As the template is a circular molecule, the

technique used is commonly known as Rolling Circle Amplification (RCA). Several variants of this method exist, as reviewed in Richardson et al., Genetic engineering, 25, 51-63. Linear RCA utilises one primer producing one concatemer from each template. Exponential RCA utilises two primers, where
5 one is complementary to the target to be amplified, while the other is complementary to the product generated by the first primer. Hence, the second primer initiates the synthesis of multiple concatemerised copies from one target polynucleotide. Multiply-primed RCA utilises a set of random hexamers as primers. These primers initiate the synthesis of multiple concatemerised copies
10 from one target polynucleotide. Secondary non-specific priming events can occur subsequently on the displaced product strands of the initial RCA step. The polymerase activity can be initiated at the 3' end of the secondary polynucleotide. Several polymerases can be used, including Sequenase, Bst DNA polymerase (large fragment), Klenow exo- DNA polymerase, which are all polymerases
15 operating at 37°C and displaying the crucial strand displacement ability necessary for making the concatemers. Also, the heat-stable Vent exo-DNA polymerase may be used. However, the enzyme shown in the literature to be most efficient on acting on circular templates is phi29 polymerase, and this is preferred.

20 Alternatively, concatemerisation takes place by ligating multiple copies of the target molecule to form one continuous single stranded molecule. In this way a circularised target polynucleotide is unnecessary.

Intervening nucleic acid sequences may also be present as shown in figure 6. This is often desirable as the intervening nucleic acid can be of known
25 sequence which will aid the hybridisation to the second polynucleotide.

Hybridisation with mask (second) polynucleotide

The hybridisation to the second polynucleotide can be carried out directly as the concatemer is produced or a single-stranded blocking molecule can be used and removed with an exonuclease after the concatemerisation has been
30 fulfilled. Accordingly, the second polynucleotide can be present during the formation of the concatemer. This is explained below.

The hybridisation of the concatemer with the second polynucleotide can occur simultaneously as the polymerase reaction proceeds on the concatemer.

In this embodiment, the circular target may be attached (ligated) to the second polynucleotide as described above, so that the polymerase product is formed in proximity to the second polynucleotide, aiding hybridisation.

5 In an alternative method, the hybridisation can also be separated from the concatemerisation reaction by "blocking" the second polynucleotide using a complementary molecule. The blocking molecule can either be synthesised in a separate reaction and then annealed to the second polynucleotide prior to the concatemerisation reaction. Alternatively, the blocking molecule can be synthesised by a polymerase directly on the second polynucleotide by using a
10 short primer. After the concatemerisation reaction, the blocking molecule can be removed using an exonuclease, and the second polynucleotide is then available for hybridisation to the concatemerised target molecule and its polymerised product.

The second polynucleotide will have at least partial complementarity to the
15 concatemer. This is achieved either by knowledge of the partial sequence of the units of sequence on the target, or by incorporating complementary sequences into the formed concatemer. The intention is to hybridise the target to the second polynucleotide, such that there are non-hybridised portions which can be interrogated. Alternatively, the method may be carried out by analysing the
20 hybridised portion. For example, the method may be designed such that restriction enzyme substrate sites are formed if perfect hybridisation occurs at selected regions of the second polynucleotide, and treating the duplex with a restriction enzyme. Monitoring any cleavage product will reveal whether the complementary sequence was present on the target, thereby revealing a
25 characteristic of the original molecule. This is shown in Figures 10 to 13.

Preferably, the second polynucleotide is said to comprise first and second sequence units in a predetermined order. The intention is to use either or both of the first and second sequence units to mask sequences on the first polynucleotide, to permit selective interrogation of the first polynucleotide to
30 occur. In one embodiment, the second polynucleotide is used to mask selected units on the first polynucleotide to thereby separate units on the first polynucleotide and permit interrogation to occur more readily. The "unmasked" units of the first polynucleotide may be in an order that represents the order of

units on the original target polynucleotide, or may be used to create a new order of units which represents specific sequences on the target polynucleotide. For example, with reference to Figure 2, the interaction between the first and second polynucleotides is carried out to reveal restriction enzyme sites occurring
5 between perfectly complementary sequences on the first and second polynucleotides. Although hybridisation occurs between non-complementary portions (see sequence in bold), this will not permit a restriction enzyme digestion to occur.

If perfect hybridisation occurs, the restriction enzyme substrate can be
10 cleaved. If one or more mismatches occurs, then there will be no cleavage. some restriction enzymes are able to cleave substrates when one or two mismatches are present. If these enzymes are to be used to characterise the target polynucleotide it is preferred that the second polynucleotide is designed so that two or more mismatches are present.

15 Alternatively, as shown in Figure 3, the interaction between the first and second polynucleotides results in non-hybridised units which can be interrogated by hybridisation with a subsequent oligonucleotide. This is explained in more detail below.

Multiple second polynucleotides may also be used in an addressable
20 array, wherein the concatemer hybridised to several second polynucleotides with subsequent interrogation permitting identification of hybridisation events. This is shown in figure 7 and is explained in more detail below.

Analysing the target polynucleotide

As explained above, the target polynucleotide comprises nucleic acid
25 sequence 'units', which represent particular characteristics (e.g. sequence information on an original molecule. The method of the invention is used to identify the units, to thereby determine the characteristics of the original molecule. The concatemer is used to identify the units at intervals more spaced apart than could be achieved if analysis was performed on the single target
30 polynucleotide. In order to carry this out, it is preferred to 'mask' selectively units on each copy of the target polynucleotide on the concatemer, with the unmasked units being interrogated and identified.

Masking is achieved by use of a second polynucleotide which will mask

(hybridize with) most of each target polynucleotide (address) on the first polynucleotide and then let each address unmask one specific bit position. This is shown in Figure 8. Address 1 will unmask bit position 1, and so on. The procedure for revealing the content of each bit position can then be to ligate the
5 bit with a padlock-probe with, for example a signal moiety, perform a polymerase fill-in reaction with a labeled nucleotide that can be detected. When all of the bit positions are labeled appropriately, the read-out step can be performed to characterise each bit. This procedure is described in more detail below.

Step 1

10 In this example, the target polynucleotide comprises at least 40 "bits" of nucleic acid sequence coding for the native (original) DNA sequence to be analysed. Each bit position can have either of two values: 0 or 1. The two bits possible for each position are at least 10bp long where two bases are unique. These two bases can be positioned anywhere in the bit coding sequence and
15 determine the value of the bit. The remaining bases (at least 8) are preferably identical for the two bit-values at each position, but different for each bit position as illustrated in figure 4.

The mask molecule (second polynucleotide) is a single-stranded molecule containing at least 40 copies of the target polynucleotide or complementary
20 sequence. As each target polynucleotide is different with respect to bit value sequence, the mask (second polynucleotide) does not contain 0 or 1 bit sequences, but "universal bits" differing in one base-pair from the value-coding bits as illustrated in figure 4.

After the target polynucleotide is concatemerised, the polymerase reaction
25 is carried out.

Step 2

The resulting hybrid between the mask and the concatemer consists of double-stranded DNA with one mismatch in every bit sequence. However, due to the design of the mask, bit-position one in target polynucleotide copy one will
30 not hybridise, but leave the bit sequence exposed. Further, bit-position two in target polynucleotide copy two will also not hybridise, exposing the sequence of the bit in position two. This mechanism is used to reveal the sequences of all the bits in the original target polynucleotide as illustrated in figure 8. Also, the

revealed bit values are physically separated by one target polynucleotide length, magnifying the signal chain at least 40-fold. The second polynucleotide may be designed so that there is no hybridisation in the bit (unit) positions to be interrogated. So, for example, in bit position 1 in target 1 the sequence on the
5 second polynucleotide may not have any complementary regions. Similarly, in bit position 2 in target 2 there may be no complementary sequences. This helps separate the bit positions to be interrogated.

Step 3

Interrogation of the resulting hybrid may be carried out using any
10 convenient read-out technique. In one embodiment it is preferred to use padlock technology (Nilsson et al., Science, 1994; 265(5181):2085-8 and Baner et al., Curr.Opin. Biotechnol., 2001; 12(1):11-15). A padlock probe is an oligonucleotide that becomes circularised by DNA ligation in the presence of an appropriate polynucleotide sequence. The reaction requires the two ends of the
15 oligonucleotide to hybridise to adjacent nucleotides on the target for ligation to occur; it is therefore highly specific. As the value of each bit is coded for by two adjacent basepairs these two basepairs can be utilised to perform sequence-specific ligation of padlock probes to the unmasked bits as illustrated in figure 9. If bit position one codes for the value 1 only the 1-specific probe will be able to
20 hybridise its two termini bases and be available for a ligation reaction. If a 0-specific probe is hybridised to a 1-bit, the padlock probe will not be available for a ligation reaction, and hence it can be washed away. When all the bit-positions have been connected to its corresponding padlock probe, the resulting molecule can be read by standard image acquiring systems provided that the 0-specific
25 and 1-specific padlock probes are labelled with 0- or 1-specific fluorophores.

Once the conversion method has been performed a "read-out step" may be carried out to obtain the sequence information encoded within.

The read-out step may be performed using any suitable technique, for example as described in WO-A-00/39333 and WO-A-04/094663.

30 One read-out strategy is to use short detectably-labelled oligonucleotides to hybridise to the units on the converted polynucleotide (first or second polynucleotide) and to detect any hybridisation event. The short oligonucleotides have a sequence complementary to specific units of the converted

polynucleotide. This is shown in Figure 7 (a).

The following example illustrates the invention.

Example

In order to demonstrate the "roll-back" principle a 114nt single-stranded
5 molecule was used as a second polynucleotide substrate and a 38bp circular target molecule. The substrate molecule was immobilized on 1µM streptavidin coated paramagnetic beads using biotin to "anchor" the second polynucleotide.

The target molecule is hybridised to the second polynucleotide and phi29
DNA polymerase is added. The polymerase performs an extension using the
10 target as a template. The extended strand is complementary to the second polynucleotide, and will, according to the roll-back theory, hybridise to the second polynucleotide forming an 114bp double-stranded molecule. Depending on the sequence of the target, the double stranded molecule will contain recognition sites for certain restriction endonucleases (figures 10, 11 and 12):

15 As shown in figure 10, the target with the unit sequence 0100 creates a recognition site for BamH1 in the 2nd bit position in the 2nd target polynucleotide copy. The 2nd bit position in copy 1 and 3 are inactivated by two single basepair mismatch hybridizations in the recognition site. Provided that the roll-back has actually occurred, digestion with BamH1 will yield a 64bp molecule that can be
20 visualized on a gel.

As shown in figure 11, the target with the unit sequence 0010 creates a recognition site for HindIII in the 3rd bit position in the 3rd target copy. The 3rd bit position in copy 1 and 2 are inactivated by a single basepair mismatch in the recognition site. Provided that the rollback has actually occurred, digestion with
25 HindIII will yield a 96bp molecule that can be visualized on a gel.

As shown in figure 12, the target with the unit sequence 0110 creates recognition sites for both BamH1 and HindIII at the 2nd position in the 2nd target copy and the 3rd position in the 3rd target respectively. The other potential restriction sites in the concatemer is inactivated as described for the target 0100
30 and 0010.

Results

Target 0100 in all the experiments a band is visible on the gel at around the 60bp marker (see lane 3, figure 13).

For target 0010, in all the experiments, a band is visible on the gel at around the 100bp marker when digested with HindIII. However, in addition to the expected 96bp band to other bands appear on the gel at around 50 and 40bps. (see lane 4, figure 13). These results suggest that the HindIII digestion is not completely inhibited by a one basepair mismatch. If partial digestion occurs in the 2nd target copy the expected molecules would be 96, 54 and 38. Digestion in the 1st target copy should yield two bands of 38 and 16 basepairs. The fact that we do not see these bands on the gel might indicate that the restriction site is too close to the end of the molecule to cleave to any detectable degree. However, the fact that detection of the 54 and 38 basepair bands is found, indicates that a double-stranded DNA is formed during the rollback process.

For target 0110, in all the experiments, the same restriction pattern is seen when cut with BamH1 and HindIII as seen for Target 0100 and 0010 respectively (see lanes 5 and 6, figure 13).

The content of all the publications referred to herein are incorporated herein by reference.

CLAIMS

1. A method for analysing a target polynucleotide having distinct units of nucleic acid sequence, comprising:
 - 5 (i) forming a first polynucleotide which is a concatemer having multiple repeating target polynucleotide sequences;
 - (ii) hybridising a portion of the first polynucleotide to a second polynucleotide such that the portion hybridised, or the portion not hybridised, corresponds to at least a sequence unit on the target, and determining the
10 sequence unit on the target.
2. A method for the conversion of a target polynucleotide having distinct units of nucleic acid sequence into a polynucleotide having nucleic acid sequences separating one or more of the distinct nucleic acid sequence units, comprising:
 - 15 (i) forming a second polynucleotide having defined first and second sequence units in a predetermined order;
 - (ii) forming on the second polynucleotide, a first polynucleotide made up of a series of target polynucleotides, wherein the order of the first and second sequence units permits the specific interaction between the first
20 and second polynucleotides, such that distinct units of nucleic acid sequence on the second polynucleotide can be distinguished from other distinct units by the extent of interaction; and optionally
 - (iii) determining the sequence and/or order of the distinct sequence units on the basis of the interaction to thereby determine one or more specific
25 sequence units on the target polynucleotide.
3. A method according to claim 1, wherein the concatemer is formed by circularising the target polynucleotide and carrying out a polymerase reaction using the circular target polynucleotide as the template.
4. A method according to claim 3, wherein the circular target
30 polynucleotide comprises additional nucleic acid sequence.
5. A method according to claim 3 or claim 4, wherein the target polynucleotide is circularised by hybridising the target polynucleotide to the second polynucleotide, wherein 5' and 3' regions of the target polynucleotide

hybridise to the second polynucleotide such that the 5' and 3' ends are in proximity, and ligating the 5' and 3' ends.

6. A method according to claim 3 or claim 4, wherein the target polynucleotide is circularised by hybridising an oligonucleotide to the target polynucleotide, wherein the oligonucleotide is complementary to both 5' and 3' regions of the target, such that the 5' and 3' ends are in proximity, and ligating the 5' and 3' ends.
7. A method according to claim 6, wherein the oligonucleotide is removed by an exonuclease prior to the polymerase reaction.
8. A method according to claim 2, wherein step (ii) is carried out by hybridisation between the first and second polynucleotides, wherein the first polynucleotide comprises sequence units which interact with the first sequence units of the second polynucleotide, but not to the second sequence units, and wherein the order of the first and second sequence units is such that the order of non-hybridised sequence units on the first polynucleotide represents a defined order of sequence units on the target polynucleotide.
9. A method according to any preceding claim, wherein the units of nucleic acid sequence on the target polynucleotide are of either a first defined sequence or a second defined sequence.
10. A method according to any preceding claim, wherein each unit on the target polynucleotide comprises at least 4 nucleotides, preferably at least 6 nucleotides, and wherein the first sequence differs from the second sequence by two nucleotides.
11. A method according to claim 2, wherein the second polynucleotide is formed by the rolling-circle polymerase reaction.
12. A method according to any preceding claim, wherein the target polynucleotide comprises at least 40, preferably at least 50, units of nucleic acid sequence.
13. A method according to any of claims 1 to 9, wherein the sequence units on the target are unique to that unit position.
14. A method according to claim 13, wherein the second polynucleotide is designed to hybridise to each target polynucleotide of the first polynucleotide, but not to a predetermined sequence unit on each target, such that, after

hybridisation, there is one sequence unit within one or more target polynucleotides that is not hybridised.

15. A method according to claim 15, wherein the sequence unit in the one or more target polynucleotides that is not hybridised, is determined.

5 16. A method for the formation of a double-stranded polynucleotide, comprising:

(i) forming a first polynucleotide;

(ii) carrying out a rolling circle amplification reaction from a primer molecule attached to the first polynucleotide, wherein the amplification
10 reaction utilises a circular polynucleotide molecule having a sequence that is at least partially complementary to repeating units of the first polynucleotide.

17. A method according to claim 16, wherein the circular polynucleotide is circularised by hybridising a single stranded polynucleotide to the first polynucleotide, wherein the 5' and 3' regions of the single stranded
15 polynucleotide hybridise to the first polynucleotide such that the 5' and 3' ends are in proximity and ligating the 5' and 3' ends.

18. A method according to claim 16, wherein the circular polynucleotide is circularised by hybridising an oligonucleotide to the single stranded form of the polynucleotide, wherein the oligonucleotide is complementary to both 5'
20 and 3' regions of the polynucleotide such that the 5' and 3' ends are in proximity, and ligating the 5' and 3' ends.