

PROBE WITH MULTIPLE TARGET REGION SPECIFICITY AND OF TRIPARTITE CHARACTER

FIELD OF THE INVENTION

The present invention relates to a nucleic acid probe for detecting target nucleic acid sequences, and more particularly to a probe for detecting variant nucleic acid sequences from genomic or amplified DNA. Further aspects of the invention relate to methods for detecting variant nucleic acid sequences.

BACKGROUND TO THE INVENTION

Nucleic acid probes are often used to identify the presence of specific target sequences in genomic or amplified DNA. The annealing or melting temperature of the probe to the target is affected by the length of the complementary region shared by the probe and the target, and by the existence of any mismatches between the otherwise complementary base pairs. This can be used to detect the presence of variants, for example SNPs or multiple repeats. A probe can be designed to have a first melting temperature (T_m) to a wild type sequence, and the annealing of the probe to the target monitored, for example through development of fluorescence on annealing. If the T_m is different to the expected value, then the target sequence includes a variant.

However, this process may be difficult to implement with longer probe regions. In particular, a long probe may not provide adequate variation in T_m between a wild type and a variant sequence.

The present invention is intended to address these and other disadvantages with conventional probes.

US2007/0065847 describes a collection of nucleotide probes labelled with a common sequence to act as a tag. The probes are used for probing libraries of compounds.

WO2011/027966 describes probes having first and second regions, the first region including a reporter molecule and the second region including a quencher molecule. When the probe hybridises to a target, exonucleases can cleave the reporter from the quencher, leading to signal detection. When there is a mismatch such that the reporter portion is not hybridised, the reporter is not cleaved, and there is no signal detection.

WO2007/058499 describes probes for respiratory viruses having first and second hybridisation portions separated by a separation portion.

5 US2007/0259347 describes probes for screening arrays, the probes having two portions separated by a linker where the linker segment is selected to minimise homology noise associated with hybridisation.

10 US2003/0170711 describes oligonucleotide probes including universal bases, including 2-deoxyinosine.

WO2011/087928 describes oligonucleotide probes for detecting KRAS and PIK3CA SNPs in a background of wild type sequences, including polydeoxyinosine linkers.

SUMMARY OF THE INVENTION

15 According to a first aspect of the present invention, there is provided a nucleic acid probe comprising a first nucleic acid sequence being complementary to a first target nucleic acid sequence; a second nucleic acid sequence being complementary to a second target nucleic acid sequence; and a linker nucleic acid sequence joining the first and second nucleic acid sequences; wherein the linker separates the two first and
20 second sequences such that the melting temperature of the first sequence annealed to the first target nucleic acid sequence and of the second sequence annealed to the second target nucleic acid sequence are discrete.

25 The presence of the linker region allows the probes to be split into functional elements that have different hybridisation characteristics. Inclusion of these linkers creates 'bubble' structures, isolating the elements of the probe from a thermodynamic perspective, to provide regions with different binding characteristics. Further, the presence of the linker nucleic acid sequence allows the whole probe to have the characteristics of a single polynucleotide molecule, but to behave as if composed of
30 separate shorter nucleic acid probes. The linker region may fold to form a loop out when the first and second sequences hybridise to their respective target sequences.

In certain embodiments, the probe may be further extended to have further multiple hybridisation regions separated by multiple linkers.

This can be used in a number of different ways. For example, the probe structure allows probing of contiguous regions, where longer probes (for example, a single probe spanning both first and second target regions) would not provide adequate reporting through T_m analysis to differentiate variants.

5

Preferably the linker is a nucleoside linker; more preferably the linker comprises polydeoxyribonucleotides; most preferably the linker comprises or consists of polydeoxyinosine. Deoxyinosine has a low melting temperature relative to natural bases due to weaker hydrogen bonding. Other nucleosides may be used.

10

Preferably the linker is up to 5, 10, 15, 20, 30, 40, 50 nucleotides in length.

15

At least one of the first and second nucleic acid sequences is a reporter region. A reporter region includes a labelled moiety; preferably a fluorescent label. This allows detection of the probe in the event of binding to a target sequence, and monitoring of annealing over a temperature range in order to determine the presence of any variant target sequences. The probe preferably does not comprise a quencher moiety, nor is the label intended to be used with a quencher. Suitable labels include FAM, TET, HEX, ROX, TAMRA, Cy3, and Cy5. Other suitable labels will be known to the skilled person.

20

Preferably the label is incorporated on to a T nucleotide, although any suitable nucleotide may be used..

25

The reporter region is preferably 15-200 nt in length, more preferably 15-150, more preferably still 15-100, or 20-100, 30-80, 40-60, or around 50 nt in length.

30

The reporter region is designed to have a first T_m in respect of an entirely complementary target sequence, and an altered, preferably lower, T_m , in respect of a variant target sequence. For example, the variant may represent a SNP (single nucleotide polymorphism), an insertion, a deletion, an inversion, or a repeated sequence.

In this embodiment, the melting temperature of the first sequence annealed to the first target nucleic acid sequence may be different to that of the second sequence annealed to the second target nucleic acid sequence. Preferably, however, both first and second

nucleic acid sequences are selected to have similar T_m to their respective target sequences.

5 The reporter region may further comprise a blocking region; that is, a portion which serves to block extension of the nucleic acid strand by DNA polymerase, so preventing strand extension during, for example, PCR. A polymerase enzyme blocking group is one which should have the functional properties of blocking further elongation of the polymer. A blocking group may be any chemical group which can be attached to a nucleotide which will allow the 5' end of the modified nucleotide to attach to a 3' end of
10 another nucleotide in a DNA chain but will not allow attachment of a nucleotide to the 3'hydroxyl group of the modified nucleotide. Suitably, the absence of an OH group in the 3' position will prevent further elongation by polymerase activity. In a particularly preferred embodiment, the blocking group is selected from acetyl, CH_3 , glycyl, leucyl and alanyl groups. In another embodiment, the blocking group may be in the form of a
15 di or tri peptide.

In an embodiment of the invention, both the first and second nucleic acid sequences are reporter regions. They may include different labels. Such a probe may be used as a multiplex reporter, allowing detection of target sequences over an extended range with
20 a single probe.

In this embodiment, the probe is split into multiple discrete reporter regions. Each reporter region has different annealing temperatures and has 1 or more fluorescent nucleotides, preferably FAM-T, or different / multiple colours. The reporter may be used
25 to report the presence of specific sequences or sequence variants, SNPs, insertions, deletions, etc. This allows multiple sequences over an extended range to be detected with a single probe. Each region is tuned to have a similar T_m in the case of the wild type target sequence, but a shifted T_m in the case of a mutation; this means that a user only has to detect the shifted T_m to know the variant is present.

30

As an alternative to detecting multiple target sequences, the probe may be structured for example to provide a control binding region and a test binding region, to improve the reliability of an assay.

In an alternative embodiment of the invention, the second nucleic acid sequence may be a reporter region, and the first nucleic acid sequence is an anchor region. In this embodiment, the melting temperature of the first sequence annealed to the first target nucleic acid sequence is higher than that of the second sequence annealed to the second target nucleic acid sequence. The linker sequence may form a loop-out region or may hybridise to any inter-probe nucleic acid sequence. The T_m of the anchor region to the target sequence is significantly higher (at least 0.5, 1, 1.5°C but more generally up to 5°C to 20°C higher than the T_m of the reporter region with the target sequence.

The anchor region is preferably at least 50 nt in length, more preferably at least 60, 70, 80, 90, 100, 125, 150 nt. The anchor region is preferably no more than 200 nt in length and no less than 20 nt in length.

In use, as the anchor region has a substantially higher annealing temperature it serves to anchor the probe to the correct region. The reporter region has a lower annealing temperature to the anchor region and has 1 or more fluorescent nucleotides, preferably FAM-T. The reporter is used to report the presence of a specific sequence or sequence variants, SNPs, insertions, deletions, etc.

In preferred embodiments, the probe is used to detect size/length polymorphisms e.g. STR profiling, Fragile X, etc. The anchor region is used to anneal to the downstream of the repeat region, while the reporter region is homologous to a short stretch of repeats. Since the reporter is isolated from the main anchor region, and since the fluorophores are only on the reporter region, it makes no difference if the reporter binds at different sites to cause loop outs in the amplicon - normally this would alter the T_m of the product, making measurement difficult. In certain embodiments, the anchor region may be designed to be homologous to a short stretch of the repeat region as well as a flanking region; for example, the anchor region may include up to 50 nt of repeat sequence and a longer stretch, eg 150 nt, of flanking sequence.

A further aspect of the invention provides a nucleic acid probe comprising a first anchor nucleic acid sequence being complementary to a first target nucleic acid sequence; a second reporter nucleic acid sequence being complementary to a second target nucleic acid sequence; and a linker nucleic acid sequence joining the first and second nucleic

acid sequences; wherein the linker separates the first and second sequences such that the melting temperatures of the first and second sequences hybridised to their respective target sequences is discrete, and where the melting temperature of the anchor sequence annealed to the first target nucleic acid sequence is higher than that of the reporter sequence annealed to the second target nucleic acid sequence; and wherein each reporter sequence comprises at least one detectable label.

The reporter preferably comprises a plurality of repeated sequence units. Preferably at least 3, 4, or 5 repeated units.

The linker preferably comprises polydeoxyinosine.

The anchor may comprise one or more repeated sequence units, but preferably comprises unique sequence. The anchor is preferably from 50 – 200 nt in length.

The probe preferably further comprises a blocking sequence to prevent strand extension during PCR.

The probe may further comprise one or more additional linker and reporter sequences.

A yet further aspect of the invention provides a nucleic acid probe comprising a first reporter nucleic acid sequence being complementary to a first target nucleic acid sequence; a second reporter nucleic acid sequence being complementary to a second target nucleic acid sequence; and a linker nucleic acid sequence joining the first and second nucleic acid sequences; wherein the melting temperature of the first reporter sequence annealed to the first target nucleic acid sequence is similar or equal to that of the second reporter sequence annealed to the second target nucleic acid sequence; and wherein each reporter sequence comprises at least one detectable label.

The probe preferably further comprises a blocking sequence to prevent strand extension during PCR.

A further aspect of the invention provides a method for detecting one or more polymorphisms in a target nucleic acid sequence, the method comprising:

providing a probe according to an aspect of the present invention;

contacting the probe with a test nucleic acid sequence, the test sequence having a first target nucleic acid sequence and a second target nucleic acid sequence;

allowing the first nucleic acid sequence of the probe to hybridise to the first target nucleic acid sequence; and the second nucleic acid sequence of the probe to hybridise to the second target nucleic acid sequence; and

determining the T_m of the first probe sequence hybridised with the first target sequence; and determining the T_m of the second probe sequence hybridised with the second target sequence;

wherein the T_m is indicative of the presence of a wild type sequence, or of a variant sequence.

This method may also be used to improve the utility of melting temperature whereby the T_m of a first reporting region acts as a reference to measure the T_m of the variant sequence region. Generally the position of the recorded melting temperature which is useful in determining the underlying sequence, may be affected by the concentration of salts in the reaction mixture. By comparing the melting position of the first reporter region as a known sequence, it is possible to more accurately determine the sequence under the second reporter sequence since both regions will be equally affected by interfering substances such as external salts.

20

Thus, a further aspect of the invention provides a method for detecting one or more polymorphisms in a target nucleic acid sequence, the method comprising:

providing a probe according to an aspect of the present invention;

contacting the probe with a test nucleic acid sequence, the test sequence having a first target nucleic acid sequence and a second target nucleic acid sequence;

allowing the first nucleic acid sequence of the probe to hybridise to the first target nucleic acid sequence; and the second nucleic acid sequence of the probe to hybridise to the second target nucleic acid sequence; and

determining the T_m of the first probe sequence hybridised with the first target sequence; and determining the T_m of the second probe sequence hybridised with the second target sequence;

comparing the determined T_m of the first probe sequence with the expected T_m of the first probe sequence;

normalising the determined T_m of the second probe sequence based on the difference of the determined and expected T_m of the first probe sequence;

wherein the normalised T_m of the second probe sequence is indicative of the presence of a wild type sequence, or of a variant sequence.

5 A still further aspect of the invention provides a method for detecting tandem repeats in a target nucleic acid sequence, the method comprising:

providing a probe according to an aspect of the present invention, wherein the reporter sequence comprises a number of tandem repeat units;

10 contacting the probe with a test nucleic acid sequence, the test sequence having a first target nucleic acid sequence and a second target nucleic acid sequence, wherein the first target sequence is complementary to the probe anchor sequence, and the second target sequence comprises a number of tandem repeats complementary to the tandem repeat units of the reporter sequence;

15 allowing the anchor nucleic acid sequence of the probe to hybridise to the first target nucleic acid sequence; and the second nucleic acid sequence of the probe to hybridise to the second target nucleic acid sequence; and

detecting binding or otherwise of the reporter sequence to the test nucleic acid sequence.

20 Binding may be detected by determining the T_m of the reporter sequence to the test nucleic acid sequence, and/or by detecting label (eg, fluorescence) from the reporter sequence. Binding indicates that the number of repeats in the target sequence is greater than or equal to the number of repeats in the reporter sequence; if there are more repeats in the reporter sequence, then there is either no binding or reduced binding. In certain embodiments, the T_m may be considered to be generally
25 proportional to the number of repeats in the target sequence, and the T_m may be determined in order to determine the number of repeats. Alternatively, or in addition, the reporter sequence of the probe may include one or more labelled moieties in each repeat unit; in this way, the detected label may be taken as being proportional to the number of repeats in the target sequence.

30

BRIEF DESCRIPTION OF THE DRAWINGS

These and other aspects of the present invention will now be described by way of example only and with reference to the accompanying drawings, in which:

35 Figure 1 shows a first probe architecture in accordance with an embodiment of the invention;

Figures 2 to 5 show alternative duplex structures formed through binding of the probe of figure 1 to a target sequence;

Figure 6 shows a second probe architecture in accordance with an alternative embodiment of the invention;

5 Figure 7 shows common mutant codons in the *M. tuberculosis* *rpoB* gene;

Figure 8 shows the sequence of a portion of the *M. tuberculosis* *rpoB* gene, together with the sequences of certain mutant genes and the sequence of a probe in accordance with the invention for detecting the presence of such mutants;

Figure 9 shows alternative duplex structures formed through binding of the probe of

10 Figure 8 to target sequences;

Figure 10 shows example melt curves obtained from the duplexes of Figure 9;

Figure 11 shows a sequence alignment of a portion of the vaccinia genome with that of other pox viruses;

Figure 12 shows a probe for detection of the pox viruses;

15 Figure 13 compares sensitivity of the linker probe of Figure 12 with that of a conventional vaccinia probe;

Figure 14 shows melt curve analysis of the linker probe of Figure 12 and a conventional probe with various amplified pox virus genomic sequences;

20 Figure 15 shows primer sequences and amplification conditions for amplifying the pox virus fragment used in Figure 14; and

Figure 16 shows primer and probe sequences, and melt curve analysis, for a short tandem repeat in the human tyrosine hydroxylase gene.

DETAILED DESCRIPTION OF THE DRAWINGS

25 The present invention relates to a nucleic acid probe which is split into two or more functional elements, separated by a nucleic acid linker. The linker is preferably deoxyinosine, which has a low T_m relative to the naturally occurring DNA bases. Other suitable linkers may be used.

30 Figure 1 shows schematically a first embodiment of a probe according to the present invention. The probe includes an anchor region, of around 100 – 200 nt in length, separated by a deoxyinosine linker of 5 nt in length from a reporter region, of from 15 – 200 nt in length. The anchor region is designed to be complementary to a first target region which is downstream from a repeated sequence region in a target nucleic acid.

35 The reporter region is complementary to the repeated sequence region, and includes a

known number of repeat units. The reporter region also includes one or more labelled nucleotides, preferably FAM-T. In certain embodiments, each repeat unit includes a labelled nucleotide.

5 The probe sequences are designed such that the T_m of the anchor region to the target is significantly higher than the T_m of the reporter to the target, which in turn is significantly higher than the T_m of the linker to any of the sequences in the target nucleic acid.

10 Figures 2 to 5 show various different scenarios for the binding of the probe to a target. In these scenarios, the probe is being used to determine size/length polymorphisms e.g. STR profiling, Fragile X, etc. The anchor region is used to anneal the probe to the downstream of the repeat region, while the reporter region is homologous to a short stretch of repeats. Since the reporter is isolated from the main anchor region, and since
15 the fluorophores are only on the reporter region, it makes no difference if the reporter binds at different sites to cause loop outs in the amplicon - normally this would alter the T_m of the product, making measurement difficult.

Where the length of the target repeated sequence region is less than that of the
20 reporter region (Figure 2), then the reporter region does not bind to the target, and there is no fluorescence from the label, and no melt curve obtained by measuring T_m of the reporter. However, as the anchor region has bound to the target, this can be measured (eg, by determining T_m), in order to confirm that the probe does bind and that the repeated sequence region is absent.

25 Where the length of the target repeated sequence region is equal to or longer than the length of the reporter region, then any of the scenarios in Figures 3 to 5 may occur. In Figure 3, the reporter is kept small to 1, 2, 3, 4 or 5 repeat lengths which limits the impact of any loop-out effects on the T_m of the reporter:amplicon duplex. Examples of
30 use are for large repeat structures such as Fragile X where it is important to 'band' the number of repeats; e.g. < 50, > 100, > 200. Alternatively, a small loop-out may occur (Figure 4); again, since the reporter is small, there is minimal effect on the T_m of the reporter:amplicon duplex, and the presence of at least the number of repeats on the probe can be determined.

35

If the reporter length is increased to accommodate multiple repeats (Figure 5), then the T_m of the reporter amplicon duplex will be proportional to the number of repeats. Unbound portions of the reporter will not fluoresce due to self-quenching of the label. One or two nucleotides of each repeat in the reporter are labeled. This methodology
5 may be important in STR analysis where you are interested in the number of 3, 4 or 5bp repeats up to 15x or 20x.

An alternative probe architecture is shown in Figure 6. This probe is a multiplexed reporter, and includes two separate reporter regions joined by a linker. Each reporter
10 region has different annealing temperatures and has 1 or more fluorescent nucleotides, preferably FAM-T, or different / multiple colours. The reporter is used to report the presence of a specific sequence or sequence variants (eg, SNPs, insertions, deletions, etc). This allows multiple sequences over an extended range to be detected with a single probe. Each region is tuned to have a similar – but not identical – T_m in the case
15 of the wild type sequence, but a shifted T_m in the case of a mutation so that a user only has to detect the shifted T_m to know the variant is present. By “similar” is meant that the T_m differs by at most 2, 1.5, 1, 0.5 deg C.

The probe could also be structured for example to provide a control binding region and
20 test binding region.

An example of use of a multiplexed reporter probe to detect variants in the *Mycobacterium tuberculosis* *rpoB* gene is now given. Multi drug resistance in *M. tuberculosis* is complex. Rifampin is a first line *M. tuberculosis* medication and is the
25 main target to identify in the field prior to treatment. Rifampin resistant *M. tuberculosis* have mutations in the 81-bp core region of the *rpoB* gene, which encodes the β -subunit of RNA polymerase. 96% of Rifampin resistant clinical isolates of *M. tuberculosis* have mutations in this gene. Mutations in codons 516, 526, or 531 result in high level Rifampin resistance. However, detecting mutations across an 81-bp gene region would
30 typically require multiple conventional probes, several of which would need to overlap, so requiring multiple detection steps.

Figure 7 shows wild type and variant codons in a section of the *rpoB* gene; note that 9 codons are shown from a 21 codon (63 nt) region. The present inventors have
35 designed a probe (shown in Figure 8, labeled PROBE) to detect and report the

presence of the any of the potential *M. tuberculosis* genotypes with mutations in the main resistance codons 516, 526, or 531.

Figure 8 also shows the wild type (WT) sequence of the *rpoB* gene, together with three mutant sequences (MUT516, MUT526, and MUT531). The boxed regions of the WT sequence are those to which the probe is designed to hybridise. The probe itself includes a first reporter sequence spanning codons 531 to 525, including potential mutant codons 526 and 531, a linker comprising five polydeoxyinosine residues (indicated as "I"), and a second reporter sequence spanning codons 518 to 510, including potential mutant codon 516. Each of the reporter sequences includes two labeled T residues (highlighted). The second reporter sequence is also adjacent a PCR blocker moiety.

The probe binding to WT sequence in both positions gives the same T_m of 60C for each reporter sequence. The presence of mutant sequences shifts this to 52C-56C.

Figure 9 shows the possible binding patterns of the probe to target sequence. Where both reporter sequences bind to the wild type (WT:WT), the T_m of the first reporter (T_{m1}) equals the T_m of the second reporter (T_{m2}). A single mutant in the first reporter (MUT:WT) lowers the T_m of the first reporter to T_{m3} , which is lower than T_{m2} . A single mutant in the second reporter (WT:MUT) lowers the T_m of the second reporter to T_{m4} , which is lower than T_{m1} . A single mutant in each reporter (MUT:MUT) shifts the respective T_m s to T_{m3} and T_{m4} . Finally, the presence of a second mutant in the first reporter target (MUT:WT) lowers the T_m to T_{m5} , which is lower than T_{m3} .

Each of these T_m s is distinguishable from one another, and so the single probe may be used to determine the presence of three distinct mutations in the *rpoB* gene. Figure 10 gives representative melt curves from such experiments. The presence of wild type of mutant sequences can be clearly distinguished. Thus, this probe architecture permits simple multiplex detection of multiple mutations across a relatively long stretch of genome.

A further illustrative use of such a probe is shown in Figures 11 to 15, which illustrate use of a probe for detection of the Vaccinia virus, and distinguishing this virus from related pox viruses.

Figure 11 compares a 30 nucleotide portion of the genomic sequence of Vaccinia with that of the monkeypox, cowpox, ectromelia, and camelpox viruses. It will be seen that there are eight bases which may differ across this sequence.

5

Figure 12 illustrates a probe for detection of this region of the virus genome and distinguishing the various viruses from one another. The probe includes a first reporter region separated from a second reporter region by a sequence of five inosine residues. The first reporter region includes a trimethoxystilbene cap. Each reporter region includes two fluorescently labelled T residues.

10

The probe sequence is:

5'-GAG T*AT TTG T*CA TTT IIII TAT ATT T*GT TGG CT*G-3' : SEQ ID NO: 1

Capped with a 5' trimethoxystilbene, and a 3' phosphate. I = inosine. T* = fluoresceine labelled T.

15

The two reporter regions span the 30 nucleotide sequence shown in Figure 11, and both cover a number of potential variation sites.

Figure 13 compares the sensitivity of the probe with that of a standard vaccinia probe spanning the same 30 nt sequence. The linker probe is significantly more sensitive at the same copy number. To compare the sensitivity, a portion of the genome is amplified using the primers and conditions indicated in Figure 15, and then a melt curve assay conducted using the relevant probe on the amplicon.

20

Figure 14 shows melt curves for the vaccinia probe ("linker probe") and a conventional vaccinia probe ("standard probe") when hybridised to different pox virus target sequences. The range of T_m spanned by the linker probe is much smaller, but the linker probe provides improved discrimination between the various pox viruses, and allows the five viruses to be clearly identified and distinguished using a single probe. The use of a linker probe compared with a conventional probe lowers the T_m of the product, and lowers the T_m range to enable multiplexing higher in the range. The probes of the invention allow up to five base mismatches over a short region of 30 bases to be identified, so allowing identification of the pox viruses.

25

30

35

Figure 16 shows primer and probe sequences for analysing a short tandem repeat (STR) in the human tyrosine hydroxylase gene.

The forward and reverse primers are:

5 TH01-FWD

ATT CAA AGG GTA TCT GGG CTC TGG : SEQ ID NO: 2

TH01-REV

GTG GGC TGA AAA GCT CCC GAT TAT: SEQ ID NO: 3

10

The TH gene includes a short tandem repeat sequence of (AATG)_n, which is known to be polymorphic. To confirm that the present invention is capable of distinguishing between different lengths of STR, two probes were compared for sensitivity and accuracy. Both probes were of the same sequence, probe 1 was prepared without a linker, while probe 2 includes a five inosine base sequence acting as a linker between the two portions of the probe. Probe sequences were:

15

TH01-Probe1

20 M ACAGACTCCAFG G~~FF~~GAATGAA~~FF~~GAATGAGGGAAATAAGGG AGGA P: SEQ ID NO: 4

TH01-Probe2

M ACAGACTCCAFG G~~FF~~GAATGAA~~FF~~GAATGA***** GGGAAATAAGGG AGGA P: SEQ ID NO: 5

25

(M = trimethoxystilbene; P = phosphate; * = inosine; F = fluoresceine labelled T).

30

After amplifying the target sequence with the primers, melt curve analysis was performed with the amplicon and either probe 1 or probe 2. The curves are shown in Figure 16. Although both probes are capable of distinguishing between 3, 6, 9, and 12 copies of the repeat, probe 2, with the linker, provides discrimination between different numbers of repeats over a much narrower temperature range.

CLAIMS:

- 5 1. A nucleic acid probe comprising a first nucleic acid sequence being complementary to a first target nucleic acid sequence; a second nucleic acid sequence being complementary to a second target nucleic acid sequence; and a linker nucleic acid sequence joining the first and second nucleic acid sequences; wherein the linker separates the two first and second sequences such that the melting temperature of the first sequence annealed to the first target nucleic acid sequence and of the second
10 sequence annealed to the second target nucleic acid sequence are discrete..
2. The probe of claim 1 wherein the linker comprises a polydeoxyribonucleotide.
3. The probe of claim 1 or 2 wherein the linker comprises or consists of
15 polydeoxyinosine.
4. The probe of any preceding claim wherein the linker is up to 5, 10, 15, 20, 30, 40, 50 nucleotides in length.
- 20 5. The probe of any preceding claim wherein at least one of the first and second nucleic acid sequences is a reporter region including a labelled moiety; preferably a fluorescent label.
6. The probe of any preceding claim wherein the probe does not comprise a quencher
25 moiety.
7. The probe of claim 5 wherein the reporter region is preferably 15-200 nt in length, more preferably 15-150, more preferably still 15-100, or 20-100, 30-80, 40-60, or
30 around 50 nt in length.
8. The probe of claim 5 or 7 wherein the reporter region is designed to have a first T_m in respect of an entirely complementary target sequence, and an altered, preferably lower, T_m , in respect of a variant target sequence.

9. The probe of any preceding claim further comprising a blocking region which serves to block extension of the nucleic acid strand by DNA polymerase.

5 10. The probe of any preceding claim wherein both the first and second nucleic acid sequences are reporter regions.

10 11. The probe of claim 10 wherein both reporter regions are selected to have a similar T_m in the case of binding to the wild type target sequence, and reduced T_m s in the case of binding to a mutant target sequence; preferably the reduced T_m for the first reporter region differs from that of the second reporter region.

15 12. The probe of any of claims 1 to 9 wherein the second nucleic acid sequence is a reporter region, and the first nucleic acid sequence is an anchor region, wherein the melting temperature of the anchor region annealed to the first target nucleic acid sequence is higher than that of the second sequence annealed to the second target nucleic acid sequence.

20 13. The probe of claim 12 wherein the T_m of the anchor region : target sequence duplex is significantly higher (at least 0.5, 1, 1.5, but more generally up to 5 to 20 deg C higher) than the T_m of the reporter region to the target sequence.

14. The probe of claim 12 or 13 wherein the anchor region is at least 50 nt in length, more preferably at least 60, 70, 80, 90, 100, 125, 150 nt.

25 15. The probe of claim 12, 13, or 14 wherein the anchor region is no more than 200 nt in length.

30 16. The probe of any of claims 12 to 15 wherein the reporter region comprises tandem repeat sequences.

17. The probe of any of claims 12 to 16 wherein the anchor region comprises tandem repeat sequences and unique sequences.

35 18. The probe of any of claims 1 to 11 wherein the first and second target nucleic acid sequences are found in the M tuberculosis genome.

19. The probe of any of claims 1 to 11 wherein the first and second target nucleic acid sequences are found in the vaccinia pox virus genome.

5 20. The probe of any preceding claim wherein the first and second target sequences are within 200, 150, 100, 50, 40, 30, 20, 15, 10, 5, 4, 3, 2, 1 nucleotides of one another in a genomic sequence.

10 21. A nucleic acid probe comprising a first anchor nucleic acid sequence being complementary to a first target nucleic acid sequence; a second reporter nucleic acid sequence being complementary to a second target nucleic acid sequence; and a linker nucleic acid sequence joining the first and second nucleic acid sequences; wherein the linker separates the first and second sequences such that the melting temperatures of the first and second sequences hybridised to their respective target sequences is
15 discrete, and where the melting temperature of the anchor sequence annealed to the first target nucleic acid sequence is higher than that of the reporter sequence annealed to the second target nucleic acid sequence; and wherein each reporter sequence comprises at least one detectable label.

20 22. A nucleic acid probe comprising a first reporter nucleic acid sequence being complementary to a first target nucleic acid sequence; a second reporter nucleic acid sequence being complementary to a second target nucleic acid sequence; and a linker nucleic acid sequence joining the first and second nucleic acid sequences; wherein the melting temperature of the first reporter sequence annealed to the first target nucleic
25 acid sequence is similar or equal to that of the second reporter sequence annealed to the second target nucleic acid sequence; and wherein each reporter sequence comprises at least one detectable label.

30 23. A nucleic acid probe according to any preceding claim further comprising one or more further linker sequences, and one or more further nucleic acid sequences for hybridisation to one or more further target sequences.

24. A method for detecting one or more polymorphisms in a target nucleic acid sequence, the method comprising:
35 providing a probe according to any of claims 1 to 23;

contacting the probe with a test nucleic acid sequence, the test sequence having a first target nucleic acid sequence and a second target nucleic acid sequence;

allowing the first nucleic acid sequence of the probe to hybridise to the first target nucleic acid sequence; and the second nucleic acid sequence of the probe to hybridise to the second target nucleic acid sequence; and

determining the T_m of the first probe sequence hybridised with the first target sequence; and determining the T_m of the second probe sequence hybridised with the second target sequence;

wherein the T_m is indicative of the presence of a wild type sequence, or of a variant sequence.

25. A method for detecting tandem repeats in a target nucleic acid sequence, the method comprising:

providing a probe according to any of claims 12 to 17, the probe comprising a reporter sequence having a number of tandem repeat units;

contacting the probe with a test nucleic acid sequence, the test sequence having a first target nucleic acid sequence and a second target nucleic acid sequence, wherein the first target sequence is complementary to the probe anchor sequence, and the second target sequence comprises a number of tandem repeats complementary to the tandem repeat units of the reporter sequence;

allowing the anchor nucleic acid sequence of the probe to hybridise to the first target nucleic acid sequence; and the second nucleic acid sequence of the probe to hybridise to the second target nucleic acid sequence; and

detecting binding or otherwise of the reporter sequence to the test nucleic acid sequence.

26. A method for detecting one or more polymorphisms in a target nucleic acid sequence, the method comprising:

providing a probe according to any one of claims 1 to 23;

contacting the probe with a test nucleic acid sequence, the test sequence having a first target nucleic acid sequence and a second target nucleic acid sequence;

allowing the first nucleic acid sequence of the probe to hybridise to the first target nucleic acid sequence; and the second nucleic acid sequence of the probe to hybridise to the second target nucleic acid sequence; and

determining the T_m of the first probe sequence hybridised with the first target sequence; and determining the T_m of the second probe sequence hybridised with the second target sequence;

5 comparing the determined T_m of the first probe sequence with the expected T_m of the first probe sequence;

 normalising the determined T_m of the second probe sequence based on the difference of the determined and expected T_m of the first probe sequence;

 wherein the normalised T_m of the second probe sequence is indicative of the presence of a wild type sequence, or of a variant sequence.

10

27. The method of any one of claims 24 or 26 wherein the variant sequence is a deletion, a repeat, a SNP, or an insertion.