PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12Q 1/68	A1	(11) International Publication Number: **WO 00/23620**
		(43) International Publication Date: 27 April 2000 (27.04.00)

(21) International Application Number: PCT/NL99/00643

(22) International Filing Date: 18 October 1999 (18.10.99)

(30) Priority Data:
98203481.1 16 October 1998 (16.10.98) EP

(71) Applicant *(for all designated States except US)*: KEYGENE N.V. [NL/NL]; Agro Business Park 90, NL–4708 PW Wageningen (NL).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: VERBAKEL, Henricus, Maria [NL/NL]; De Esdoorns 18, NL–5427 TK Boekel (NL). SEGERS, Hubertus, Mechtilda, Petrus [NL/NL]; Heldenseweg 22, NL–5995 RP Kessel (NL). VAN EIJK, Michael, Josephus, Theresia [NL/NL]; Tolsteegplantsoen 43, NL–3523 AN Utrecht (NL). SCHOUTEN, Johannes, Petrus [NL/NL]; Hudsonstraat 68, NL–1057 SN Amsterdam (NL). CHAN, Yin–Lay [NL/NL]; Anna Polaklaan 19, NL–1187 PW Amstelveen (NL).

(74) Agent: OTTEVANGERS, S., U.; Vereenigde Octrooibureaux, Nieuwe Parklaan 97, NL–2587 BN The Hague (NL).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: METHOD FOR THE GENERATION OF DNA FINGERPRINTS

(57) Abstract

A process for the analysis of nucleic acids comprising the following steps: fragmenting DNA to produce a collection of fragments with at least three different types of ends; ligating at least two adapters to the different types of ends; providing at least a first oligomer primer complementary to the first adapter and a second oligomer complementary to the second adapter; amplifying said DNA fragments, whereby at least one type of DNA end does not take part in the amplification reaction; analysis of the amplified products.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

Title: Title: Method for the generation of DNA fingerprints.

The present invention relates to nucleic acid amplification methods, nucleic acid detection methods and probes for the detection of nucleic acids. The invention more specifically relates to methods for the generation of DNA fingerprints for the identification and typing of nucleic acid in a sample. The DNA genome of living organisms (except viruses) consists of a sequence of 500.000 - 5.000.000.000 or more building blocks (bases). In order to detect (small) differences between different genomes, a limited fraction of this complex DNA has to be highlighted.
In case no sequence information of the genome is known, a usual starting point for the generation of DNA fingerprints is the fragmentation of the DNA preparation in smaller parts by bacterial enzymes (restriction-endonucleases).
Subsequently a limited number of the DNA fragments obtained is highlighted, either by hybridisation with a labelled probe or by amplification of a subset of the fragments, e.g. with the use of the polymerase chain reaction (PCR).

PCR has been used extensively during the last years to amplify simultaneously *different fragments* of the chromosomal DNA from a organism, in order to obtain a "fingerprint" of the particular genome. In these cases it is often not important which fragments are amplified, provided that the reaction is completely reproducible and the number of fragments that are amplified are roughly controllable. These DNA fingerprints are used to show similarities and differences between DNA samples obtained from different individuals, races or species. Some of the techniques used are RAPD (Random Amplified Polymorphic DNA) which suffers from insufficient reproducibility between different laboratories and AFLP (Amplified Fragment Length Polymorphism, Vos et al 1995).

The AFLP technique involves four steps : 1) Fragmentation of the DNA by restriction-endonucleases. 2) Ligation of adapters to the ends of the DNA fragments. 3) Amplification of a subset of the fragments by a set of carefully selected oligonucleotides, one of which is labelled, e.g. with a fluorescent tag. 4) Visualisation of the amplified fragments by gel-electrophoresis and a suitable detection method. For fragmentation of the genome in a reproducible way, the AFLP method uses restriction-endonucleases that generate a single stranded tail of usually 2 - 4 bases on each fragment. Using a DNA ligase enzyme, a small DNA fragment (adapter) can be attached to these single strands, provided that the adapter also has a single stranded part and that the two single stranded parts are complementary each other. These adapters have a known DNA sequence. Apart from the adapter sequence, also the sequence recognised by the restriction enzyme is known. This information can be used to design primers for the amplification of all fragments starting with an adapter and followed by the restriction endonuclease site. In order to limit the number of fragments amplified, some random bases can be added to the 3' end of one or both PCR primers (selective bases). Adding a first selective nucleotide results in a primer that is not completely complementary anymore to the adapter and/or adjacent restriction enzyme site. Instead, depending on the code of the first nucleotide following the restriction enzyme site, the primer will have a mismatch in three of the four fragments. Thus the number of fragments preferentially amplified is reduced by four. For each additional selective base added to the three prime end of the primer, the number of fragments preferentially amplified with this primer decreases by a further factor of four.

A drawback of the AFLP method is that it usually requires two consecutive cycles of PCR for the analysis of complex genomes. The technique is also rather sensitive to changes in

protocol, preventing the replacement of the expensive Mse 1 enzyme by other restriction endonucleases. Another drawback of the AFLP technique as it is currently used, is the sensitivity to minor impurities in the DNA preparations. Another drawback of the AFLP technique is that the number of bands obtained with complex genomes often cannot be limited to less than approx. 30 as the number of selective bases used cannot be increased to more than three without the risk of extra bands due to mismatching of misalignment between PCR primers and template. For certain applications a further reduction in the number of bands is desirable.

In the present invention we provide improved methods for DNA-fingerprinting. Limiting the number of fragments that are amplified is one of the features of the invention which is important to 1) improve the reproducibility of DNA-fingerprinting, 2) reduce the sensitivity of DNA-fingerprinting techniques to impurities in the sample, 3) reduce the problems associated with two consecutive cycles of PCR and 4) limiting the number of fragments detected thus improving interpretation of the result
The present invention, called DNA-PARTIAL ADAPTER LIGATION-SELECTIVE AMPLIFICATION, has proven to be much less sensitive to changes in protocol and to impurities in the DNA preparations. In a preferred embodiment at least one restriction endonuclease more is used than in the AFLP technique, the number of suitable enzymes is large permitting the choice of cheap and stable enzymes. Using a preferred embodiment of the DNA-PARTIAL ADAPTER LIGATION-SELECTIVE AMPLIFICATION invention, discussed below, the number of efficiently amplified DNA fragments is reduced 20-fold compared to the AFLP technique, which limits the number of selective nucleotides needed on the primers used and permits a further reduction of the number of bands amplified.

BRIEF DESCRIPTION OF THE DRAWINGS.

Figure 1 shows a graphic outline of the general selective amplification mechanism of the DNA-PARTIAL ADAPTER LIGATION-SELECTIVE AMPLIFICATION invention.

Figure 2 shows a 6 % polyacrylamide gel with DNA fingerprints obtained by applying the DNA-PARTIAL ADAPTER LIGATION-SELECTIVE AMPLIFICATION invention to E. coli genomic DNA.

Figure 3 shows a 6 % polyacrylamide gel with DNA fingerprints obtained by applying the DNA-PARTIAL ADAPTER LIGATION-SELECTIVE AMPLIFICATION invention to different Cyncta individuals.

Figure 4 shows a 6 % polyacrylamide gel with DNA fingerprints obtained by applying the DNA-PARTIAL ADAPTER LIGATION-SELECTIVE AMPLIFICATION invention to different Lycopersicum strains.

Figure 5 shows a 6 % polyacrylamide gel with a DNA fingerprint obtained by applying the DNA-PARTIAL ADAPTER LIGATION-SELECTIVE AMPLIFICATION invention to human genomic DNA.

Figure 6 shows a detail of Figure 5.

Figure 7 Two possible ways to amplify part of the sequence of an individual AFLP marker using an internal amplification primer P1 or P2. In this figure, the AFLP marker is an EcoRI +2/ MseI +3 AFLP marker but the general principle applies to other restriction enzyme combinations and numbers of selective nucleotides in the AFLP primer(s).
EcoRI +2 refers to an AFLP primer with 2 selective bases, MseI +3 refers to an MseI AFLP primer with three selective bases, internal P1 refers to an internal PCR primer in forward direction, and internal P2 refers to an internal PCR primer in the reverse direction

Figure 8 Schematic presentation of a result of an AFLP sequence tagging.

DETAILED DESCRIPTION OF THE DNA-PARTIAL ADAPTER LIGATION-SELECTIVE AMPLIFICATION INVENTION.

Definitions

As used herein, the terms "restriction endonucleases" and "restriction enzymes" refer to bacterial enzymes each of which cut double-stranded DNA at or near a specific nucleotide sequence.

As used herein, the terms "oligonucleotide" and "oligomer" are defined as molecules comprised of two or more deoxyribonucleotides or ribonucleotides, preferably more than three. One or more of the nucleotides can be modified e.g. by addition of a methyl group, a biotin or digoxigenin moiety, a fluorescent tag or by using radioactive nucleotides.

As used herein, the term "primer" refers to an oligonucleotide, whether occurring naturally as in a purified restriction digest or produced synthetically, which is capable of acting as a point of initiation of nucleic acid sequence synthesis when placed under conditions in which synthesis of a primer extension product which is complementary to a nucleic acid strand is induced, i.e. in the presence of four different nucleotide triphosphates and a polymerase in an appropriate buffer ("buffer" includes pH, ionic strength, cofactors etc.) and at a suitable temperature. One or more of the nucleotides of the primer can be modified e.g. by addition of a methyl group, a biotin or digoxigenin moiety, a fluorescent tag or by using radioactive nucleotides.

As used herein, the term "DNA polymorphism" refers to the condition in which two or more different nucleotide sequences can exist at a particular site in the DNA.

As used herein, the term "nucleotide variation in sequence" refers to any type of nucleotide variation being caused either by single or multiple nucleotide substitutions, or by

deletions or insertions. These nucleotide variations may be mutant or polymorphic allele variations.
As used herein the term "restriction fragment length polymorphism"("RFLP") refers to the differences among individuals in the lengths of restriction fragments formed by digestion with a particular restriction endonuclease.
As used herein the term "PCR" refers to the polymerase chain reaction (Mulis et al U.S.Pat.Nos. 4,683,195, 4,683,202 and 4,800,159). The PCR amplification process results in the exponential increase of discrete DNA fragments whose length is defined by the 5' ends of the oligonucleotide primers.

Conventional techniques of molecular biology and recombinant DNA techniques, which are in the skill of the art, are explained fully in the literature. See, e.g., Sambrook, Fritsch and Maniatis, Molecular Cloning; A Laboratory Manual, Second Edition (1989) and a series, Methods in Enzymology (Academic Press, Inc.).

A method for the selective amplification of a small number of DNA fragments from a complex mixture of fragments has been published (Vos et al, 1985 ; European patent application no. 534 858 A1). This method, known as amplified fragment length polymorphism (AFLP), is based on the digestion of a DNA preparation with two restriction-endonucleases, followed by the ligation of DNA oligomers to the ends of these fragments. A subset of the resulting fragments are amplified by the Polymerase Chain Reaction (PCR), separated by gel-electrophoresis, and visualised by a suitable detection method, e.g. autoradiography.
Using two restriction enzymes generating different single stranded tails, and two different adapters, A and B, a DNA mixture is obtained with three different kinds of fragments :
1) Fragments having two ends generated by the same restriction-enzyme and having type A adapters attached to

both ends; 2) Fragments having two B-type adapters and 3) Fragments having one A and one B type adapter.
In the AFLP technique, the number of fragments amplified is reduced in three ways:
First, the two primers complementary to the adapters differ in their annealing temperature. The primer with the higher annealing temperature is preferentially used during a PCR reaction favouring the amplification of fragments with at least one adapter complementary to this primer. The restriction enzymes for the digestion of the DNA are chosen in a way that a vast majority of the fragments contain adapters that are both complementary to the primer with the lowest annealing temperature and will thus be amplified less efficiently.
Second, during the PCR process, the fragments having two different adapters attached are preferentially amplified, as the ends of fragments having two identical adapters are able to anneal to each other as well as to the PCR primers.
The third way to limit the number of fragments amplified is the inclusion of selective bases in the primers. Fragments need to have an adapter attached and thus have ends consisting of an adapter sequence and a restriction endonuclease recognition site, but also need to have a specific sequence of 1 - 3 bases next to the restriction endonuclease recognition site in order to be amplified.
The AFLP method as described is to some extent sensitive to changes in the protocol, such as a change in restriction enzymes used, and impurities in the DNA preparations.
These drawbacks of the AFLP method are due to the fact that all DNA-fragments contain adapter sequences at both ends. Each DNA fragment thus can act as a template during the polymerase chain reaction and all fragments produced during a PCR cycle will be a template during the next PCR cycle, as all fragments produced have an adapter sequence complementary to one of the PCR primers. Due to competition during the PCR

cycles, the fragments having two different adapters attached are enriched using the described protocol.

In the AFLP technique, two different restriction-endonucleases are used for digestion of the genomic DNA. One frequent-cutter, such as Mse I and a rare-cutter, such as EcoR I. Assuming a completely random DNA sequence, the EcoR1 restriction endonuclease, recognising the DNA-sequence GAATTC, will cut the DNA on average once every 4096 bp. The Mse I restriction endonuclease, recognising the DNA-sequence TTAA, will cut the DNA on average once every 256 bp. Both enzymes leave different protruding ends on each fragment produced e.g. the rare-cutter EcoR1 will produce fragments with the sequence TTAA on each 5' end of the digested DNA. The frequent cutter Mse I will leave a 2 bp extension AT on each 5' end.
To these different ends, different DNA oligonucleotides (adapters) can be ligated using the enzyme T4-DNA-ligase. As a result, the DNA will be cut on average 17 times in every 4096 bp. The resulting sample will contain a large number of small fragments, most containing the Mse-adapter sequence at both ends (approx. 88 %). A small number of fragments will contain one Mse 1 and one EcoR1 adapter (approx. 10 %), and a very small number will contain two EcoR1 adapters (approx. 1,4 %). In the PCR part of the AFLP technique, all fragments may be amplified as every fragment has two adapters attached which serve as recognition point for the PCR primers. All single stranded DNA fragments produced by the polymerase reaction and the subsequent heating step can serve as template for the next round of amplification.
During the many cycles of amplification, the fragments containing one EcoR1 and one Mse 1 adapter will be amplified by the Polymerase Chain Reaction more efficiently than the other fragments. The amplification step of the AFLP technique should be considered as a competition in which only the carefully selected conditions will give successful results.

By using PCR-primers complementary to the adapter sequence and the restriction endonuclease recognition site, the number of fragments amplified is usually too large for analysis. The number of different fragments amplified is therefore limited by the addition of up to three selective bases to each of the primers. Theoretically the number of amplified fragments is reduced by a factor 4 for every selective base added to one of the primers.

In a preferred embodiment of the invention a process is provided for the analysis of nucleic acids comprising providing said nucleic acids in a double stranded DNA form wherein said process further comprises fragmenting said DNA to produce a collection of DNA fragments with at least three different types of ends. Upon the fragmentation of the DNA, at least two types of oligomer adapters are ligated to at least two different types of ends. At least a first oligomer primer complementary to the first adapter and a second oligomer complementary to the second adapter are provided and said DNA fragments are amplified, whereby at least one type of DNA end does not take part in the amplification reaction. Following the amplification reaction the amplified products are analysed with a suitable method, preferable by, but not limited to gel-electrophoresis. With DNA ends not taking part in the amplification reaction is meant those DNA ends from which no amplification is initiated in the reaction. Since a DNA fragment of the invention has two DNA ends, amplification of the DNA end not taking part in the amplification reaction may be initiated from the opposing DNA end. However, since in this case the amplification is not exponential, said fragment will not be a predominant amplification product and not contribute significantly to the DNA-fingerprint. A suitable but non-limiting method to prevent a DNA end from taking part in the amplification reaction is not to supply a primer complementary to said DNA end in the amplification reaction.

Fragmentation of the DNA is preferably accomplished by means of one or more restriction enzyme digestions. However, other DNA fragmenting enzymes may also be used for the present invention as long as they produce specific fragments and do not cut apparently at random such as a DNase.

In a preferred embodiment of the invention the DNA ends taking part in the amplification reaction are fewer in number than the DNA ends not taking part in the amplification reaction. In a particularly preferred embodiment of this invention, called the three restriction enzyme embodiment, the fragmenting of DNA is achieved with three restriction enzymes, two of which preferably recognise a 6 bp sequence and the third preferably recognises a 4 bp sequence, wherein preferably the enzyme recognising a 4 bp sequence is used to produce the DNA fragment end that is not taking part in the amplification reaction. In a preferred embodiment of the invention the fragmenting of the DNA occurs simultaneously to the ligation of the adapters.

In a preferred embodiment of the invention one primer has an essentially lower annealing temperature than the another primer. In another preferred embodiment of the invention at least one primer is labelled, preferably the primer with the lowest annealing temperature. Labelling may occur through any suitable method such as through biotin, digoxigenin, fluorescent tags or radioactivity. In a preferred embodiment of the invention the labelled primer is labelled at its 5'end. In yet another preferred embodiment of the invention at least one primer further comprises one or more selective nucleotides, leading to the preferred amplification of a subset of all DNA fragments containing the adapter to which the primer was complementary.

In another aspect of the invention, mRNA in a sample is converted into double stranded DNA, using methods known in the art such as reverse transcription, prior to performing the method of the invention. The resulting DNA fingerprint is, in this case, in effect a fingerprint of the mRNA in a

sample. In another aspect of the invention is provided a kit useful for performing the method of the invention wherein the kit minimally comprises at least one method for fragmenting DNA to produce a collection of DNA fragments with at least three different types of ends.

In a preferred embodiment of the invention the amplification of the fragments is achieved by means of polymerase chain reaction. The invention is further illustrated using this amplification method. However, the invention is not limited to amplification through polymerase chain reaction. The NASBA amplification reaction or other methods may also be used.

In one embodiment of the invention only a fraction of the DNA fragments generated in the DNA fragmenting step are further processed in the method of the invention, i.e. a fraction of the fragments is purified from the pool of fragments generated. Said purification may occur through any means applicable. For instance, when part of the sequence of interest is known, specific purification oligomers can be generated that comprise of a sequence complementary to the sequence of interest and further comprise a means of purification, such as a biotin tag or attachment to an immobile substrate. In a preferred embodiment of the invention one of the primers added is not complementary to the known sequence of the adapters but instead complementary to a different known sequence, resulting in the specific amplification of only those fragments containing the known sequence. This specific embodiment may be useful in, for instance but not limited to, the characterisation of retroviral or transposon integration sites. In a preferred embodiment the primer not complementary to the adapter sequences is complementary to a retroviral or transposon sequence. Another preferred embodiment where one of the primers added is not complementary to the known sequence of the adapters but instead complementary to a different known sequence, resulting in the specific amplification of only those fragments containing the known sequence is termed AFLP

sequence tagging. The process termed AFLP® sequence tagging, amplifies a single AFLP marker from an AFLP template or pre-amplification reactions, with the aim to develop a dominant PCR test for this AFLP marker. This method is based on the use of an AFLP primer, in combination with an oligonucleotide primer that is derived from the internal sequence of the AFLP marker. Compared to a "standard" AFLP AFLP sequence tagging can be used to obtain an essentially specific signal (one or a few fragments) for the relevant polymorphism(s), essentially without the large number of fragments resulting from "standard" AFLP. Thereby enabling an easier detection of said polymorphism. The necessity of having at least one primer present comprising a specificity for an adapter is probably best exemplified in the following non-limiting example. When an AFLP template is generated, information for the specific polymorphism lies in the adapter ligated to the exact cutting site plus eventual selective bases in the primer directed to said adapter. To amplify this information one may use a further primer directed toward an adapter ligated to the other end of the fragment. However, this usually results in a rather complicated fragment pattern due to the relatively large number of fragments amplified. However, in one embodiment of the AFLP sequence tagging technique of the invention, instead of the primer for an adapter at the other end of the fragment a specific primer is used, directed against a known sequence in the fragment, thereby eliminating a large number of essentially irrelevant fragments. In this one embodiment the primer directed toward the adapter comprising the polymorphic information cannot be replaced by a specific primer. One may envisage cloning of DNA adjacent to the site comprising the polymorphic information such that an amplification with two specific primers may be devised. However, apart from the fact that this not always possible, cloning of adjacent DNA is cumbersome and time consuming.

The AFLP sequence tagging method of the invention is now further exemplified below. The method is based on the use of an AFLP primer, in combination with a oligonucleotide primer that is derived from the internal sequence of the AFLP marker. Since AFLP templates are prepared using two restriction enzyme combinations (followed by adapter ligation), there are two possible ways to amplify a particular AFLP fragment from an AFLP template when using a combination of an AFLP primer and a primer derived from the internal sequence of an AFLP marker. This is schematically shown in Figure 7: in this figure one possibility is to use internal primer P1 in combination with the MseI + 3 AFLP primer and the other possibility is to use the EcoRI +2 primer in combination with the internal P2 primer.

Since the presence of AFLP *marker* in an AFLP fingerprint is the result of a polymorphism in at least one restriction enzyme site and/or the adjacent selective nucleotides, there are different expected outcomes of the use of the two primer combinations EcoRI+2- internal P2 or MseI +3 - internal P1, depending on the molecular polymorphism that is responsible for the particular AFLP marker.

For individuals where the AFLP marker is present in the +2/+3 fingerprint, generally, both primer combinations (EcoRI+2- internal P2 or MseI +3 - internal P1) will result in a PCR product, when taking AFLP templates or preamplification reactions as starting material.

However, for individuals where the AFLP markers is absent in the +2/+3 fingerprint (due to polymorphism in the restriction enzyme sites and/or selective bases) , PCR product formation with the two primer combinations (EcoRI+2- P2 or MseI +3 - P1) depends on the position of the AFLP marker polymorphism in the EcoRI restriction enzyme site, MseI restriction enzyme

site, a selective base on the MseI site and/or a selective base on the EcoRI site. In at least one of these positions a polymorphism is present resulting in the absence of the AFLP markers in the fingerprint of the individual. For the amplification with the two primer combinations EcoRI+2- P2 or MseI +3 -P1 this means that:

1) when the AFLP marker polymorphism is located in the EcoRI restriction enzyme **site** PCR product is expected with primer combination MseI+3 - internal P1, starting from AFLP templates. However, no PCR product is expected with primer combination EcoRI+2- P2. No PCR product is expected with primer combination MseI+3 - internal P1 when using +0/+0 or +0/+3 pre-amplifications as starting material for the PCR because the EcoRI site is not present and the next available EcoRI site is usually not located within amplifiable distance.
2) when there is only polymorphism in the MseI restriction enzyme **site**, a PCR product is expected with primer combination EcoRI+2 - internal P2, when taking AFLP templates as starting material. PCR product formation starting from +0/+0,+1/+0 or +2/+0 AFLP preamplification reactions is dependent on the position of the nearest MseI site in the genome. For relatively small AFLP markers there is a fair chance that the nearest MseI site is located within amplifiable distance in which case PCR product formation occurs with primer combination EcoRI+2 - internal P2. In this case the PCR product will have a different length.
3) when there is only polymorphism in one of the **selective nucleotides** on the EcoRI side, PCR product is expected with primer combination MseI+3 - internal P1, starting from AFLP templates, +0/+0 and +0/+3 preamplification reactions but not when starting from +2/+0 preamplification reactions. However, PCR product formation with primer combination EcoRI+0 - internal P2 is

dependent upon the position of the polymorphism at the +1 or +2 selective EcoRI base, but PCR product will be formed with when AFLP templates,+0/+0 and +0/+3 preamplifications are used as the starting material.

4) when there is only polymorphism located at one of the **selective nucleotides** on the MseI side, PCR product is expected with primer combination EcoRI - internal P2 when using AFLP templates or +0/+0 and +2/+0 preamplification reactions as starting material but not when using +0/+3 preamplifications. However, PCR product formation with primer combination MseI+0 - internal P1 is dependent on the position of the polymorphism at the +1,+2 or +3 selective MseI base, but PCR product is expected when using AFLP templates and +0/+0 preamplification reactions and possibly also +2/+1 or +2/+2 (pre)amplification reactions as starting material.

Note:
In all the above situations PCR product refers to specific PCR product that is derived from the AFLP marker that is targeted with the internal P1 or internal P2 primer.

Based on the above it can be concluded that the amplification of a part of the sequence of an AFLP marker with one primer combination consisting of an AFLP primer and a primer derived from the internal sequence of the AFLP marker of interest can result in a dominant PCR assay where PCR product formation is correlates with presence of absence of the AFLP marker in the fingerprint. We have named this method AFLP-sequence tagging.

Non-limited applications of AFLP sequence tagging include the development of a dominant PCR assay for markers of interest (such as AFLP markers with predictive value for (a) trait(s) of interest, amplification of AFLP fragments derived from gene-families members, targeting of AFLP markers containing

specific conserved domain sequences, such as those present in certain types of plant disease resistance genes, and targeting of retrotransposon containing restriction fragments.

In the present invention, the number of amplified fragments is limited by digesting the DNA with two restriction enzymes generating different ends, and ligating different adapters to these ends, as well as digesting the DNA with one or more other fragmenting enzymes for which no adapters are added. Preferably the fragmenting enzymes, are restriction enzymes. In a preferred embodiment of the invention the ends for which no specific adapters are added are fragmented by enzymes that cut the DNA (much) more often than the two restriction-enzymes that are used to produce the ends for which adapters are added. As a result 4 different kinds of fragments are produced: 1) Fragments without any adapter sequence (the majority of fragments in the preferred embodiment). These will not be amplified at all. 2) Fragments having one adapter sequence. These may be amplified but not in an exponential way, as the fragments produced will have no sequence complementary to one of the PCR primers. 3) Fragments having two identical adapters and 4) Fragments having two different adapters. These latter two kinds of fragments will be exponentially amplified as with the usual AFLP protocol. In a preferred embodiment these two types of fragments are present in roughly equimolar amounts.
During the amplification reaction the fragments with two different adapters are preferentially amplified. Furthermore the annealing temperature of the primers is chosen such that the primer with the lower annealing temperature is labelled (In contrast to the AFLP method in which the labelled primer has the highest annealing temperature.). Fragments containing two unlabeled primers may be amplified but are not detected. Amplification of fragments with two labelled primers is not efficient as the annealing temperature of these primers is

low. Amplification of fragments with one labelled and one unlabeled primer is strongly favoured.
In a preferred embodiment of the three restriction enzyme embodiment, the mixture of DNA fragments is obtained by digesting the DNA with a mixture of 3 restriction-endonucleases. Two of these enzymes recognise a 6 bp sequence and the third enzyme recognises a 4 bp sequence. In this preferred embodiment of the three restriction enzyme embodiment the adapters are ligated to all ends made by digestion with the two 6 bp cutters. The digestion and the ligation reaction may be performed simultaneously. Part of the digestion-ligation reaction is used as a template in an amplification reaction such as PCR. The amplification primers used are complementary to the adapters used in the digestion-ligation step but may contain selective nucleotides in order to reduce the number of fragments amplified.
The results can be visualised e.g. by gel-electrophoresis. The resulting gel-patterns ("fingerprints") can be used e.g. to distinguish closely related organisms such as different strains of one organism.

THEORETICAL BACKGROUND OF THE DNA-PARTIAL ADAPTER LIGATION-SELECTIVE AMPLIFICATION METHOD, ILLUSTRATED BY, BUT NOT LIMITED TO, EMBODIMENTS OF THE THREE RESTRICTION ENZYME EMBODIMENT.

Suppose a DNA genome is cut by three different Restriction-endonucleases: A, B and C, having a, b and c cutting sites. For a circular DNA, the number of fragments obtained is a + b + c. The number of DNA-ends is thus 2(a+b+c). Using three enzymes, 6 different kinds of fragments are generated.
The fraction of fragments with 2 ends both generated by the A-enzyme is $a^2/(a+b+c)2$
The fraction of fragments with 2 ends both generated by the B-enzyme is b2/(a+b+c)2

The fraction of fragments with 2 ends both generated by the C-enzyme is c2/(a+b+c)2
The fraction of fragments having one end generated by the A-enzyme and one end generated by the B-enzyme is 2ab/(a+b+c)2.
The fraction of fragments having one end generated by the A-enzyme and one end generated by the C-enzyme is 2ac/(a+b+c)2.
The fraction of fragments having one end generated by the B-enzyme and one end generated by the C-enzyme is 2bc/(a+b+c)2.

In case adapters are ligated to the DNA-ends generated by enzymes B and C, and that during the amplification reaction the fragments having an adapter on both ends are amplified, the fraction of fragments that can be amplified is b2+c 2+2bc/(a+b+c)2. The total number of fragments that can be amplified is b2+c 2+2bc/(a+b+c).
Assuming that only the fragments having two different adapters are amplified efficiently, the fraction of the fragments that are amplified efficiently is 2bc/(a+b+c)2. The total number of fragments that can be efficiently amplified is 2bc/(a+b+c)

Typical situations are :
Traditional AFLP : No A enzyme is present. The B-enzyme has approximately 16 times more cutting sites than the C-enzyme. All fragments can be amplified. The fraction of fragments that are amplified efficiently is 2 x 16c x c / (17c)2 = 32 / 289 = 11,1 %.

DNA-partial adapter ligation-selective amplification using two enzymes with a 6 bp recognition site and one enzyme with a 4 bp recognition site : The B- and C-enzymes (both with an 6 bp recognition site) have approximately the same number of cutting sites. The A-enzyme (having a 4 bp recognition site), for which no adapters are added, has approx. 16 times as much cutting sites. The fraction of fragments that can be amplified is 4b2/(18b)2 = 4/324 = 1,23 %. The number of

fragments that can be efficiently amplified is 2b2/(18b)2 = 2/324 = 0,62 %.
Instead of two enzymes with a 6 bp recognition site, an enzyme such as Ban 2 can be used. Ban 2 has several different 6 bp recognition sites to which different adapters can be ligated.

DNA-partial adapter ligation-selective amplification using two enzymes with a 7 bp recognition site and one enzyme with a 4 bp recognition site : The B- and C-enzymes (both with an 7 bp recognition site) have approximately the same number of cutting sites. The A-enzyme (having a 4 bp recognition site), for which no adapters are added, has approx. 64 times as much cutting sites. The fraction of fragments that can be amplified is 4b2/(66b)2 = 4/4356 = 0.1 %. The number of fragments that can be efficiently amplified is 2b2/(66b)2 = 2/4356 = 0,05 %.
Instead of two enzymes with a 7 bp recognition site, an enzyme such as Rsr 2 can be used, having two different 7 bp recognition sites to which two different adapters can be ligated.
Using the DNA-partial adapter ligation-selective amplification technique, the number of fragments that are potentially amplified can thus be easily reduced 80 - 1000 fold compared to the AFLP method (100 % vs. 1,23 % or 0.1 %). The number of fragments that are efficiently amplified is easily reduced 20 - 220 fold (11 % vs. 1,23 % or 0.05 %).
In one embodiment of the three restriction enzyme embodiment of the invention, using two restriction enzymes recognising 6 bases for producing the ends taking part in the amplification and using a restriction enzyme recognising 4 bases for producing the ends not taking part in the amplification reaction, only approx. 1 % of the fragments can be amplified exponentially, compared to 100 % of the fragments in the AFLP method. In the DNA-partial adapter ligation-selective amplification method the approx. 0.6 % of the fragments that

have two different adapters has to compete only with the 0.6 % of the fragments that have two identical adapters. With the present invention it is possible to use longer primers and higher annealing temperatures during the PCR cycles, resulting in more stringent amplifications and lower background.

Assume the genome is a typical bacterial genome of 4.000.000 bp. The A enzyme for which no adapters are added cuts 15.625 times (on average every 256 bp). The B and C enzymes for which adapters are added each cut 976 times (every 4096 bp.). The total number of fragments that can be amplified will be 216. OF these, 108 will be amplified efficiently as they have two different adapters attached to their ends. For bacterial genomes, it will therefore not be necessary to further reduce the number of fragments by using selective bases on the amplification primers in case the fragments are separated on a long acrylamide gel. Addition of one selective nucleotide to each of the PCR primers will further reduce the average number of fragments amplified to 7, permitting the use of agarose gels with ethidium bromide for separation and detection.

In case of a human genome of 3.000.000.000 bp, the number of efficiently amplified fragments will be 750 times more : 81.000. By using three selective nucleotides on each primer, this number is reduced to only 20.

In contrast the AFLP method will generate on average almost 2000 fragments on a bacterial 4.000.000 bp genome (without selective nucleotides) and on average 400 fragments on a human genome when using three selective nucleotides on each PCR primer.

Due to the non-random arrangement of base pairs in a genome, certain restriction sequences may be substantially underrepresented. It is therefore often possible to select for a certain organism restriction endonucleases which recognise 6 bp and which cleave on average less than once

every 4096 bp. In the AFLP method the number of fragments is reduced 10 fold when the enzyme used cleaves only once every 40960 bp. In the three restriction enzyme embodiment of the invention, using two enzymes that cleave on average only once every 40960 bp. the number of fragments amplified is reduced 100-fold.

EXAMPLE 1.

We noticed that the DNA-PARTIAL ADAPTER LIGATION-SELECTIVE AMPLIFICATION method is much more robust than the AFLP technique described by Zabeau and Vos. When using the AFLP technique on complex DNA (e.g. human or plant DNA) it is necessary to perform a "pre amplification" PCR reaction. We obtained excellent DNA-PARTIAL ADAPTER LIGATION-SELECTIVE AMPLIFICATION fingerprints on these DNA's using a combined digestion-ligation reaction of 1 - 2 hr, followed by a single PCR reaction. As a result a typical DNA-PARTIAL ADAPTER LIGATION-SELECTIVE AMPLIFICATION fingerprint is obtained in less time, and at reduced costs compared to a typical AFLP fingerprint.

In order to show the robustness of the DNA-PARTIAL ADAPTER LIGATION-SELECTIVE AMPLIFICATION method, the effects of several possible interfering components and variables were evaluated. This evaluation was done using an implementation of the invention based on the three restriction enzyme embodiment (see also example 5).

Digestion-ligation reaction.

No change in pattern was observed when DNA samples of between 1 and 1000 ng tomato DNA / 20 ul reaction were used. The pattern changed when only 0.2 ng DNA / 20 ul was used. As only 0.5 ul of the digestion/ligation reaction was used for the PCR reaction, this corresponds to 5 pg tomato DNA / PCR reaction, or only approx. 5 haploid genome equivalents.

Using 40 ng tomato DNA / 20 ul digestion/ligation reaction, no change in pattern was observed when reaction times were reduced to 30 minutes or extended to 4 hours, or if the incubation temperature was raised to 37 °C or lowered to 25 °C (at an 1.5 hr incubation time).

Using 40 ng tomato DNA / 20 ul reaction, 1.5 hr incubation time at 30 °C. and 0.5 ul of the digestion/ligation reaction as template for the PCR reaction, no change in pattern was observed if the concentration of any reaction component used for the digestion/ligation reaction was increased two-fold or reduced two-fold, including : All enzymes used, DNA-oligomers, ATP, Tris-HCl, NaCl, MgCl2 and DTT.

PCR reaction.

No change in pattern was observed if the annealing temperatures were increased or decreased 3 °C, if the number of cycles was increased with 10 cycles, if the Mg concentration was varied between 1.3 and 1.6 mM, the KCl concentration was increased with 20 mM, the reaction components were mixed at room-temperature or on ice either immediately before, or 1 hour before the start of the PCR reaction.

The DNA-PARTIAL ADAPTER LIGATION-SELECTIVE AMPLIFICATION method proved to be sensitive however to the following changes :

The use of excess Dde 1 : More than 0.25 unit in the DNA sample added to the PCR reaction disturbed the pattern (5 - 10 units used / 20 ul digestion-ligation reaction).

Excess Mg : If the Mg concentration was raised above 1.75 mM (including the Mg present in the DNA sample, typically containing 6 mM MgCl2) the relative amounts of the PCR products changed.

Excess adapters : If more than 0.1 pMol of each of the ligation-adapter was present in a 50 ul PCR reaction, the pattern obtained was disturbed in case adapters were used

with an annealing temperature higher than those of the PCR primers.

Although specific restriction endonucleases are recited in the Examples, it will be recognised that isoschizomers, i.e. enzymes that have the same recognition sequence, or other restriction enzymes can be substituted and the same or equivalent results will be achieved. The following examples illustrate but do not limit the invention.

Enzyme-mixtures used contain 1 unit/ul of each restriction-endonuclease + 0.25 Weiss units T4-DNA-ligase. In these experiments, we define 1 unit restriction-endonuclease as the amount of enzyme needed to digest to completion in 1 hr at 30 °C, 1 ug of a hypothetical DNA with 1 site in every 4096 bp. for each enzyme with a 6 bp recognition site and 1 site in every 256 bp. for each enzyme with a 4 bp. recognition site. The buffer used contains 10 mM Tris-HCl pH 7,6 ; 50 mM NaCl ; 6 mM MgCl2 and 0.5 mM ATP. Activity of enzymes was measured with unmethylated Lambda DNA as substrate and was afterwards corrected for site density in Lambda DNA.
Mbo I for instance has 116 cleavage sites in the 48.502 bp Lambda genome, or 1 site in every 418 bp. If 0.1 ul of an Mbo I (4 bp recognition site) preparation was sufficient to digest 1 ug unmethylated Lambda DNA, the activity was stated to be 10 u/ul x 256/418 = 6.1 units/ul.
The term "unit" for the activity of T4-DNA-ligase is defined as the amount of enzyme able to convert 1 nmol (32P) from pyrophosphate into Norit-absorbable material in 20 min at 37 °C (Weiss units).
Two of the restriction-enzymes used are not yet commercially available. They can be replaced with slightly less satisfactory results by well known isoschizomers. Fun 2 is an isoschizomer of EcoR1, but has no star activity and has a higher activity at temperatures lower than 37 °C. SpaH1 is an isoschizomer of Sph 1.

EXAMPLE 2.

This example pertains to the generation of a DNA fingerprint of DNA isolated from phages Lambda, T7 and Adenovirus. As the DNA sequence of these viruses is exactly known, the size of the fragments amplified with the DNA-PARTIAL ADAPTER LIGATION-SELECTIVE AMPLIFICATION method can be predicted. We used Genbank accession numbers J02459 (Lambda), V01146 (T7) and J01917 (Adenovirus).

In a typical reaction we used 10 ng of T7 DNA, 10 pg adeno DNA and 10 pg Lambda DNA. This mix was added to a digestion/ligation reaction in a final volume of 20 ul containing 10 mM Tris-HCl pH 7,6 ; 50 mM NaCl ; 10 mM MgCl2 ; 1 mM DTT ; 2 pMol of each of four oligomers (SEQ ID NO 3, SEQ ID NO 4, SEQ ID NO 5 and SEQ ID NO 6), 0.5 unit T4-DNA ligase ; 0.5 unit Rsa 1 ; 0.5 unit Fun 2 ; 0.5 unit SpaH1. Incubation was performed for 1.5 hr at 30 °C. Following the digestion/ligation step, 0.5 ul of this mixture was used as a template for PCR amplification in 500 ul reaction tubes using a Perkin Elmer thermal cycler and using the following conditions :

a) 2.5 minute denaturation at 95 °C.

b) 30 cycles consisting of 30 second denaturation at 95 °C ; 90 second annealing + elongation at 72 °C.

The 50 ul PCR reaction contained 1 ul of the digestion/ligation reaction as template, as well as 10 pMol of an 5'-FITC labelled primer complementary to the Fun 2 adapter and having one selective nucleotide on its 3'end (T ; SEQ ID NO 7), 10 pMol of an unlabeled primer complementary to the SpaH1 adapter (no selective nucleotides, SEQ ID NO 8), 15 mM Tris-HCl pH 8.5 , 50 mM KCl , 1.5 mM MgCl2 , 0.01 % non-ionic detergent such as Triton X-100 , 200 uM of each of the four dNTP's and 0.5 unit Taq polymerase. The PCR mixture was overlayered with mineral oil.

Following the PCR reaction, 8 ul of this reaction was analysed on a 1.8 % agarose gel in 0.5 x TBE buffer. The results obtained were strong bands on a agarose gel with lengths of approx. 1500, 480 and 270 bp. This is in excellent agreement with the bands expected from the known nucleotide sequences of these phages : 1509 bp (including primer sequences) from the fragment between the SpaH1 sites at position 3652 and 5127 in Adenovirus DNA ; 496 bp from the Adenovirus SpaH1 site at 31.802 till the Fun 2 site at 32.264 (Preceded by a T residue in the complementary strain) and a 284 bp fragment from the Lambda DNA Fun 2 site at 39168 till the SpaH1 site at 39418.
No fragments were produced from the T7 DNA template which was present in a 1.000 times larger amount.
Results obtained with many other enzyme combinations also agreed with the fragments expected from the DNA-sequence of the phages. Fragments with two identical adapters attached (like the 1475 bp fragment in the experiment described) were only present when longer than approx. 600 bp.

EXAMPLE 3.
This example pertains to the generation of DNA fingerprints of genomic DNA isolated from an E. coli strain, lysogenic for phage Lambda.
Genomic DNA was isolated from E. coli strain DM1 (Life Technologies). Samples containing 20 ng samples DNA were processed as in DNA-PARTIAL ADAPTER LIGATION-SELECTIVE AMPLIFICATION example 2. The following restriction-enzymes, adapters and PCR primers were used :
Lane 1: Fun 2, SpaH1 and Dde I. Fun 2 is an isoschizomer of EcoR1. SpaH1 is an isoschizomer of Sph 1. Adapters : Eco (SEQ ID NO 3 and SEQ ID NO 4) + Sph (SEQ ID NO 5 and SEQ ID NO 6). Primers : FITC-labelled Eco-T (SEQ ID NO: 7) and unlabeled Sph-0 (SEQ ID NO: 8)
Lane 2: As in lane 1, but Dde I was replaced by Mbo I.

Lane 3: Fun 2, Bcl I and Rsa I. Adapters : Eco (SEQ ID NO 3 and SEQ ID NO 4) + Bam (SEQ ID NO 9 and SEQ ID NO 10). Primers : FITC-labelled Eco-T (SEQ ID NO: 7) and unlabeled Bcl-0 (SEQ ID NO: 11).

Lane 4: As in lane 3, but Rsa I was replaced by Dde I.

The results are shown in FIG. 2 As expected each enzyme combination gives a different DNA-fingerprint pattern. Some bands smaller in size that are present in lane 1 comigrate and may be identical to bands in lane 2. These fragments are flanked by Fun 2 and SpaH1 sites and do not contain Dde I and Mbo I sites. Some bands smaller in size that are present in lane 3 comigrate and may be identical to bands in lane 4. These fragments are flanked by Fun 2 and Bcl 1 sites and do not contain Dde I and Rsa I sites. The strong signal detected in lanes 1 and 2 at 225 minutes comigrates with a 284 bp fragment detected when only phage Lambda was used as a test DNA. The coli strain used is lysogenic for Lambda and is dam-, permitting the use of the Mbo I and Bcl 1 enzymes. Indeed the Lambda fragment between nucleotides 39168 and 39423 of the Lambda sequence (Genbank accession NO J02459) is the only Lambda fragment flanked by EcoR1 and Sph 1 sites that does not contain a Dde 1 or Mbo I site and that contains a T nucleotide next to the EcoR1 site. After adapter ligation and PCR, the 254 bp fragment will generate a FITC labelled 284 nucleotide fragment.

EXAMPLE 4.

This example pertains to the generation of a DNA fingerprint from different specimens of a small insect (Orchesella cincta ; springtails). DNA from different individuals was purified by phenol extraction, chloroform extraction and ethanol precipitation. Approx. 250 ng DNA could be obtained from one individual.

10 ng of this DNA was added to a digestion/ligation reaction with a final volume of 20 ul and containing 10 mM Tris-HCl pH 7,6 ; 50 mM NaCl ; 6 mM MgCl2 ; 1 mM DTT ; 1 pMol of each of

four oligomers (SEQ ID NO 12, SEQ ID NO 13, SEQ ID NO 9 and SEQ ID NO 10), and 0.5 units each of T4-DNA ligase, Rsa 1, Xba 1 and BamH1.
Incubation was performed for 1.5 hr at 30 °C. Following the digestion/ligation step, 0.5 ul of this mixture was used as a template for PCR amplification in 500 ul reaction tubes using a Perkin Elmer thermal cycler and using the following conditions :
a) 2.5 minute denaturation at 95 °C.
b) 10 cycles consisting of 30 second denaturation at 95 °C ; 30 second annealing at 70 °C and 60 second elongation at 72 °C.
c) 40 cycles consisting of 30 second denaturation at 95 °C ; 30 second annealing at 60 °C and 60 second elongation at 72 °C.

The 50 ul reaction volume contained 0.5 ul of the digestion/ligation reaction as template as well as 10 pMol of an 5'-FITC labelled primer complementary to the Xba adapter and having three selective nucleotides on its 3'end (CCG ; SEQ ID NO 1), 10 pMol of an unlabeled primer complementary to the BamH1 adapter (no selective nucleotides ; SEQ ID NO 2) , 15 mM Tris-HCl pH 8.5 ; 50 mM KCl ; 1.5 mM MgCl2 ; 0.01 % non-ionic detergent such as Triton X-100 ; 200 uM of each of the four dNTP's and 0.5 unit Taq polymerase. The PCR mixture was overlayered with mineral oil.

Following the PCR reaction, 2 ul of this reaction was mixed with 2 ul of formamide containing 5 mg/ml blue dextran, heated for 5 minutes at 80 °C in order to denature the DNA and was analysed on a 6 % acrylamide gel (acrylamide-bisacrylamide 29 : 1), containing 8 M urea in 100 mM Tris-borate pH 8.3 ; 2 mM EDTA. A Pharmacia ALF apparatus was used to run the gel and detect the fluorescent PCR products. The results obtained are shown in Fig. 3 . All Cincta individuals gave distinctive fingerprints. Cincta individuals collected

in one forest had several bands in common (e.g. lanes 2-4), while individuals collected from other forests gave very distinct patterns (e.g. lane 1).

EXAMPLE 5.

This example pertains to the generation of DNA fingerprints from the genomic DNA of different Lycopersicum (tomato) strains. Genomic DNA was isolated from L. esculentum (lane 1) and L. penelli (lane 2) and from 4 different individuals of a cross between these two species (lanes 3-6).

Samples containing 20 ng samples DNA were processed as in example 4. The following restriction-enzymes, adapters and PCR primers were used :

Restriction enzymes: Xba I, BamH1 and Rsa I. Adapters : Xba (SEQ ID NO 12 and SEQ ID NO 13) + Bam (SEQ ID NO 9 and SEQ ID NO 10). Primers : FITC-labelled Xba-CCG (SEQ ID NO: 1) and unlabeled Bam-0 (SEQ ID NO: 2).

The results are shown in FIG. 4.

All fingerprints were different. As expected, all bands observed in lanes 3-6 were also present in the fingerprints of one of the two parents (lanes 1 & 2).

EXAMPLE 6.

This example pertains to the generation of DNA fingerprints from human genomic DNA (Promega). A sample containing 20 ng DNA was processed as in example 4. The same restriction-enzymes, adapters and PCR primers were used. The results are displayed in Fig. 5 and 6. In figure 5, the fragments having lengths between 65 and 710 nucleotides are displayed. Figure 6 shows a part of this same gel pattern.

EXAMPLE 7.

The technique of the invention used for the mapping of transposons. A DNA sample suspected to contain a transposon inserted in a specific gene of which no sequence information is available, is digested with a restriction-enzyme. DNA-

adapters of known sequence are ligated to the DNA ends and the DNA is denatured. Tagged or immobilised oligonucleotides complementary to part of the transposon sequence are allowed to hybridise to the transposon containing fragments. DNA fragments not hybridised to tagged or immobilised oligonucleotides are removed. The resulting DNA preparation which is strongly enriched in DNA fragments containing transposons is used as a template in an amplification reaction such as PCR using one primer complementary to the adapters ligated to the DNA fragments and one primer complementary to the transposon sequence. The resulting amplification products are analysed. In case the organism studied has more than 50 identical transposons in its genome it may be necessary to limit the number of amplified products by addition of one or more selective nucleotides to the primer complementary to the adapter-sequence.
By comparison of different samples that do or do not contain the transposon insert in the gene of interest, the DNA fragment containing both transposon-sequences as well as sequences of the gene of interest can be identified and used for the characterisation of the gene of interest.

EXAMPLE 8.
The technique of the invention used for the detection of polymorphisms. The technique of the invention can be used to detect polymorphisms between the DNA-sequence of a small part of a genome and the homologous part of another genome. Genomic DNA is digested with restriction-endonucleases to fragments of 4 - 40 Kb and the DNA is denatured. Tagged or immobilised oligonucleotides complementary to part of the DNA fragment of interest are allowed to hybridise to the DNA-sample. DNA fragments not hybridised to tagged or immobilised oligonucleotides are removed. The resulting single stranded DNA preparation which is strongly enriched in the DNA fragment of interest is made double-stranded using an oligonucleotide complementary to a sequence at the end of the

DNA molecule , a DNA polymerase and a suitable buffer containing deoxyribonucleotides. The oligonucleotide used as a primer for the polymerase should be unique for the DNA fragment of interest.
The double stranded DNA preparation obtained is digested to small fragments with one or more restriction-endonucleases. Adapters are ligated to the DNA-ends and the resulting preparation of DNA fragments is amplified using oligonucleotides complementary to the adapters used. The amplified fragments can be analysed directly, or after mixing with a control DNA preparation and one cycle of denaturation and renaturation by well known techniques that are not only sensitive to the length of the fragments obtained but also for the sequence of the DNA fragments, such as denaturing gradient gel electrophoresis (DGGE) and SSCP.

EXAMPLE 9.
The technique of the invention used for the detection of polymorphisms between the DNA-sequence of a small part of a genome and the homologous part of another genome.
Genomic DNA is digested with restriction-endonucleases to fragments of 4 - 40 Kb. Adapters are ligated to the DNA ends obtained and the DNA is denatured.
Tagged or immobilised oligonucleotides complementary to part of the DNA fragment of interest are allowed to hybridise to the DNA-sample. DNA fragments not hybridised to tagged or immobilised oligonucleotides are removed. The resulting single stranded DNA preparation which is strongly enriched in the DNA fragment of interest is used as a template in a DNA extension assay. Dependent on the sensitivity of the detection method, the DNA can be used directly in an extension assay or be amplified first e.g. with the use of the Polymerase Chain Reaction.
To the DNA-fragment are added oligonucleotides complementary to specific sequences of the enriched fragment. Each hybridised oligonucleotide acts as a primer that is elongated

by Klenow fragment of E.coli polymerase I in the presence of deoxyribonucleotides and using the right conditions of buffer-composition and temperature. Elongation of each primer halts when the next hybridised primer is reached. After inactivation of the polymerase and denaturation of the DNA, the separate labelled fragments are analysed by techniques such as denaturing gradient gel electrophoresis and the patterns produced are compared with a control sample. Preferably the primers used for the elongation reaction are chosen such that the fragments produced each have a different length and/or behaviour on the gels used for analysis. In a preferred embodiment of the elongation reaction, the primers used for the elongation reaction are labelled. In another preferred embodiment the deoxyribonucleotides used for the elongation reaction are labelled.

The technique of the invention used to analyse a specific mRNA.
An RNA preparation is incubated with a tagged or immobilised oligonucleotide or cDNA fragment that hybridises to the mRNA of interest. Non-hybridised RNA molecules are removed. the resulting RNA preparation which is strongly enriched for the mRNA of interest is used in an elongation assay as described above except that the DNA polymerase is replaced by a reverse transcriptase.

Example 10.
AFLP sequence tagging was used for the specific amplification of a single AFLP marker (PstI/MseI P13/M61-385) in maize. DNA of the parents of a recombinant inbred lines (RIL) population was used as starting material. These parentslines are named Mo17 and B73. AFLP restriction/ligation mixes (R/L mixes) were prepared with restriction enzymes PstI and MseI. PCR reactions were carried out using primercombinations Forward 11 and P13 and using 20-fold diluted +2/+3 AFLP reaction mixture from both parental lines as PCR template.

The following PCR amplification profile was used: 30 cycles of 30 sec at 94°C; 60 seconds at 56°C and 60 seconds 72°C.

The PCR reaction mixture consisted of:

PCR template	: 5.0 µl
Primer P13 (50 ng/ul)	: 1.5 µl
Primer Forward 11 (50 ng/ul)	: 1.5 µl
5 mM dNTP's	: 2.0 µl
10*PCR-buffer (PE)	: 5.0 µl
Taq. Polymerase (PE)	: 0.2 µl
H2O	: 34.8 µl

Primer sequences:

P13:	5'-GAC TGC GTA CAT GCA GAG-`3
M61:	5'-GATGAGTCCTGAGTAACTG-3'
Forward 11	5'-TCG ACC AGA ATC AAT GTT TGG C-`3

It should be understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and the scope of the appended claims. The DNA to be analysed can be genomic DNA from any organism including viruses, micro-organisms, plants, animals and human, but can also be cDNA made by reverse transcription from RNA.

References

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, van de T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., and Zabeau, M. (1995) Nucleic Acid Research 23, 4407 - 4414 .

Patent no.	Inventor	Issued	Title
4683202	Mullis	7/1987	Process for amplifying nucleic acid sequences.

CLAIMS

1. A process for the analysis of nucleic acids comprising providing said nucleic acids in a double stranded DNA form said process further comprising
   - fragmenting said DNA to produce a collection of fragments with at least three different types of ends
   - ligating at least two adapters to the different types of ends
   - providing at least a first oligomer primer complementary to the first adapter and a second oligomer complementary to the second adapter
   - amplifying said DNA fragments, whereby at least one type of DNA end does not take part in the amplification reaction
   - analysis of the amplified products.
2. A process according to claim 1 wherein the fragmenting of DNA is achieved with one or more restriction enzymes.
3. A process according to claims 1 or 2 wherein, at the start of said amplification, the DNA ends taking part in said amplification are fewer in number than those not taking part.
4. A process according to any of the claims 1-3 wherein the fragmenting of DNA is achieved with three restriction enzymes, two of which recognise a 6 bp sequence and the third recognises a 4 bp sequence.
5. A process according to any of the claims 1-4 wherein at least one primer is labelled.
6. A process according to anyone of the claims 1-5 wherein at least one primer further comprises one or more selective nucleotides, leading to the preferred amplification of a subset of all DNA fragments containing the adapter to which the primer was complementary.
7. A process according to anyone of the claims 1-6 wherein said DNA is cDNA.
8. A process according to any of the claims 1-7, wherein one primer is not complementary to the known sequence of the

adapter(s) but instead is complementary to a different known sequence.

9. A kit useful for performing a process according to anyone of the claims 1-8 comprising at least one method for fragmenting DNA to produce a collection of DNA fragments with at least three different types of ends.

10. A DNA-fingerprint generated using a process according to anyone of the claims 1-8 or a kit according to claim 9.

Figure 1, Graphic outline of the DNA-SALSA method.

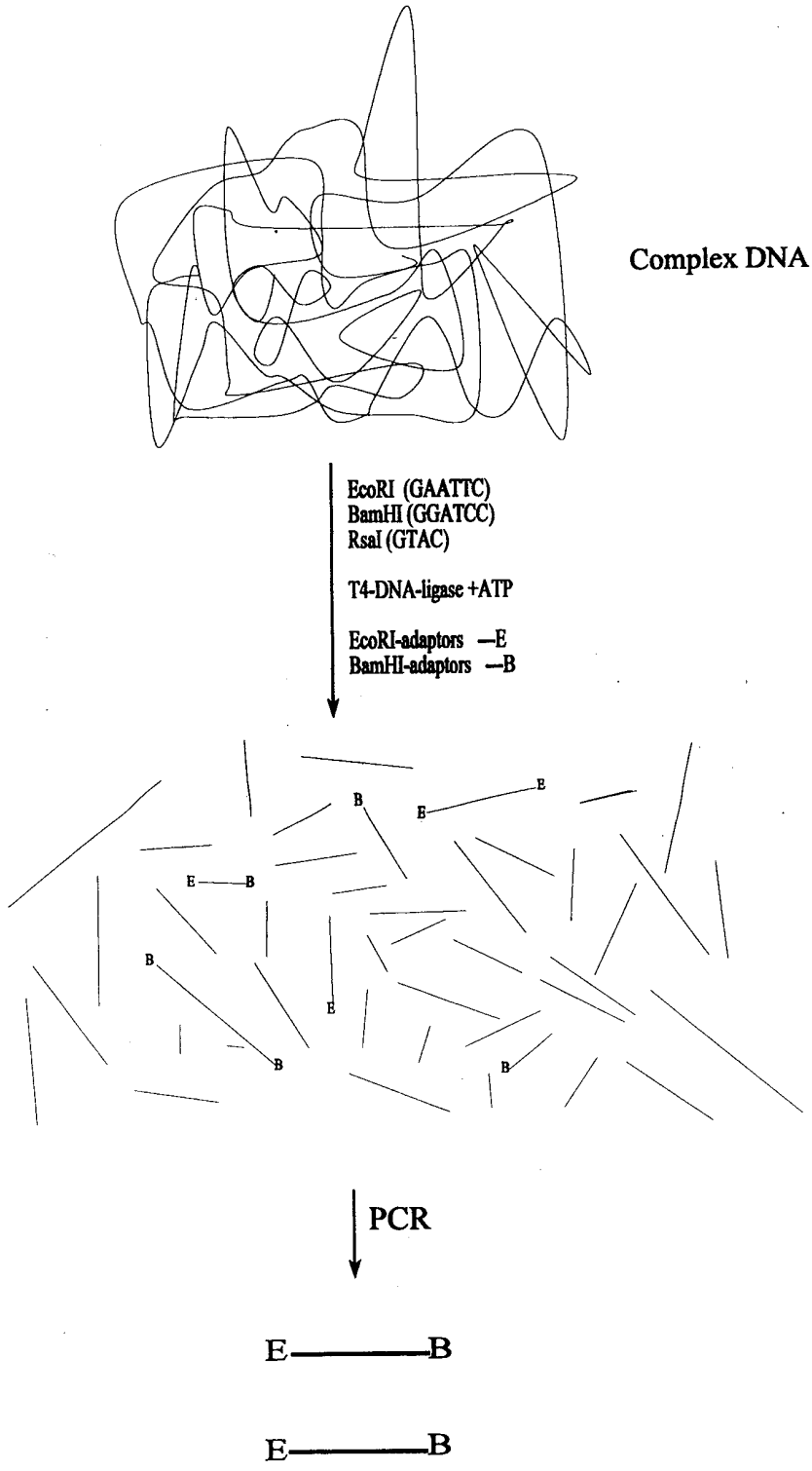


FIGURE 2.

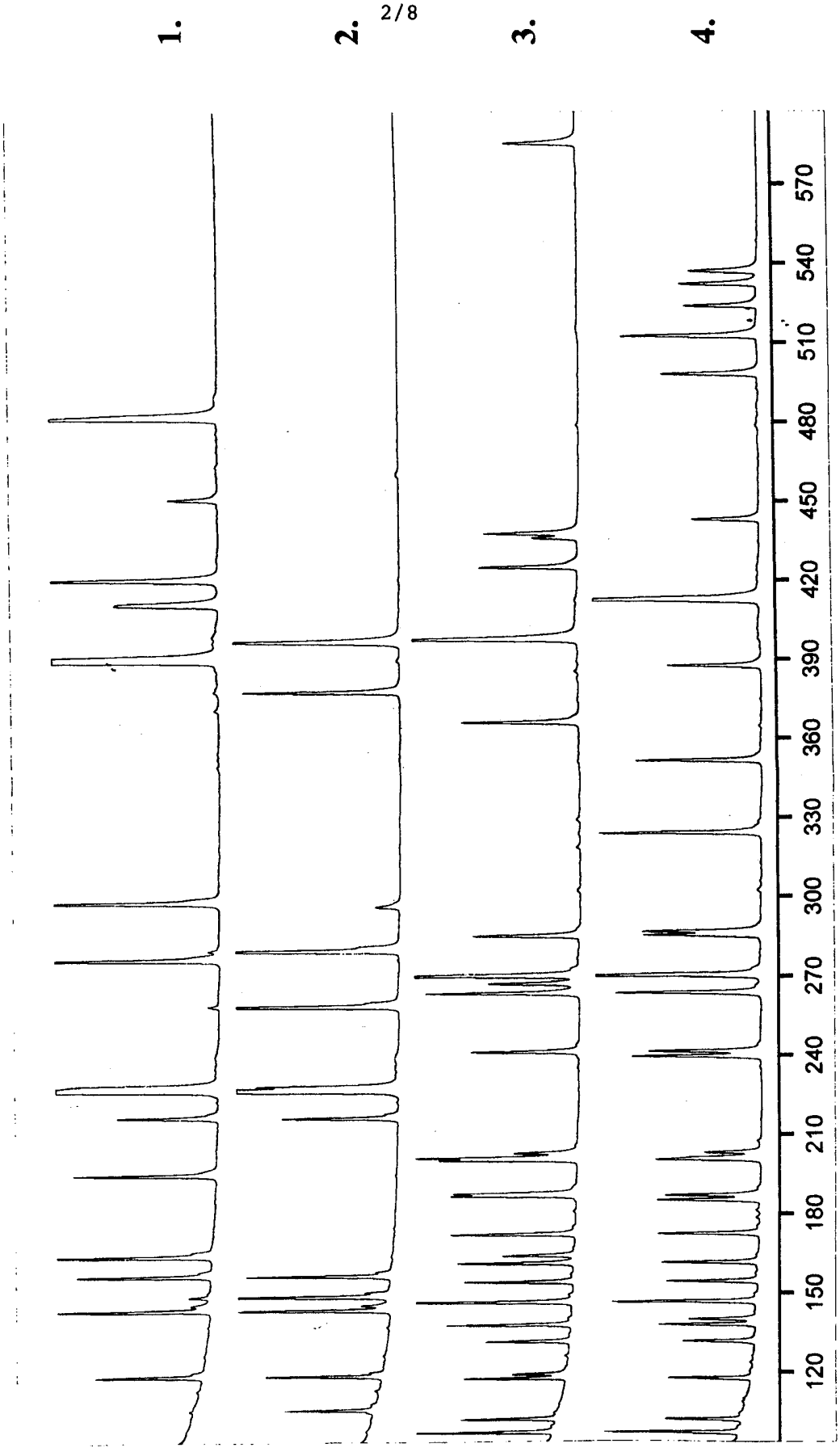

Lane
1.
2.
3.
4.
120
150
180
210
240
270
300
330
360
390
420
450
480
510
540
570

FIGURE 3.

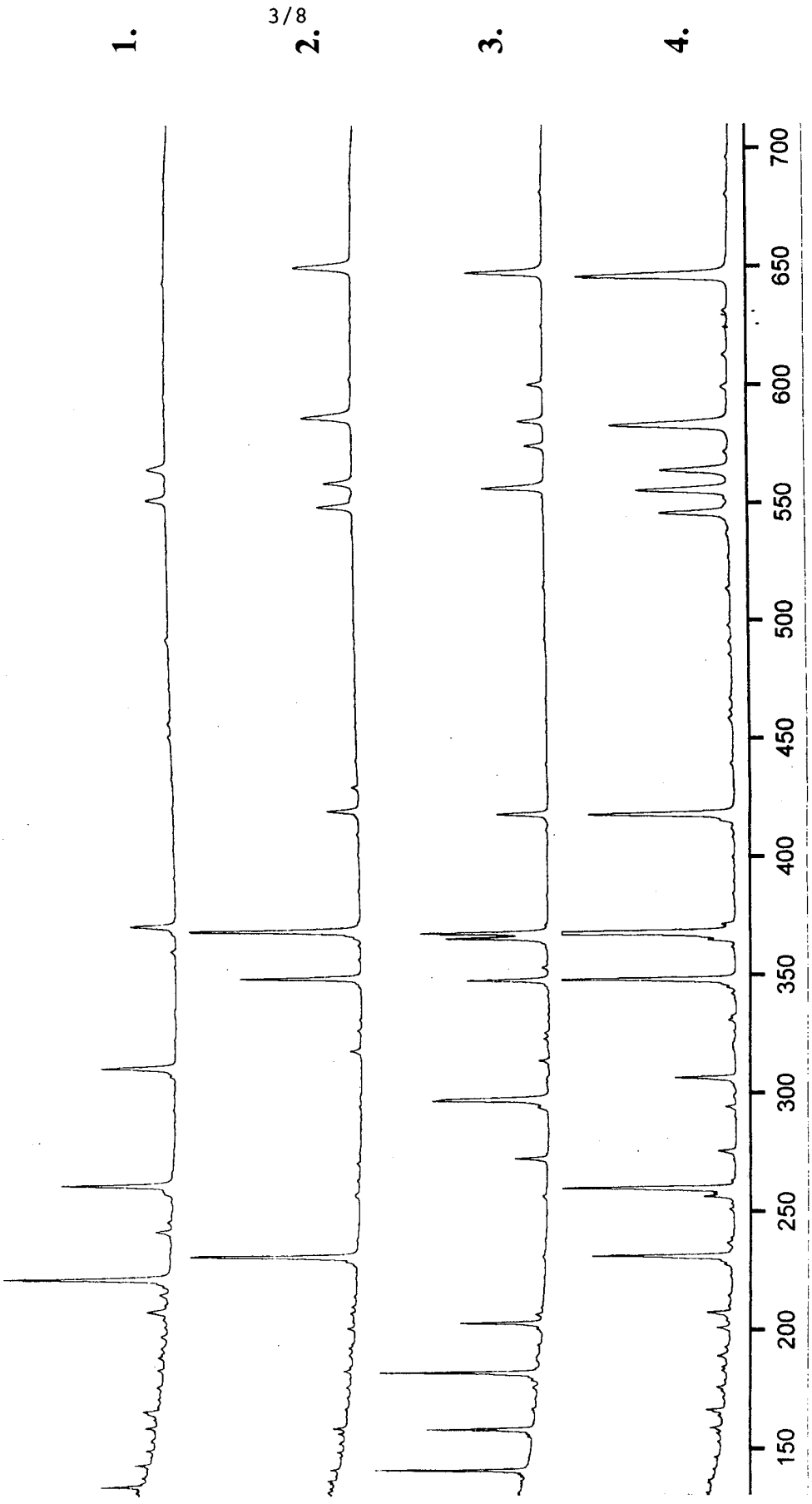


FIGURE 4.

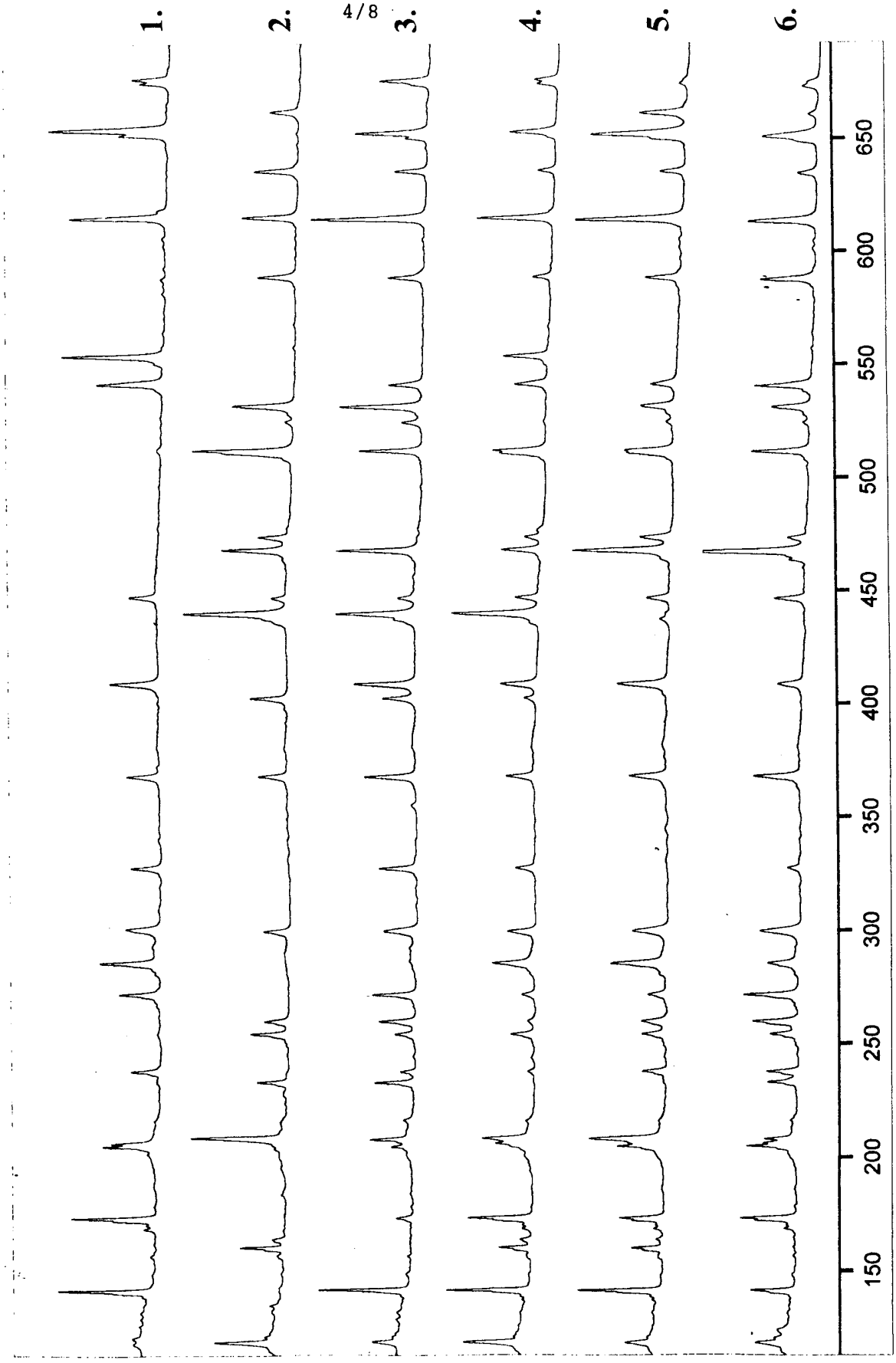


FIGURE 5.

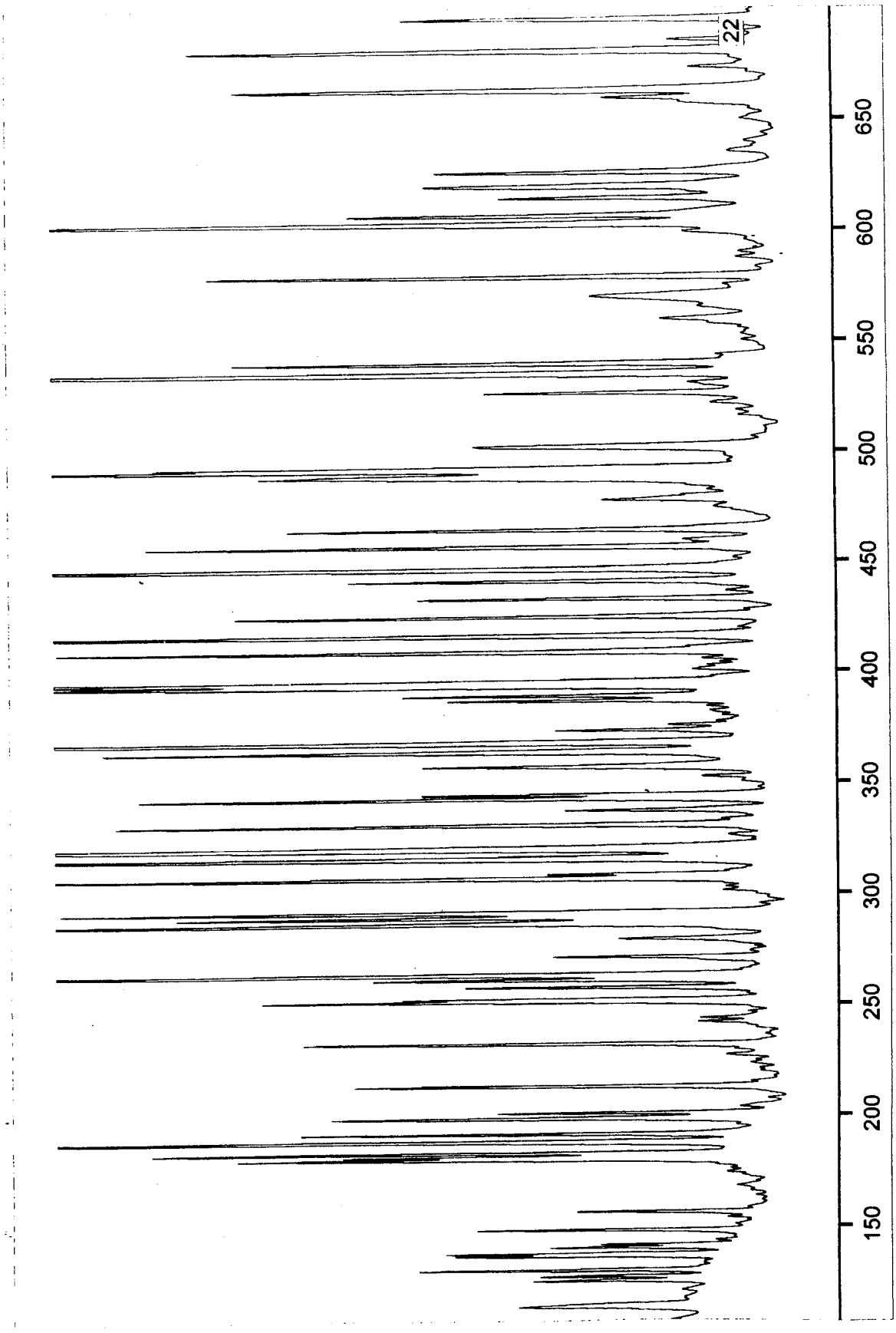


FIGURE 6.

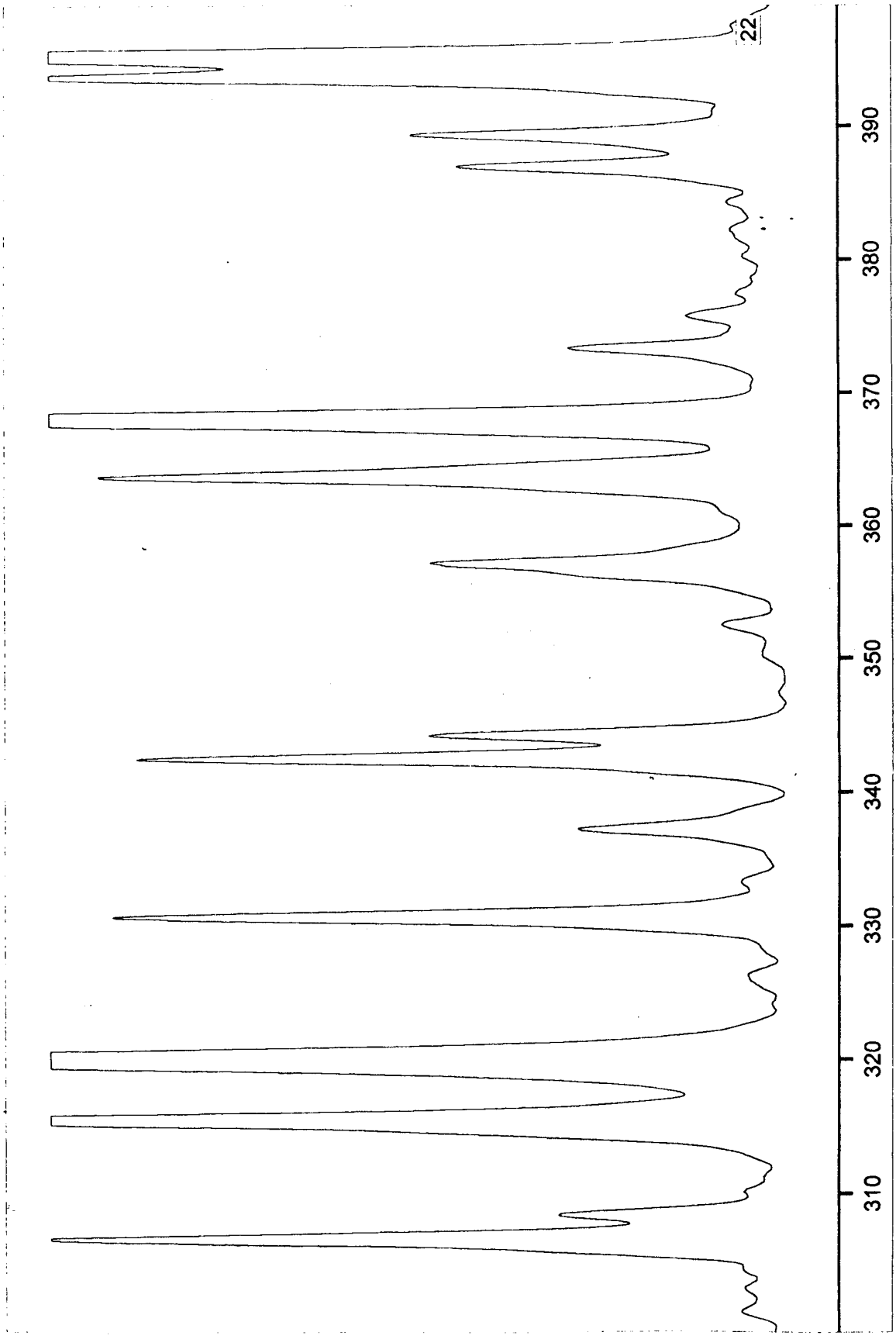
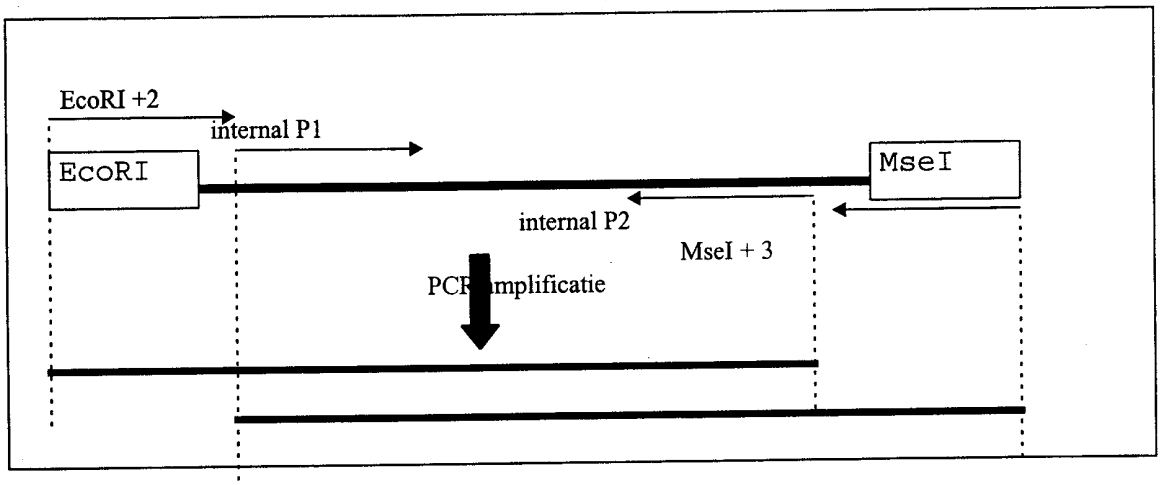


Figure 7


EcoRI +2
internal P1
EcoRI
MseI
internal P2
MseI + 3
PCR amplificatie

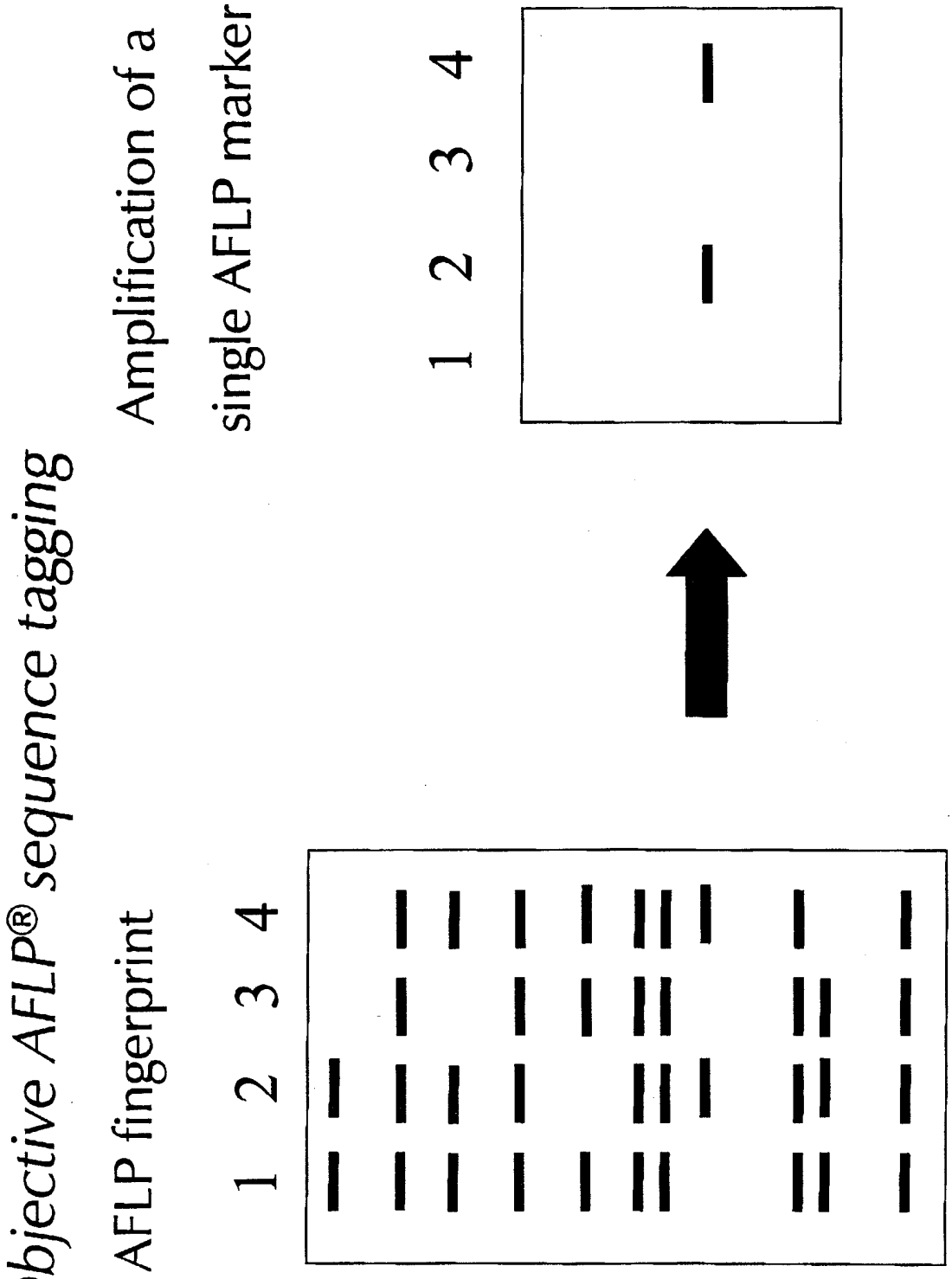

Objective AFLP® sequence tagging
AFLP fingerprint
1 2 3 4
Amplification of a
single AFLP marker
1 2 3 4

Fig. 8

SEQUENCE LISTING

(2) INFORMATION FOR SEQ ID NO:1:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

FITC-GGCGTCGAGACTAGACCG

(2) INFORMATION FOR SEQ ID NO:2:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GACGCGCCAGCAAGATCC

(2) INFORMATION FOR SEQ ID NO:3:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
TCGTAGTGGGCCGCGACCAC

(2) INFORMATION FOR SEQ ID NO:4:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
AATTGTGGTCGCGGCCC

(2) INFORMATION FOR SEQ ID NO:5:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
GTTGACGCCAGCCCGATACATG

(2) INFORMATION FOR SEQ ID NO:6:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

TATCGGGCTGGCGTC

(2) INFORMATION FOR SEQ ID NO:7:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
FITC-GTGGGCCGCGACCACAATTCT

(2) INFORMATION FOR SEQ ID NO:8:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
GACGCCAGCCCGATACATGC

(2) INFORMATION FOR SEQ ID NO:9:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
ACGAAGTCCCGCGCCAGCAA

(2) INFORMATION FOR SEQ ID NO:10:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
GATCTTGCTGGCGCGGG

(2) INFORMATION FOR SEQ ID NO:11:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
GTCCCGCGCCAGCAAGATCA

(2) INFORMATION FOR SEQ ID NO:12:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO: 12:
ACGTTGTGGCGGCGTCGAGA

(2) INFORMATION FOR SEQ ID NO:13:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO: 13:
CTAGTCTCGACGCCGCC

INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/NL 99/00643

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12Q1/68

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 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)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SIMONS G ET AL: "AFLP-BASED FINE MAPPING OF THE MLO GENE TO A 30-KB DNA SEGMENT OF THE BARLEY GENOME" GENOMICS, vol. 44, no. 1, January 1997 (1997-01), pages 64-70, XP002049472 see page 2: "AFLP Analysis"	1-3,5-10
X	WO 96 17082 A (DU PONT ;MORGANTE MICHELE (IT); VOGEL JULIE MARIE (US)) 6 June 1996 (1996-06-06) page 29 -page 33, line 4	1-3,5-10
X	WO 98 30721 A (PIONEER HI BRED INT ;BIRO RONALD L (US); FEAZEL RHONDA (US); HELEN) 16 July 1998 (1998-07-16) page 26, line 20	1-3,5-10

-/--

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

[X] Patent family members are listed in 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
16 February 2000	23/02/2000

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, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Osborne, H

2

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Inter ...nal Application No
PCT/NL 99/00643

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JANSSEN P ET AL: "Evaluation of the DNA fingerprinting method AFLP as a new tool in bacterial taxonomy" MICROBIOLOGY, vol. 142, no. 7, July 1996 (1996-07), page 1881-93 XP002097675 page 1891, paragraph 2 - paragraph 3; table 2	1-3,5-9
A	GB 2 295 228 A (UNILEVER PLC) 22 May 1996 (1996-05-22) page 8, line 14	1-10
A	GB 2 295 011 A (UNILEVER PLC) 15 May 1996 (1996-05-15) page 4, line 23 - line 31	1-10
A	WO 98 08981 A (LIFE TECHNOLOGIES INC) 5 March 1998 (1998-03-05) the whole document	1-10

2

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inter nal Application No
PCT/NL 99/00643

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9617082 A	06-06-1996	AU 704660 B	29-04-1999
		AU 4367496 A	19-06-1996
		DE 69507646 D	11-03-1999
		DE 69507646 T	16-09-1999
		EP 0804618 A	05-11-1997
		JP 10509594 T	22-09-1998
		NZ 298236 A	28-01-1999
		US 5955276 A	21-09-1999
WO 9830721 A	16-07-1998	AU 6024598 A	03-08-1998
GB 2295228 A	22-05-1996	NONE	
GB 2295011 A	15-05-1996	NONE	
WO 9808981 A	05-03-1998	AU 4174397 A	19-03-1998