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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 February 2003 (27.02.2003)

PCT

(10) International Publication Number
WO 03/016546 A1

(51) International Patent Classification[7]: C12P 19/34, C12Q 1/68, C07H 21/04

(21) International Application Number: PCT/US02/26670

(22) International Filing Date: 21 August 2002 (21.08.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/313,912 21 August 2001 (21.08.2001) US

(71) Applicant *(for all designated States except US)*: **FLINDERS TECHNOLOGIES PTY LTD.** [AU/AU]; Mark Oliphant Building, Laffer Drive, Science Park Adelaide, Bedford Park, S.A. 5042 (AU).

(71) Applicant *(for MW only)*: **KOHN, Kenneth, I.** [US/US]; Kohn & Associates PLLC, Suite 410, 30500 Northwestern Highway, Farmington Hills, MI 48334 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **BURGOYNE, Leigh, Alexander** [AU/AU]; Flinder Technologies Pty Ltd., Mark Oliphant Building, Laffer Drive, Science Park Adelaide, Bedford Park, S.A. 5042 (AU).

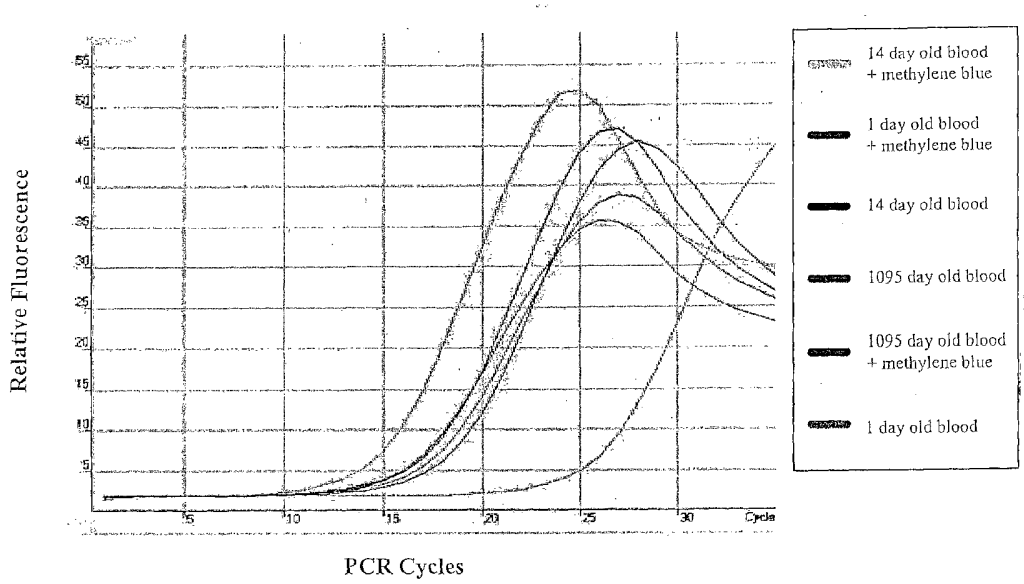
(74) Agent: **RINALDO, Amy, E.**; Kohn & Associates, PLLC, 30500 Northwestern Highway, Suite 410, Farmington Hills, MI 48334 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, 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, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: METHOD AND DEVICE FOR SIMULTANEOUSLY MOLECULARLY CLONING AND POLYLOCUS PROFILING OF GENOMES OR GENOME MIXTURES



PCR Cycles

(57) Abstract: A method for amplifying genetic material by amplifying the genetic material using a single primer sequence. A detector for detecting pathogens in a sample, the detector including a single primer sequence for use in amplification reaction whereby the primer sequence amplifies genetic material of a pathogen thereby detecting pathogens in a sample. A kit for performing the above method including a single primer sequence and a device for amplifying genetic material. A computer program for creating the primers for use in the above methods.

WO 03/016546 A1

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

METHOD AND DEVICE FOR SIMULTANEOUSLY MOLECULARLY CLONING AND POLYLOCUS PROFILING OF GENOMES OR GENOME MIXTURES

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention generally relates to the fields of molecular biology and nucleic acid analysis. More specifically, the present invention relates to a method of genetic analysis related to arbitrary sequence oligonucleotide fingerprinting.

2. Description of the Related Art

Molecular cloning is the process of selecting a nucleic acid sequence and amplifying that sequence many times. Profiling is the process of selecting a small subset of sequences out of a genome.

These processes are used for detection and identification of organisms that are either novel to the world, or are previously described organisms whose presence are totally unsuspected in a serum or plasma. Detection and identification of novel pathogens are currently accomplished by formally detecting the genome of the pathogen with specific sequences from known regions of the genome using methods such as those disclosed in U.S. Patent No. 6,255,467.

Microbial identification, i.e. identification of bacterial, viral, and mycotic species, strains, and subtypes, is a key concern in clinical microbiology for diagnosis of infectious disease, selection of effective pharmaceutical treatment, and epidemiological investigation of the source and spread of infectious disease. Microbial identification is also a vital requirement in the detection and management of biological warfare agents. Microbial identification is important in agricultural,

industrial, and environmental biomonitoring. For example, microbial identification can be used for the detection of pathogens that reduce agricultural productivity as well as for microbes that add nutrients to soil, in order to monitor industrial bioprocesses and assess biodegradation capacity in soil and waste treatment facilities.

Analysis of gene expression is another area that requires new methodologies in order to function more effectively. Transcriptional profiling, i.e., analysis of the relative abundance of messenger RNA (mRNA) transcribed from different genes, is critical to understanding patterns of gene expression that are associated with all biological processes including development, differentiation, response to environmental stresses, and other cellular and organismal functions of interest. The ability to analyze patterns of gene expression can lead to discovery of new genes associated with biological processes. A detailed understanding of gene regulation at the level of transcription is also a significant concern of the pharmaceutical industry. The understanding of gene regulation enables the identification of genetic targets for drug development that can lead to understanding the heterogeneous responses to pharmaceutical interventions. Transcriptional profiling is currently conducted by the techniques of "differential display" (Liang, P. and Pardee, A. B. (1992) Science 257:967-971; Liang, et al., (1994) Nucl. Acids Res. 22:5763-5764; Prashar, Y. and Weissman, S. M. (1996) Proc. Nat'l. Acad. Sci., U.S.A. 93:659-663.) and "representational difference analysis" (Hubank, M. and Schatz, D. G. (1994) Nucl. Acids Res. 22:5640-5648; Lisitsyn, N. A. (1995) Trends Genet. 11:303-307), both of which involve PCR, gel electrophoretic analysis of DNA fragments, and a variety of other complex manipulations. A need clearly exists for new technology, an alternative to DNA arrays, that enables more robust, rapid, and cost-efficient methods for analyzing a very large number of gene transcripts in a short period of time.

An example of such a method used in the field of detection is random amplified polymorphic DNA (RAPD) marker analysis. The method utilizes a single, short primer PCR with genomic DNA. The single, short (8-10 mer) primers have an arbitrary sequence and generate a product that can be used in gel electrophoretic

fingerprint analysis to generate numerous polymorphic markers (Williams *et al*, (1993) Methods in Enzymol. 218: 704-740; McClelland & Welsh, (1995) pg 203-211. In: Dieffenbach, C. W., Dveksler, G. S. (Eds.) PCR Primer – A Laboratory Manual. Cold Spring Harbor Laboratory Press, USA.).

Additionally, amplified fragment length polymorphism (AFLP) uses a single primer to profile nucleic acids. The primer is ligated to restricted fragments and has different principles of amplification to those disclosed herein (Vols *et al*, (1995) Nucl. Acids Res. 23(21): 4407-4414; Gibson *et al*, (1998) *J Clin Micro.* 36(9):2580-2585).

Microbial identification typically involves time-consuming and expensive culturing and biochemical procedures, as well as costly and complex immunological tests. DNA sequencing and PCR analysis also can be performed to achieve accurate microbial identification and typing, but similar current DNA typing procedures, these microbial DNA diagnostic tests require pre-knowledge of the sequences expected to be found so that PCR primers can be designed, or sequences prepared, for use as array points. When the sequences are used as point arrays, classic identification studies need to be performed, often by culturing the organism followed by classic sequencing of the genome.

The identification of highly divergent or novel genomes is typically performed by inserting DNA into a vector such as a plasmid or a virus and then selecting clones randomly and sequencing the inserts. This is time-consuming, usually requiring weeks to be completed. Additionally, the method requires a relatively large amount of the nucleic acid. Because of these limitations, the method is almost never used in routine diagnostics.

Although not suitable for routine diagnostics, the cloning and sequencing procedure of traditional molecular biology does provide sequences sampled from the subject genome that can be observed and fully analyzed, even though the sequences have never been previously seen. A limitation of this approach is that culturing unknown organisms to acquire sufficient nucleic acid to perform the procedure is unreliable.

In summary, most current technologies for the "rapid" identification of sequences requires precise information about a sequence, or identification of the sequence following a long and tedious procedure. Thus, the presence of a pathogen of completely unknown sequence is not routinely detectable because identification by PCR requires known, stable regions of the candidate genome to act as sources for the primer sequences. Also, the identification of sequences on a sequence-array requires a very large menu of candidate sequences to be physically attached to that chip at individual spots. Therefore, neither of these techniques is suitable for highly divergent or novel genomes.

It would therefore be useful to develop a method that does not require culturing of an organism to obtain conventionally useable amounts of DNA, and that is much faster and more readily automatable than the methods disclosed above. It also would be useful if this method could deliver an output of sequences that can be analyzed against the world's databases for match or non-match and can be analyzed for relationships, even distant relationships, to known sequences.

It would also be useful to develop an effective method for rapidly sampling and sequencing molecularly cloned samples for rapid identification, or if novel, classification of species, strains, and sub-types, and for rapid transcriptional profiling.

SUMMARY OF THE INVENTION

According to the present invention, there is provided a method for amplifying genetic material by amplifying the genetic material using a single primer sequence. A detector for detecting pathogens in a sample, the detector including a single primer sequence for use in amplification reaction whereby the primer sequence amplifies genetic material of a pathogen thereby detecting pathogens in a sample. There is also provided a kit for performing the above method including a single primer sequence and a device for amplifying genetic material. Finally, there is

provided a computer program for creating the primers for use in the above methods.

BRIEF DESCRIPTION OF THE DRAWINGS

Other advantages of the present invention will be readily appreciated as the same becomes better understood by reference to the following detailed description when considered in connection with the accompanying drawings wherein:

Figure 1 is a graph showing the methylene blue photolysis effect versus archival time of human DNA on filter paper;

Figures 2 A through D are photographs and graphs showing the results of amplification of genetic material using the single primer method of the present invention;

Figures 3 A through D are photographs and graphs showing the results of amplification of genetic material using the single primer method of the present invention, Figure 3D shows the amplification protocol;

Figure 4 is a photograph showing the results of amplification of DNA using four stable DNA polymerases;

Figures 5 A and B are photographs showing in Figure 5A an agarose gel stained with ethidium bromide and in Figure 5B x-ray film exposed to Southern blot hybridized with ^{35}S dATP labeled soil amplification; and

Figure 6 shows a block diagram of the devices of the present invention.

DESCRIPTION OF THE INVENTION

Generally, the present invention provides a method and kit for amplifying RNA and/or DNA in a sample while simultaneously producing molecular clones that also constitute a profile of that sample. The method can be used for detecting illness and the presence of bacteria or other pathogens. The method also can be used for agricultural purposes such as testing for bacteria in soil samples or other similar purposes.

A "profile" as used herein is any set of sequences extracted from a large number of sequences such that the extracted sequences have some utility (e.g.

criminal detection) and are particularly advantageous in that they are easily amplified and/or have some high utility.

In detecting illness, the present invention functions by detecting diseases based upon changes in the amounts of nucleic acids that are present in plasma, agricultural samples, soil, or other samples. Examples of diseases that can be detected include infectious bacterial and viral diseases.

In agricultural applications, the present invention functions by detecting foreign nucleic acids that should not be found in the sample. An example of an agricultural or forensic application includes, but is not limited to, biologically profiling soils from minute samples of soil.

The method and kit of the present invention differ from many of those found in the prior art because the prior art methods and kits for amplifying RNA and/or DNA often result in what are known to those of skill in the art as "primer-dimers" and other primer concatenates. Primer concatenates are a range of short to very long products of the primers formed by template switching of polymerases of which the smallest member is a primer-dimer. These artifacts are artificial by-products of the PCR and reverse transcriptions and they quench amplification. Primer-dimers and primer concatenates generally are undesirable by-products of polymerase reactions. Therefore, the present invention is beneficial over the amplification reactions of the prior art because the present invention is extremely resistant to the formation of primer-dimers and primer concatenates.

The present invention is able to overcome the problems of the prior art by utilizing primers that readily bind to nucleic acids, but have minimal specificity for defined target sequences. The primers also do not readily associate in such a way as to allow amplifiable concatenations, even at low annealing temperatures. Typically, these primers are used for the first primer loadings, are approximately 16-30 bases long, and have properties as described below.

The prior art discloses the use of a single primer that either requires the use

of a ligase or sequence specificity for the sequence to be amplified. Unlike the prior art method patents, the present invention does not require the use of a ligase nor does it require significant sequence specificity from the primers. Further, single primers are used instead of arrays of single primers as are required by the prior art. By using a single suitable primer sequence and long extension times, the PCR can be used to sample a low concentration of nucleic acid molecules in a sample without assuming or knowing anything about the sequences included therein. A long extension time can be defined as a time long enough for at least 2kb of extension by the polymerase in the amplification mixture. If many amplification cycles are used, then the method of the present invention also enables the investigator to profile the sequences found in the sample such that the profile becomes progressively clearer and simpler with an increased number of amplification steps. Additionally, the present invention does not require large amounts of nucleic acid, as are usually required by two-primer amplifications, in order for amplification to take place. The applicability to low amounts such as, for example, that in 1µl of clean plasma or serum from a healthy person, is important because this development is useful in allowing for automated illness determination.

By utilizing a suitable, single primer sequence and short extension times, PCR also can be used to sample a very small amount of nucleic acid molecules without assuming or knowing anything about the sequences included therein. A short extension time is defined as the time it takes for the PCR system to produce approximately 200 base-pair extensions in that particular amplification mixture. The PCR amplifies DNA or RNA post reverse transcriptase. If RNA is to be amplified, the PCR is preceded by reverse transcriptase copying of the RNA to DNA, using the same single template as is used in the PCR that follows the copying. If single-stranded species, either RNA or DNA, is to be selectively favored in the amplification in the first extension cycle, either reverse transcriptase or PCR is not initiated by, nor preceded by, a denaturation step.

Many amplification cycles imposed on the very small DNA samples disclosed above enable the method to become highly selective in the amplification of sequences. This selectivity does not arise from the same principle as does the

selectivity of conventional PCR. The method instead selects particular sub-sequences of the interprimer sequence to amplify in a process termed structurally mediated interprimer selectivity (SMIPS). The selection is based upon the properties of the amplified section instead of the complimentarity of primers for target sequences as is found in conventional PCR. Importantly, the method creates a relative selectivity for long sequences. Additionally, the amplifications avoid, or at least minimize, any target sequence specificity for primers. The SMIPS process is accomplished when primers are first annealed for one or two cycles at very low specificity so that there is little or no target sequence specificity. This step is then followed by a series of amplifications that use a very high specificity of amplification such that there is no more initiation from the original template. The products of the initial amplification that preferentially amplify are those with interprimer sequences highly favorable to amplification. Therefore, the method preferentially amplifies sample sequences from more complex genomes than from very simple genomes because the more complex a genome, the higher the probability of unusually amplifiable interprimer sequences are contained therein.

Using a single primer versus a pair of primers stops the selectivity of the PCR from being dominated by the homology of the primers to the template (the traditional selectivity of the PCR) and instead makes selectivity dependent on the properties of the amplified section between the primers. This is because, with single primers, the ends of the amplification product are self-complementary. During the amplification, the complementary nature of these ends causes hairpinning and hairpinning can grossly inhibit amplification in the prior art. The prior art has overcome this problem by using two primers and/or very short primers to prevent the occurrence of hairpinning. The present invention overcomes these problems by using the primers disclosed both above and in the examples provided below. The use of single primers causes the species being amplified to have complementary termini. These termini normally interfere with amplification in both a size-dependent way such that large inserts are favored and in an interprimer sequence-dependent way such that sequences that fold to keep the ends apart are favored.

Amplification favors interprimer sequences that, when single-stranded, fold

such that their termini are held well apart. This is a strange property and is only well developed by relatively few sequences, even in a very complex genome, and is often not found in a simple genome. Further, the section of sequence between the primers commonly has a high tendency to fold in such a way as to positively oppose the proximity of its ends. This is the main basis of the "profiling" or selectivity disclosed herein.

By using combined molecular cloning and profiling (sample-amplifying), pathogenic genomes that are plasma or serum-loaded or are from other sources such as soils, then can be estimated by DNA-array-related technology or by scanning electrophoretograms or, as in this disclosure, by absolute sequencing. In traditional PCR, one primer pair gives one product so a profile requires multiple primers in a multiplex. In profiling using single primers, a single primer gives many products that have similar amplification properties characteristic of the section between the primers.

The principle of profiling complex genomes, such as that of humans as understood in the forensic and related sciences, is the selection of a few alleles for amplification that are peculiarly advantageous for particular applications. The present invention utilizes the single-primer long-extension reaction with many cycles to naturally accomplish this principle by selecting for strange, single-stranded structures. The application of this principle creates a base-process for automated molecular cloning and sequencing from unknown genomes.

The method of the present invention is a very high amplification-PCR reaction with single primer that can be preceded by reverse transcription with the same primer. The single primer of 16 to 30 bases long is used at low stringency, such that zero homology or less than 6 to 8 bases of the template have homology to the primers' 3' end that serves as an initiation site. The rest of the primer serves to inhibit amplification of primer-dimers and primer concatenates.

More specifically, the amplification first amplifies many sites on the template, but then progressively favors an ever-smaller number of sites so that a limited

number of multiple products from a complex genome slowly resolve into progressively simple banding patterns. The products thus evolve during the amplification from evenly polydispersed distributions of size to a progressively simpler subset of bands. These products comprise a profile in that they are a complex mixture of products that are diagnostic of the genome or genomes from which they originated

When many rounds of amplifications are performed, the process becomes strongly selective for the properties of the sequences between the primers. Such properties can include, but are not limited to, that is single stranded nucleic acid under the prevailing temperatures of its primer-loading stage, and in amplification, folds so as to most effectively hold the sequences termini apart. This property favors primer loading. This is only one specific condition/property; so the larger a genome, the higher the numbers of sequences within the genome that fulfill this condition. A small genome only amplifies well if it possesses a sequence that folds appropriately when single stranded, or is amplifying in the absence of more complex genomes. Therefore, the process statistically samples from large genomes.

By using the same primer to prime the reverse transcriptase as is used for the PCR, a strong favoritism is obtained that preferentially amplifies RNA sequences over DNA sequences. Another effect is that when the RNA is primed non-specifically with the single primer, the copy of the RNA, with the single primer now grafted onto it, preferentially primes its complementary genes in the contaminating DNA such that the DNA 5' to the original RNA tends to be preferentially entrained in the subsequent PCR. As the use of monoprimers causes the species being amplified to have complementary termini, all amplified species interfere with their own amplification, but in a size-dependent way such that large inserts are favored. The use of long extension times can thus support a strong selectivity for long product length.

Control of the specificity of the primer allows control of the average number of initiation sites per kilobase of any complex population of sequences. The method amplifies mixed genomes in such a way that at first it has selectivity for the most

common genomes in a mixture. An example of such a genome is any simple virus genome mixed with a larger amount of a mixture of human DNA. On a per mole basis, the simple genome usually has many more copies than any one species of the complex genome (e.g., human genome); then later in the amplification, selectivity shifts to the most complex genomes. The selection of parasitic genomes is enhanced by selectivity for single-stranded genomes due to not denaturing the template before the first extension because this allows preferential priming of single-stranded nucleic acid over double-stranded nucleic acid.

The present invention can be also used to amplify nucleic acid from any parasite in the plasma or serum and can be made to favor RNA sequences. The methods and products of the present invention can be used to detect known or unknown virions, either RNA or DNA, with equal speed and ease. The method can also detect bacteria.

Combining the ability to detect illness as disclosed in the previous embodiment with the ability to identify pathogens at high concentration provides broadly based objective evidence of illness for legal or industrial purposes. This differs from the prior art applications of PCR or RT-PCR in association with the detection of a parasitic organism in that the methods of the present application are applicable to the association of the presence of one or more organisms as well as evidence of tissue damage whereas the prior art methods are not applicable to said association.

The ability to monitor the progression of disease from the sequence drift of the pathogen is beneficial in the fight against infections. With regard to HIV patients, the determination of the sequence polymorphism of the population of virions in the serum or plasma of a patient enables researchers to determine much about the state of the disease and its prognosis.

The methods and products of the present invention can be used for field studies of virions or routine diagnostics for general practice surgeries. In mass surveys, the detection of sequence variants of known virions or previously

undetected virions or other nucleic acid-containing moieties in biological fluids by a protocol that is highly automatable, and thus amenable, to massed screenings and discovery is beneficial.

The invention is particularly applicable to systems that handle the acquisition and analysis of complex data in databases that associate clinical records with molecular data. Analysis of the amplification products by various means is common to the nucleic acid field. Examples of such amplification products include, but are not limited to, a "DNA chip," high resolution gels with data acquisition systems, post-chip technology, on-line sequencing technology, or any other suitable technology known to those of skill in the art.

The methods of the present invention can be also used in conjunction with a material that can store genetic material. This material can be beneficial because it is amenable to distance-collection and is highly automatable with an extremely useful and very broad application. This methods-material combination combines the following:

- Processing for RNA occurring on the storage media.
- Low specificity, high-gain amplification using very few primers and, in the preferred and demonstrated version, long-range PCR. (Importantly, not random primers.)
- The degree of specificity can be optimized for general use from the choice of amplification conditions, generalized to the choice of the contours of an amplification-conditions ramp.
- Final data analysis by nucleic acid arrays or on-line sequencing technology.

The present invention also allows open-ended accumulation of sequence libraries for use on chip-style devices.

The present invention can be used for measuring levels of nucleic acids. The methods measure levels of nucleic acids more sensitively than current technology, without respect to a specific organism. In other words, even though the methods

use a single primer, the scope of organisms includes all organisms with nucleic acid, including but not limited to, virions or bacteria in plasma or serum. The methods do have a statistical selectivity for the more complex genomes. The methods therefore only leave out non-nucleic-acid-infective entities such as true prions.

The methods can be also used to objectively record and catalogue large numbers of previously unidentified organisms as gel patterns for future reference. Typing complex mixtures of organisms can be also accomplished by the methods of the present invention. This includes characterizing the nucleic acids from soils for forensic purposes.

The theory of primer design is that they are to have a very high C (cytosine) content with very low G (guanine) content, such that their double-stranded products have a high melting point with the primer itself having negligible secondary structure. Purines, particularly G, are to be avoided in the five nucleotides at the 3' terminus. This is designed to create amplification products with short, high melting point termini, created by the primer, that act as cassettes for lower melting point nucleic acid sequences; or for sequences that have the property of folding so as to hold the termini apart; or for sequences with a combination of these two properties.

Shown below are primers that fulfill these conditions as well as with examples of primers that do not. The primers used in accordance with the above methods do not deliberately target sequences. This list is included for purposes of illustration and is not intended to be limiting. What homologies exist are fortuitous and undesirable. Examples of primers that have been tested are as follows:

Seq1 ACACCTCCCGCAACCCCC3'; moderate performance (Seq. ID No: 1)
Seq2 ACGCACCCATCCACCCCC3'; moderate performance (Seq. ID No: 2)
Seq3 CCACCCAGCATCACCCCC3'; moderate performance (Seq. ID No: 3)
Seq4 GCCATCCCACCAACCTCC3'; satisfactory performance (Seq. ID No: 4)
Seq5 CCCTCGAACACCACCTCC3'; very good performance (Seq. ID No: 5)
Seq6 ACCCGCCACTCAACCTCC3'; satisfactory performance (Seq. ID No: 6)
Seq7 CGACACCAACCTCCGACC3'; bad performance (Seq. ID No: 7)

Seq8 CCCACAAGCCACTCGACC3'; bad performance (Seq. ID No: 8)
Seq9 CAACCGACTCCACCGACC3' bad performance (Seq. ID No: 9)

Suitable primers (those with the CCTCC 3' end) were generated using a program that randomly generated a particular sequence-type. Sequences rich in C but low in G, combine a relatively high melting point with negligible self-complementarity. Any program can be used for obtaining suitable primers.

With regard to the device that can be used in conjunction with the above methods, there are three mechanization options. The devices are disclosed below. All three options begin with a conventional PCR processing laboratory robot that carries out either a one-stage PCR reaction (devices A and C), or a two-stage amplification (device B). The device is generally shown in Figure 6.

Device A is a nucleic acid profiling machine. This machine resembles a laboratory robot that can handle 96 well plates that are set beside a capillary electrophoresis machine. The laboratory robot is for two-stage PCR that has a minor difference from conventional processing. The robot is attached to a standard capillary electrophoresis device or HPLC device for separating and observing patterns of DNA bands, but not necessarily observing the bands. The device separates the bands of clones and delivers them to an on-line sequencing device such as a mass spectroscope or by the Sanger procedure that delivers the peaks one by one to a multicapillary sequencing device. The first device must either operate so as to separate the strands of the bands being studied, or in the technically simplest version, the preferred version, must operate on bands that are cut by a restriction endonuclease before loading and then separated as single strands. The endonuclease cutting ensures that the fragments are small enough to separate well with current technology and also ensures that the fragments have only one end either probe-labeled or complementary to the single primer that can be required for the sequencing technology of device B or device C. Sequencing can be also performed by the use of a second low-specificity primer loading, with an alternate primer, thus providing each strand with a separate primer for sequencing.

Device B is to be used in line with and subsequent to device A. Device B takes a short, single strand of DNA and sequences it by Sanger sequencing fluorescent primer chain terminating technology or related technology. The restricted band from the gel is ejected into a reaction vessel such as one well of a 96-well plate. A robot then adds a sequencing mixture containing a primer with the same sequence as the single primer used for the original amplification. After sufficient reaction time as is required to develop the spectrum of terminated sequences characteristic of this technology, the mixture is loaded onto a resolving apparatus, for example, a high resolution capillary gel. The sequence data is acquired automatically and compared to the world's database of sequences.

Device C is to be used in line with and subsequent to device A. This also takes a short, single strand of DNA and sequences it. In this case, the sequencing is accomplished by mass spectroscopic (MS) technology. In one example of MS technology, after the small double-stranded band is exposed to a second primer such that a short series of diprimer amplifications can occur to allow ejection from the stage-A separation, one strand of the double-strand fragment to be selected using a reporter-group or handle such as biotin is incorporated into the 5' end of one of the two primers to separate the strands so that only one strand is presented to the MS apparatus and subsequently sequenced. In another version, the strand separation occurs on the first stage of separation, stage A, by using separation conditions such that denatured and partly refolded single strands are separated from each other by the properties of the folding structure they form. In another version of this apparatus, the sequence of the single primer or its complement is used to separate the strands by preferentially binding or delaying the movement of one strand before presenting them to the MS apparatus. In all cases, the resultant sequence data is collected and compared to the world's database of sequences. In an additional version of this apparatus, prior to sequencing, a second low stringency primer loading step is carried out using an alternate sequencing primer. The purpose of this is to introduce a second primer that is different from the first primer, to allow separate sequencing of each strand.

The general method requires the following steps. First, samples are

collected. The samples can be maintained on storage media. Second, the samples are processed via phenolic or other methods with similar purification effect. Third, the sample is reverse transcribed with the protocol set forth in Table 1. Fourth, the PCR protocol set forth in Table 2 is followed. RT-PCR products are ammonium acetate and alcohol precipitated prior to visualization and then run on an acrylamide gel. Fifth and finally, analysis of the amplification products is accomplished using any of the above-described methods.

Table 1 Reverse transcription protocol used to generate cDNA

Temperature	Time
42°C	40 minutes
45°C	5 minutes
50°C	5 minutes
55°C	5 minutes
60°C	5 minutes
95°C	5 minutes

Table 2 PCR protocol used in the first and second round amplifications

Temperature	Time	Cycles
94°C	10 minutes	1
60°C	30 seconds	40
55°C	30 seconds	40
50°C	30 seconds	40
72°C	2 minutes	40

The above discussion provides a factual basis for the use, the methods, and the kits of the present invention. The method used with, and the utility of, the present invention can be shown by the following non-limiting examples.

EXAMPLES

METHODS:

General methods in molecular biology: Standard molecular biology techniques known in the art and not specifically described were generally followed as in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York (1989), and in Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, Maryland (1989) and in Perbal, *A Practical Guide to Molecular Cloning*, John Wiley & Sons, New York (1988), and in Watson et al., *Recombinant DNA*, Scientific American Books, New York, and in Birren et al. (eds) *Genome Analysis: A Laboratory Manual Series, Vols. 1-4* Cold Spring Harbor Laboratory Press, New York (1998) and methodology as set forth in United States Patents 4,666,828; 4,683,202; 4,801,531; 5,192,659 and 5,272,057; and incorporated herein by reference. Polymerase chain reaction (PCR) was carried out generally as in *PCR Protocols: A Guide To Methods And Applications*, Academic Press, San Diego, CA (1990). *In-situ* (In-cell) PCR in combination with flow cytometry can be used for detection of cells containing specific DNA and mRNA sequences (Testoni et al, 1996, Blood 87:3822.).

General methods in immunology: Standard methods in immunology known in the art and not specifically described are generally followed as in Stites et al. (eds), *Basic and Clinical Immunology* (8th Edition), Appleton & Lange, Norwalk, CT (1994) and Mishell and Shiigi (eds), *Selected Methods in Cellular Immunology*, W.H. Freeman and Co., New York (1980).

EXAMPLE 1:

Analysis of Infectious Disease State / Nucleic Acid Estimation

Serological clinical samples of patients carrying hepatitis C (HCV), and human immunodeficiency virus (HIV), as well as from healthy patients were manipulated on FTA® paper (Whatman House, UK, Whatman plc, Whatman International Ltd., Whatman House St. Leonard's Road 20/20 Maidstone Kent ME16 0LS). However, other methods of viral manipulation can be also used. 1mm discs

loaded with clean sera, HCV, or HIV infected sera were routinely used in analysis.

All samples were phenol processed prior to reverse transcription (RT) and PCR, with the same primer, Seq#005, used in both sets of reactions. The RT temperature protocol was as follows: 42°C for 40 minutes, 45°C for 5 minutes, 50°C for 5 minutes, 55°C for 5 minutes, 60°C for 5 minutes, 94°C for 5 minutes. The entire exhausted RT reaction was routinely used as the template for the PCR. The PCR temperature cycling protocol was as follows: 94°C for 10 minutes, 94°C for 10 seconds, 60°C for 30 seconds, 55°C for 30 seconds, 50°C for 30 seconds, 72°C for 2 minutes; for 35 cycles. A second round of PCR with a protocol identical to the first round was often done, using 1µl of the first round amplification as template.

First and second round amplification results were visualized on 7% polyacrylamide gels, which were stained with ethidium bromide and viewed under UV light. Examples of results generated using this procedure are set forth in Figure 2.

Discussion / Conclusion

The interpretation of the data in Figure 2 is not governed by any particular banding pattern, but by the visibility and/or frequency of band formation. The frequency of band formation inversely correlates with relative amounts of free nucleic acids in a given sample. All samples (except where specified) were obtained from sera, and thus were expected to contain large amounts of cell-free nucleic acids (template) due to the cell breakdown that occurs during the clotting process (Lee *et al.*, (2001) *Transfusion.* 41:276-282). Additionally HCV infected sera samples were expected to contain more cell-free nucleic acid due to the cytopathic effect of the virus on hepatocyte cells within the patient. First and second round amplifications generated from HCV and clean sera samples resolved as a smooth polydispersed pattern, with a wide range of molecular weight. This pattern is an indicator of high levels of template.

In a healthy individual, samples of plasma contain minimal cell-free nucleic acids, and thus there is limited template for a PCR reaction. Amplification generated

from clinical plasma samples was visualized as a set of discrete bands, indicating minimal template.

One important practical aspect of estimation of trace template by single-primer non-specific amplifications is that it allows template estimation by two totally different principles: the first being the determination of the amounts of DNA by conventional estimation of the amounts of amplification product after a set standard reaction; and the second being the determination of the amounts of template by usage of the loss of molecular diversity during amplification. Loss of molecular diversity can be estimated by a range of methods known to those in the field. Examples of such methods include observation of renaturing rates of products during the process of amplification; observing the ratio of material found in discrete peaks as compared to that found in background smear in a size fractionation of the products; and usage of loss of molecular diversity as a measure of the amount of original template from any given genomic source.

Amplification from the HIV-infected sera samples produced a discrete set of bands, similar to those observed from plasma. This type of pattern is an indicator of limited template and is not unexpected from this type of sample that is expected to be depleted in cells. A part of the life cycle of HIV is to infect cells containing the CD4+ receptor (Kuby, (1994) Immunology 2nd Edition, W.H. Freeman and Company, USA; Collier *et al*, (1993) Human Virology, Oxford University Press Inc., USA). Thus HIV infects white blood cells and eventually causes cell death. Literature indicates that advanced cases of HIV result in depletion of white blood cells from the blood (Kuby, (1994) Immunology 2nd Edition, W.H. Freeman and Company, USA; Collier *et al*, (1993) Human Virology, Oxford University Press Inc., USA). The cytopathic effect of HIV on white blood cells results in a significant decrease in nucleated cells in the blood, and thus less cell-free nucleic acid produced during clotting.

Further analysis of the first and second round amplification products via sequencing reveals that the type of nucleic acids amplified also divulges useful information about the disease state. Sequences obtained from HIV sera samples contained an increased number that originated from the mitochondria with respect to

the nucleus, in conjunction with a lack of whole cell genomic nucleic acid. Mitochondrial sequences were rarely observed in amplifications from non-HIV subjects. Additionally, sequences generated from HCV infected sera with the non-specific primers specified provided further information about secondary infections, for example sequences were obtained from *Pseudomonas aeruginosa* (Picot *et al*, (2001) Microbes & Infection. 3(2): 985-995), *Neisseria meningtidis* (Parkjill *et al*, (2000) Nature. 404(6777): 502-506) and *Burkholderia cepacia* (Mohr *et al*, (2001) Microbes & Infection. 3(5): 425-435), all of which have been documented as disease-related opportunistic pathogens.

This type of analysis is not limited to infectious disease, but could also be a powerful tool in cancer diagnosis. Published studies on the serum from tumor patients detected tumor DNA (Jahr *et al*, (2001) Cancer Research. 61: 1695-1665), but this failed to give insight into the size or location of the tumor. Study of the sequence content of mRNA fragments in serum or plasma, because of the expected short half-life of RNA, can be a powerful tool in the detection of the tissue that is dying around the tumor and/or the nature of the tumor itself.

Confirmation of the presence of the infectious agent was also observed with a further pathogen specific amplification. Figure 3 displays amplification results obtained when a CVB-4 specific amplification was carried out on a range of products (5µl) from first round PCRs. The pathogen specific reactions generated the 110 bp fragments that indicated the presence of CVB-4. The 110 bp fragment was sequenced and confirmed CVB-4.

The technology described in this patent has novel application to infectious disease diagnostics because, in addition to supplying information about the presence of a particular pathogen, the technology can, on the same blood sample, provide information about the health or disease state of an individual. The technology also has the ability to detect microbial entities that may not be expected or suspected (i.e. secondary infections by opportunistic pathogens). Additionally the technology also has clear applications to the diagnosis of unknown diseases, or in situations where diagnostic tests are not available or do not exist.

EXAMPLE 2:

Detection of DNA Breakage

All blood and bacterial samples used in the study were stored on FTA® paper, however FTA® paper was not essential to generate results, and was merely a convenient medium for the manipulation of biological samples. 1mm discs containing blood samples of age one day, 14 days and 1095 days were used, additionally 1mm disc samples of *E.coli* of one day old were also used.

All samples were phenol processed prior to amplification or methylene blue photolysis. When used, methylene blue treatment consisted of 5µl of methylene blue solution (50µg/ml methylene blue, 2mM TE, pH 7.5) applied to a 1mm disc and then exposed to an 8-Watt white fluorescent lamp for a period of 15 minutes at a distance of 6 cm. The intensity of the lamp output at this distance was 3.02 kWm^{-2} (Murov, (1973) Handbook of Photochemistry, Marcel Dekker Inc, New York). Methylene blue treated samples were further phenolically processed before amplification.

Amplification of the 1mm samples was conducted using the temperature cycling protocol in Table 2, the primer Seq#005 and SYBR® Green (other fluorescent ligands may also be used). Amplification and analysis was carried out on a Rotor-Gene 2000 Real-Time Cycler, with acquisition of fluorescence during the extension phase of PCR using an excitation wavelength of 470nm and an emission wavelength of 510nm. Results can be observed in Figure 1.

Discussion / Conclusion

The data displayed in Figure 1 shows a large discrepancy between the amplification intensity of one day old whole blood + / - methylene blue photolytic treatment. The apparent difference in amplification intensity appears to decrease with storage time on the FTA® paper. Literature indicates that methylene blue is a

photolytic agent capable of causing single and double strand breaks in DNA (Epe *et al*, 1989; Hong, 2000; Schneider *et al*, 1990). The SMIPS amplification process provides a straightforward method for the monitoring of genome integrity and DNA damage.

The observation that the effect of the methylene blue treatment decreased in some sort of relation to storage time on FTA® paper is not unexpected. Evidence from previous studies involved in the investigation of nucleic acid storage on FTA® suggest that variations in temperature storage conditions may possibly induce nucleic acid bond breakage either and/or via DNA conformation changes or mechanical stresses directed from the FTA® paper (Hong, (2000) PhD Thesis, Flinders University of South Australia).

The applications of the methylene blue light induced effects are significant for both commercial and scientific research. It shows the existence of a unique and sensitive method for DNA damage monitoring and detection.

It has been occasionally noted that large templates are strangely bad templates for some PCR primers but this effect has been described in such a way that it is only recognized as an annoying disadvantage of no practical application and it has been described with template specific primers. This application discloses the use of conditions that maximize the effect by using non-template specific primers and conditions.

The detection and monitoring of DNA damage, particularly DNA circulating as a result of cell death against a background of normal DNA in white blood cells can be accomplished by the present invention. Examples of monitoring that can include monitoring tissue breakdown and the monitoring of DNA modifications in "at-risk" professionals (e.g. nuclear industry workers after accidents), or the evaluation of the side effects of radiation therapy and chemotherapy on patients. Analysis of a patient's current level of DNA modification (e.g. accumulated mutations or strand breaks in circulating DNA) aids practitioners in accurately diagnosing the correct levels of treatment without risking the patient. Additionally it is desirable and

essential for preventative medicine in determining if an individual has a low DNA repair capacity, for example individuals with a family history of breast cancer (Leong et al, (2000) International Journal of Radiation Oncology, Biology, Physics. 48(4): 959-965), and is thus at greater risk in the development of a life threatening malignancy.

EXAMPLE 3:

Microbial Population Genetics

DNA Extraction and Amplification: Soil samples were collected from five distinctly different geographical sites. DNA was extracted from each soil sample using the MoBio Soil DNA Extraction kit, (address and contact details) however the kit was a convenience item and other kits and/or methods can be also applied.

A range of DNA concentrations were used as templates for the first round of PCR, 25ng, 2.5ng, 250pg and 25pg (represented as lanes 1, 2, 3 & 4 in Figure 5 respectively). A first round amplification was carried out on each of the DNA concentrations from each of the soil samples using Seq5 as the primer. The PCR protocol was the same as in Table 2. The template for the second round amplification consisted of 5µl of the first round amplification product. The second round amplification used Seq5 as the primer and the PCR protocol in Table 2.

Probe Labeling: The probe was generated by using 5µl of the first round PCR product of soil B-3 (see Figure 5), and re-amplified (using the same primer and protocol) in the presence of 50µCi S^{35}dATP.

Southern Blotting: The blotting was performed as described in Sambrook *et al*, (1989) Molecular Cloning – A Laboratory Manual, 2nd Ed. Cold Spring Harbour Laboratory Press, USA, except the membrane used was ZetaProbe (BioRad).

Probe Hybridization & X-Ray Film Exposure: These methods were performed as described in Sambrook *et al*, (1989) Molecular Cloning – A Laboratory

Manual, 2nd Ed. Cold Spring Harbour Laboratory Press, USA.

Results / Discussion

The amplification patterns observed from the five soil samples are in agreement with previous SMIPS amplification results (e.g. the infectious disease example). Samples containing greater quantities of template generate a larger range of amplification species, and thus molecular-competition sufficient to cause band formation requires more amplification cycles.

The Southern blotting does indicate that amplification species generated from each PCR are unique to the sample of origin, as no cross-reactivity was observed after hybridization (Figure 5). However cross reactivity was detected between PCR from the same soil sample, as the probe generated from soil B-3 also hybridized to other amplification products from soil B.

The evidence in this example are in accord with the expectation that the SMIPS amplifications generate a "profile" of a highly complex and unknown mixture that is a, representative of a given template mixture without necessarily being a full cross section of a given template mixture. It is a profile that is unique to that sample when compared to an identical amplification generated from other template mixtures, even though portions of the templates may share similar sequences.

These unique sequence-characteristics of the amplification of species between complex template samples provide an opportunity for the adaptation of this technology to areas such as microbial profiling and forensics.

Microbial Profiling: The profile generated from organism has the potential to act as a marker of close relatedness as the closer two organisms are genetically related, the more amplified species they have in common.

Forensics: The example provided in Figure 5 demonstrates that the microbial flora present in a given soil sample is the distinctive property of a given soil

at a given site or region. This observation can be a useful tool in a forensic investigation in which samples exist that contain mixtures of nucleic acid that maybe compared to an original source.

EXAMPLE 4:

Quality Control

SMIPS amplifications were carried out on a variety of human DNA concentrations, ranging from 310ng, 31ng, 3.1ng and zero. Amplifications were carried out using four different enzyme suppliers, but all using the Seq5 primer and the PCR protocol in Table 2. The resultant amplification products were visualized on a 1.5% agarose gel and stained with ethidium bromide.

Discussion / Conclusion

The results displayed in Figure 4 are an example in the detection of trace levels of nucleic acids, or quality control in commercial DNA polymerase preparations. The presence of even the smallest concentrations of contaminating nucleic acids in an enzyme preparation can have a dramatic effect on the outcome of the amplification reaction (e.g. production of false and misleading amplification species). Figure 4 displays the effect of contaminating nucleic acid on negative controls. Enzyme mixes from supplier R and supplier A, both have produced false positives, however enzyme mixes from supplier D and T appear to be clean. The amplifications observed from supplier R are thus suspect, as the patterns seen in the negative controls are similar to those seen in the positive samples, making it difficult to discern the desired product.

Another example in which the detection of unknown and unsuspected nucleic acids is important is in health products are for monitoring water and food quality. All unsuspected nucleic acids are undesirable contaminants of injected materials.

It is essential that healthcare products, particularly those that are used intravenously (e.g. antibiotics, vaccines, other blood related products) do not contain unsuspected nucleic acids, for example, fragments of viral or bacterial genomes. The consequences of modulating an infection with extraneous nucleic acids via the use of contaminated healthcare products are large in terms of personal health and potential litigation. Free nucleic acids are well known to transfer drug resistances and virulence between microorganisms.

Free nucleic acids can be incorporated into and expressed in human cells with the possibility of insertional mutagenesis and/or the expression of strange polypeptides with unknown consequences.

The screening of consumable materials such as water and food for unknown nucleic acids is also of prime importance, as the contaminating nucleic acids can be an early indicator of bacterial or viral contamination. Consumption of contaminated food or water is a serious health risk, routine monitoring of foods and water with the SMIPS based technology can significantly decrease the incidences of infection and disease, decreasing the risk to the consumer.

Example 5:

Any program can be utilized to create suitable primers. The program must be able to create a sequence rich in C but low in G. The program can run on any operating system.

The example program, below, is for 18 bases long and is as follows:

'Prog1 - a program in Borland's Turbobasic. Black text is comment and red text is compile-able code.
'This makes primers according to a set of rules.
' The products of this family of primers are to be 18 base pairs long.
'The products are to start with a fixed sequence at the 3' end of (5'CCTCC3')

'The products are to then have 13 additional bases added with the following characteristics: 55% C (7C), 30% A (4A) and 15% G and T (1G and 1T)
'Effective self complementarity is avoided by choice of high C with only 1G .
'There should be no very long runs of homopolymer. (Runs of >3 of any base will cause a candidate sequence to be rejected and runs of A longer than 2 are also rejected.)

'

'START CODE -start by declaring the array stores.

```
DIM PRIM$(21)
DIM R$(400)
```

'Declare the base content as an unrandomised sequence to be used in the ,eg., 13 positions of the 5' end. Place them in the string store, WD$, which will act as the menu of bases to be scrambled.

```
WD$="CCCCCCCAAAAGT"
```

'Declare the name of the output file from the keyboard and set randomiser

```
PRINT "INPUT THE NAME OF THE OUTPUT FILE FOR THE NEW
SEQUENCES.
PRINT "with no more than six characters in the name."
```

'At the same time, initialise the random generator that Borland provides in their software.
'the human delay in inputting the file name allows the internal clock to generate a highly variable seed.

```
PRINT "INPUT THE DESIRED OUTPUT FILE NAME":INPUT NM$
RANDOMIZE TIMER
```

'Clear the screen of the computer

```
CLS:PRINT " "
```

'Begin cycles of producing candidate sequences. Clear the main counter POI% and

```
POI%=0
FOR R%= 1 TO 2000
POI%=POI%+1:IF POI%=>398 THEN GOTO YNDUPFINAL
PRINT "CYCLE NUMBER- ";R%
```

' Begin a cycle by fixing the first five bases from the 3' end of the primer, in the string array store PRIM$(1) where PRIM$(1) is the 3' base.

```
PRIM$(1)="C"
PRIM$(2)="C"
PRIM$(3)="T"
PRIM$(4)="C"
PRIM$(5)="C"
```

'Now fill the rest of the primer/oligo in 13 operations filling oligo places 5 TO 18 randomly using a list of numbers from 1 to 13 and then scramble them.

'first, fill a little array with ordered numbers.

```
FOR J%=1 TO 13:NUM%(J%)=J%:NEXT J%
```

'Now scramble (disorder) these numbers in the little array.

```
FOR JJ%=1 TO 13
AGIN:
SS$=INKEY$:IF SS$<>"" THEN GOTO YNDUPFINAL:'this line of code is
only to allow an overriding termination from the keyboard
```

'call for a random number from 0 to 14 (but reject it if it is not between 1 and 13 inclusive).

```
CALL RAN:IF NU%<1 OR NU%>13 THEN GOTO AGIN
```

'use the random number in a scramble.

```
IHOLD%=NUM%(JJ%):NUM%(JJ%)=NUM%(NU%):NUM%(NU%)=IHOLD%
NEXT JJ%
```

'Now use the small list of scrambled numbers to pick out the bases from the menu of bases in WD$ and place them from the end of the 3' zone to the end of the 5' zone with operations proceeding in the 3' to 5' direction.

```
FOR J%=1 TO 13
PRIM$(J%+5)=MID$(WD$,NUM%(J%),1)
NEXT J%
```

'now write the candidate oligo's sequence in the reverse (the conventional) direction in the array store R$( ) while checking for homopolymer runs.

```
FOR J%=18 TO 1 STEP -1
R$(POI%)=R$(POI%)+PRIM$(J%)
NEXT J%
```

'Now look for any forbidden runs of any homopolymers.

```
IKILL%=0:IKILL2%=0:IKILL3%=0:IKILL4%=0
```

```
FOR J%=1 TO 17
C$=MID$(R$(POI%),J%,1):D$=MID$(R$(POI%),(J%+1),1)
IF C$=D$ THEN IKILL%=IKILL%+1:IF IKILL%>2 THEN IKILL2%=1
IF C$<>D$ THEN IKILL%=0
```

'The next lines are AAA forbidders.

```
IF C$=D$ AND D$="A" THEN IKILL3%=IKILL3%+1:IF IKILL3%>1 THEN
IKILL4%=1
IF C$<>D$ THEN IKILL3%=0
NEXT J%
```

'The next lines execute the rejections on either of these two grounds.

```
IF IKILL2%>0 THEN R$(POI%)="":POI%=POI%-1:GOTO ENDOFRR:'reject a
candidate due to homopolymer content.
IF IKILL4%>0 THEN R$(POI%)="":POI%=POI%-1:GOTO ENDOFRR:'reject a
candidate due to AAA homopolymer content.
```

'Now scan each candidate oligo against all those already listed and reject any sequences already found once.

```
FOR K%=1 TO (POI%-1)
IF K%=1 AND POI%=1 THEN GOTO SKIPIT
IF R$(K%)=R$(POI%) THEN ICOU%=ICOU%+1:' noting the number of
candidates rejected for being duplicates before actually rejecting them in the
next line.
IF R$(K%)=R$(POI%) THEN R$(POI%)="":POI%=POI%-1:GOTO ENDOFRR
SKIPIT:
NEXT K%
PRINT R$(POI%)
ENDOFRR:
IF POI%<0 THEN POI%=0
SS$=INKEY$:IF SS$<>"" THEN GOTO YNDUPFINAL:'(another line to allow
an overriding termination from the keyboard.)
NEXT R%
YNDUPFINAL:
```

'Open the disc to output the list of acceptable sequences and output them..

```
OPEN NM$ FOR OUTPUT AS #2
```

```
        PRINT#2, "The number of identicals observed and rejected in this run =
        ";ICOU%
        PRINT#2, "The number of sequences to print-"; POI%
        FOR FIN%=1 TO POI%
        PRINT#2,R$(FIN%)
        NEXT FIN%
        PRINT#2,"end"
        CLOSE#2
        PRINT "Your output is in the file you called- ";NM$
        PRINT "NOW TOUCH A LETTER ON THE KEY BOARD TO END UP THIS
        PROGRAM."
        PAUSEND:
        SS$=INKEY$:IF SS$="" THEN GOTO PAUSEND
        END
'*****************
```

'This small subroutine called RAN simply generates random numbers between 0 and 14 and delivers them to NU% for return to the main program.

```
        SUB RAN
        SHARED NU%
        LOCAL ULIM,LLIM
        LOCAL NU,DF,X
        ULIM=14:LLIM=0
        DF=ULIM-LLIM
        GETANUTHER:
        X=RND(5)
        NU=(X*DF)+LLIM:NU%=NU:IF NU%<LLIM OR NU%>ULIM THEN GOTO
        GETANUTHER
        END SUB
'******************************
```

Throughout this application various publications are referenced by author and year. Full citations for the publications are listed below. The disclosures of these publications in their entireties are hereby incorporated by reference into this

application in order to more fully describe the state of the art to which this invention pertains.

The invention has been described in an illustrative manner, and it is to be understood that the terminology that has been used is intended to be in the nature of words of description rather than of limitation.

Obviously, many modifications and variations of the present invention are possible in light of the above teachings. It is, therefore, to be understood that within the scope of the appended claims, the invention may be practiced otherwise than as specifically described.

CLAIMS

1. A method for amplifying genetic material by amplifying the genetic material using a single primer sequence.

2. The method according to claim1, further including amplifying the genetic material using short extension times.

3. The method according to claim 1, wherein said amplifying step includes amplifying with a polymerase reaction.

4. A method for detecting the presence of pathogens in a sample by:
amplifying genetic material for a pathogen in the sample using a single primer in an amplification process.

5. The method according to claim 4, wherein said amplifying step includes amplifying the genetic material of a pathogen via a polymerase reaction.

6. The method according to claim 5, wherein said amplifying step includes amplifying the pathogen genetic material with short extension times.

7. A detector for detecting pathogens in a sample, said detector comprising a single primer sequence for use in an amplification reaction, whereby said primer sequence amplifies pathogen genetic material thereby detecting pathogens in a sample.

8. The detector according to claim 7, wherein said primer sequence is between 16 and 30 bases long.

9. The detector according to claim 7, wherein said sample is selected from the group consisting of biological samples and agricultural samples.

10. A kit for performing the method of claim 1, said kit comprising:
a single primer sequence; and
a device for amplifying genetic material.

11. A method for amplifying unknown genetic material in a sample by:
adding a single primer sequence to the sample thereby preventing primer-dimers and primer concatenates;
amplifying the genetic material with multiple long extension times thereby enabling highly specific amplification of the unknown genetic material.

12. The method according to claim 11, wherein said amplifying step includes amplifying the genetic material for enough time to create at least 2kb of extension of the material.

13. The method according to claim 11, wherein said adding step includes adding a ligase independent, sequence non-specific primer to a smaller amount of the genetic material.

14. A method for amplifying unknown genetic material in a sample by:
adding a single primer sequence to the sample thereby preventing primer-dimers and primer concatenates; and
amplifying the genetic material with multiple short extension times thereby enabling highly specific amplification of the unknown genetic material.

15. The method according to claim 14, wherein said amplifying step includes amplifying the genetic material long enough for 200bp extensions.

16. The method according to claim 14, wherein said amplifying step includes:
annealing the primer for at least one cycle at very low specificity; and
repeatedly amplifying the genetic material with a higher specificity wherein there is no amplification from the original template, thereby making

selectivity dependent upon the properties of the sequence found between the amplified sections.

17. The method according to claim 14 for use in recording and cataloguing unidentified organisms.

18. A device for performing the method of claim 1, said device comprising:
a robot for performing the method; and
DNA separating and observing means functionally connected to said robot whereby said robot runs said DNA separating and observing means.

19. The device according to claim 18, wherein said DNA separating and observing means is selected from the group consisting essentially of a capillary electrophoresis machine, an HPLC device, Sanger sequencing fluorescent primer chain terminating technology, and mass spectroscopic technology.

20. A method for amplifying unknown genetic material in a sample by:
collecting the sample;
maintaining the sample in storage media;
purifying the sample;
reverse transcribing genetic material of the sample;
performing PCR on the genetic material; and
analyzing the results of the PCR.

21. A computer program for creating primers for use in the method of claim 1.

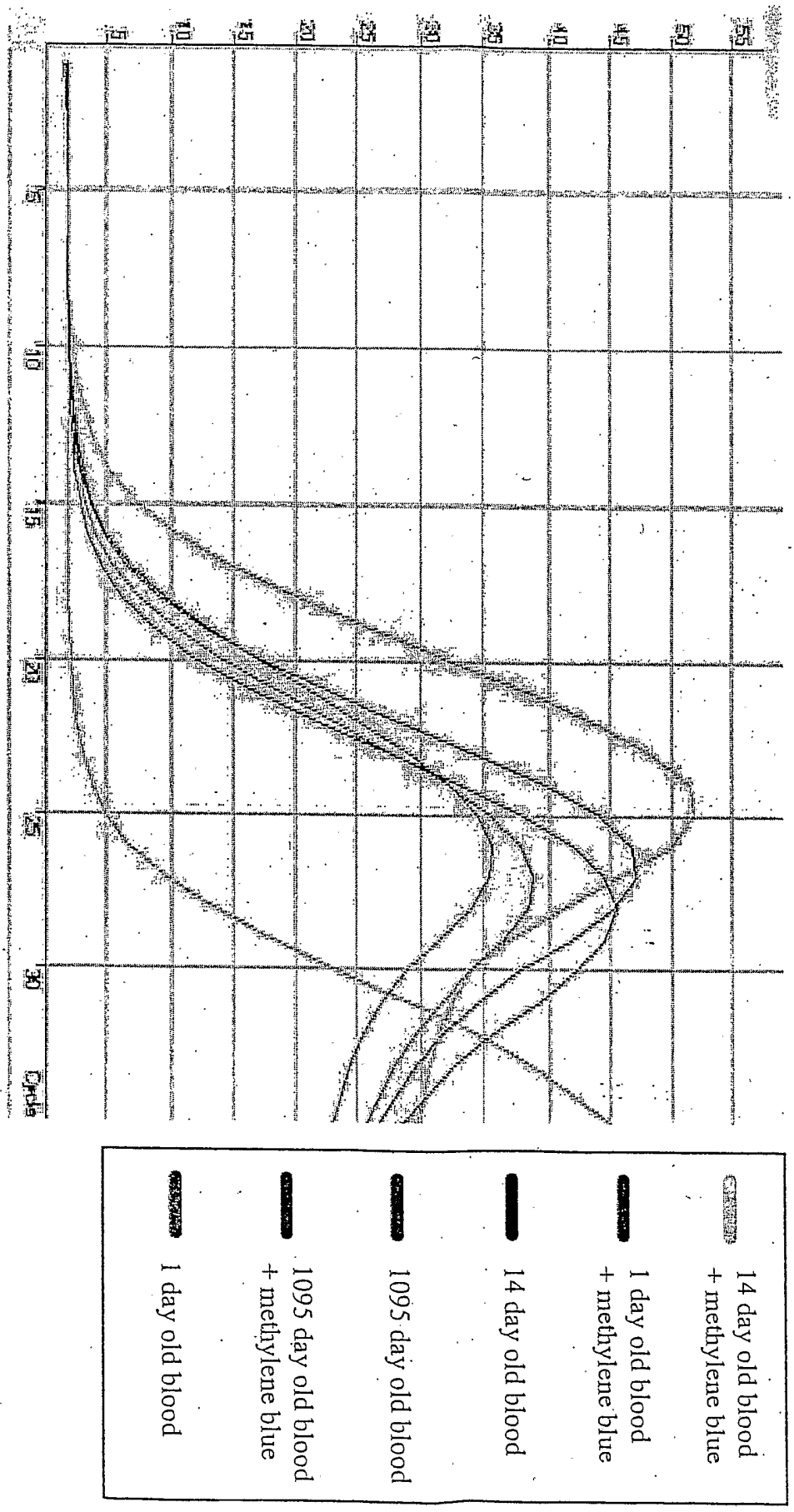

Relative Fluorescence
PCR Cycles
Cycle
14 day old blood + methylene blue
1 day old blood + methylene blue
14 day old blood
1095 day old blood
1095 day old blood + methylene blue
1 day old blood

FIG. 1

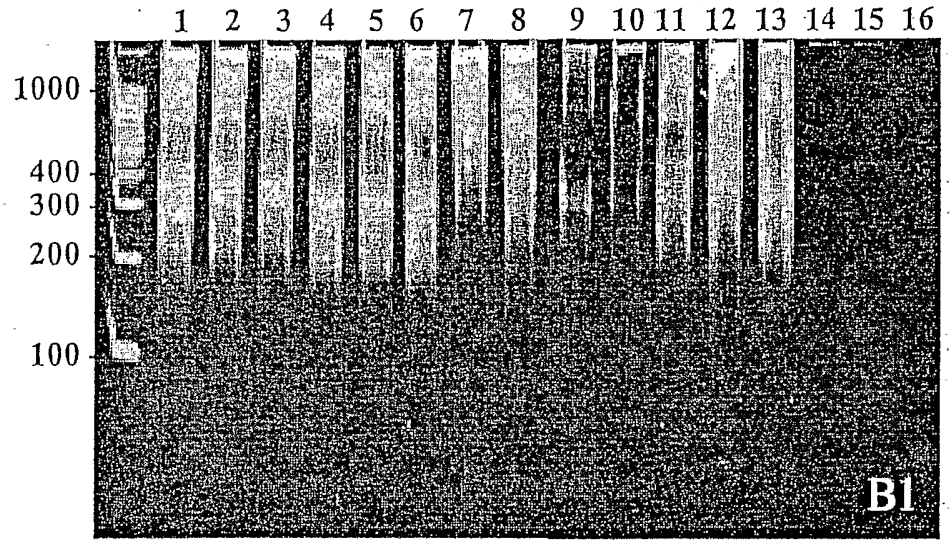


FIG. 2A

Second round amplification using 1ul of first round amplification as template

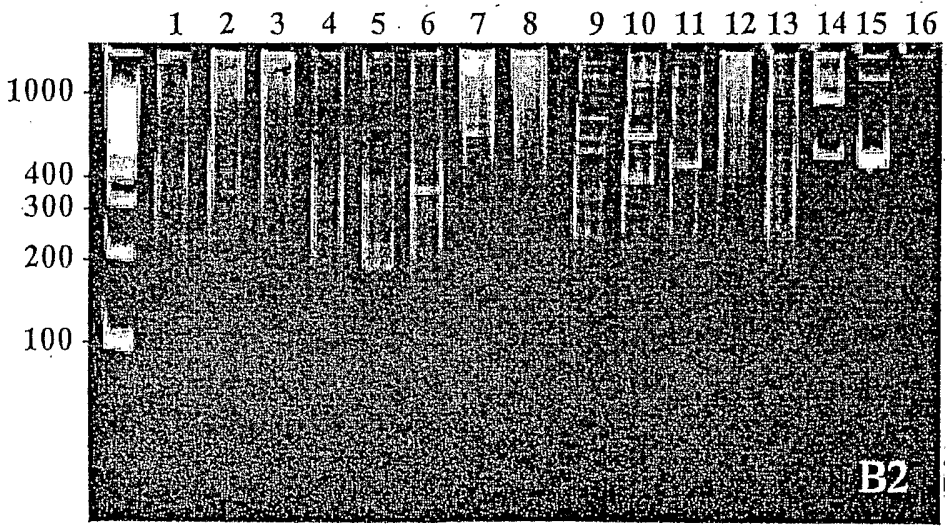


FIG. 2B

FIG. 2C

Primer	Sequence (5' - 3')
Seq5	CCCTCGAACACCACCTCC

Table B1 Seq5 primer sequence

FIG. 2D

Lane	Sample
1	Hepatitis C #1
2	Hepatitis C #2
3	Hepatitis C #3
4	HIV #1
5	HIV #2
6	HIV #3
7	CVB-4 #1
8	CVB-4 #2
9	uninfected plasma #1
10	uninfected plasma #2
11	uninfected plasma #3
12	uninfected serum #1
13	uninfected serum #2
14	FTA negative control
15	RT-PCR negative control
16	PCR negative control

Table B2 Contents for figures B1 & B2

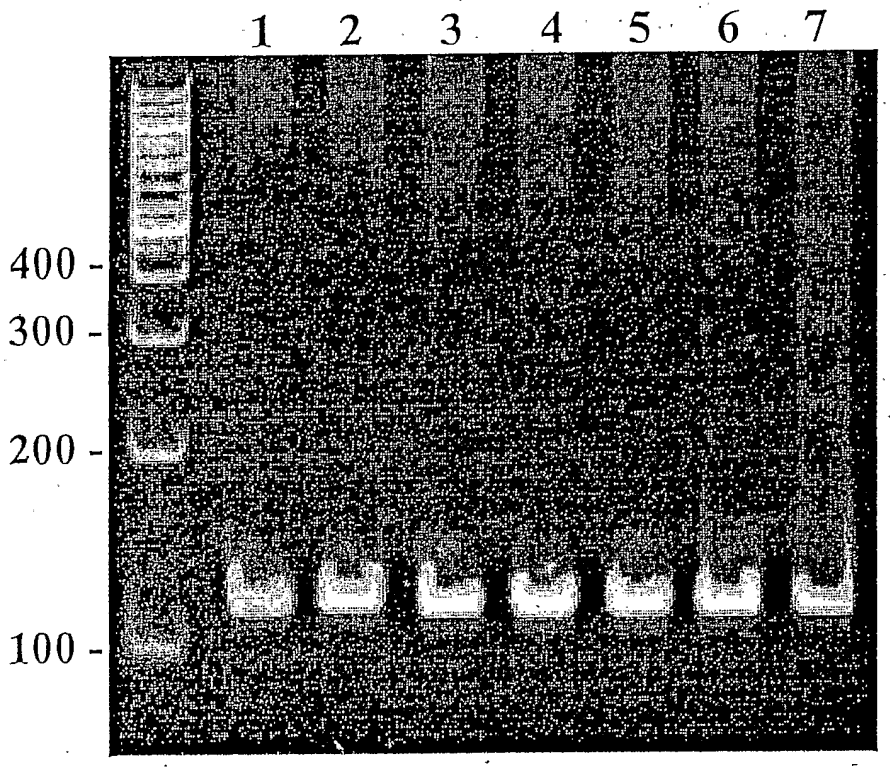


FIG. 3A

FIG. 3B

Lane	Sample
1	CVB-4 culture sample #1
2	CVB-4 culture sample #2
3	CVB-4 culture sample #3
4	CVB-4 culture sample #4
5	CVB-4 culture sample #5
6	CVB-4 culture sample #6
7	CVB-4 culture sample #7

Table C1 Contents of figure C

FIG. 3C

EP2	TCCTCCGGCCCCTGAATGCG
GS2	AAACACGGACACCCAAAGTA

Table C2 EP2 and GS2 primer sequences

FIG. 3D

94°C for 10 min	
94°C forr 10 sec	35 cycles
55°C for 30 sec	
72°C for 1 min	
72°C for 5 min	

Table C3 CVB-4 PCR amplification protocol

SUBSTITUTE SHEET (RULE 26)

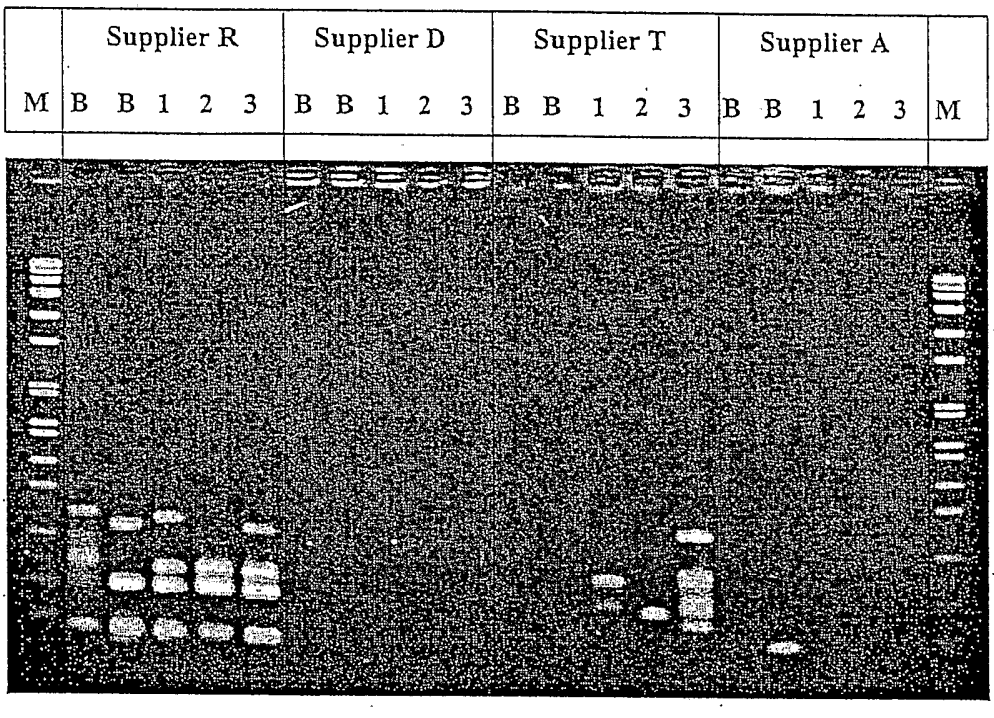

Supplier R
Supplier D
Supplier T
Supplier A
M B B 1 2 3 B B 1 2 3 B B 1 2 3 B B 1 2 3 M

FIG. 4

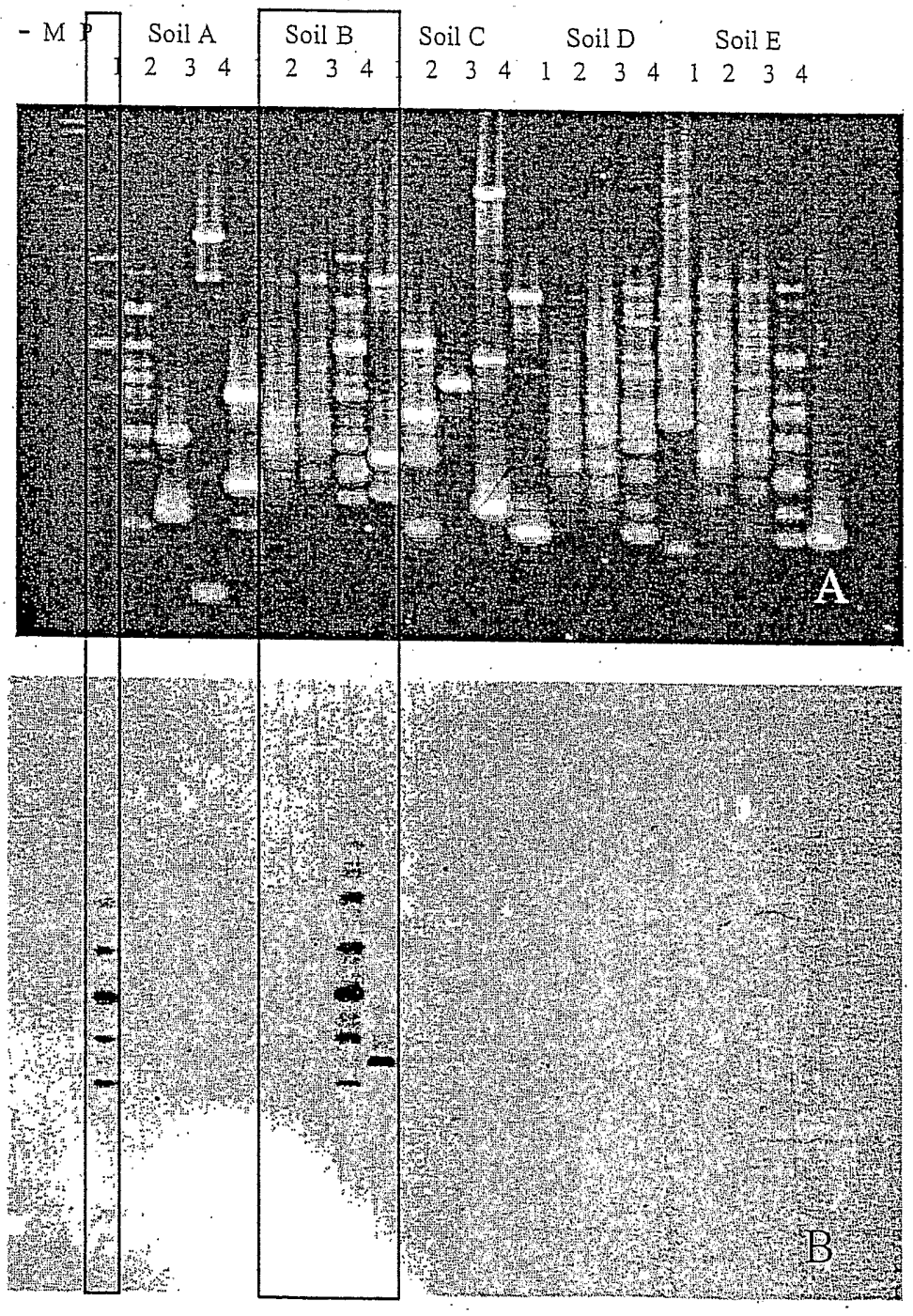


FIG. 5

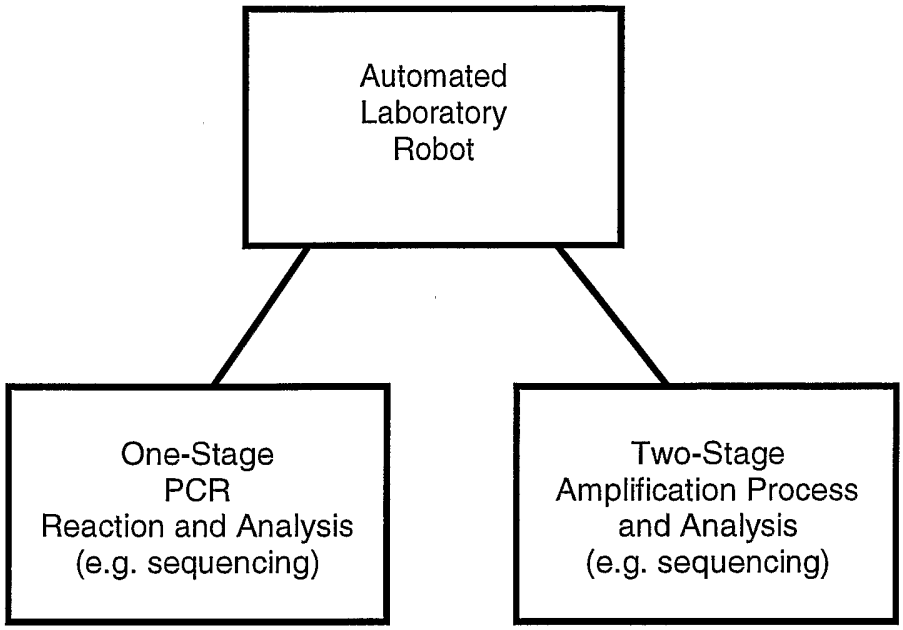

Automated
Laboratory
Robot
One-Stage
PCR
Reaction and Analysis
(e.g. sequencing)
Two-Stage
Amplification Process
and Analysis
(e.g. sequencing)

FIG. 6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/26670

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C12P 19/34; C12Q 1/68; CO7H 21/04
US CL : 435/91.1, 91.2, 91.21, 91.5, 91.51, 6; 536/23.1, 24.33, 23.7, 23.72, 23.74

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)
U.S. : 435/91.1, 91.2, 91.21, 91.5, 91.51, 6; 536/23.1, 24.33, 23.7, 23.72, 23.74

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 practicable, search terms used)
WEST 2.0, Medline, Caplus, Embase, Biosis, Scisearch, biotechds, Cancerlit,

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,683,879 A (LANEY et al) 04 NOVEMBER 1997, see entire patent, especially columns 4-12, 18, 19, 21-26.	1-6, 11-17
X	US 5,800,980 A (PERRON et al) 01 SEPTEMBER 1998, see column 23, lines 24-29.	1, 3-5
X	US 5,747,257 A (JENSEN) 05 MAY 1998, see entire patent, especially column 9, lines 30-57).	1, 3-5, 11, 12
X	US 5,972,386 A (BURGOYNE) 26 OCTOBER 1999, see entire patent.	20
X	US 5,976,572 A (BURGOYNE) 02 NOVEMBER 1999 see entire patent.	20
A	US 6,294,323 B1 (ULLMAN et al) 25 SEPTEMBER 2001, entire patent	1-6, 11-17

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent 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
26 November 2002 (26.11.2002)	27 JAN 2003
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703)305-3230	Authorized officer Gary Benzion Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/26670

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claim Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claim Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claim Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
Please See Continuation Sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-6, 11-17 and 20

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

Group 1, claim(s) 1-6,11-17, and 20, drawn to a method of amplifying genetic material.

Group 2, claim(s) 7-9, drawn to detector.

Group 3, claim(s) 10 , drawn to kit comprising primer and device.

Group 4, claim(s) 18,19, drawn to a device.

Group 5, claim(s) 21, drawn to a computer program.

The inventions listed as Groups 1-5 do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: Groups 1-5 lack unity because the "special technical feature", namely the use of a single primer in amplification, is not new and does not provide a contribution over the prior art. PCT Rule 13.2 requires that unity of invention only exist when a shared technical feature is a contribution over the art (and therefore constitutes a "special" technical feature). As methods for amplifying genetic material using a single primer sequence were known in the prior art (See Jensen, US 5,747,257, 05, May 1998, col. 9, lines 41-47 and Perron et al. US 5,800,980 01, Sept. 1998, col. 23, lines 24-29), the technical feature does not constitute a "special" technical feature and unity of he invention is lacking.

Additionally, the different inventions have different modes of operations and different effects.