(54) Title: SIMPLIFIED DE NOVO PHYSICAL MAP GENERATION FROM CLONE LIBRARIES

(57) Abstract: The current invention relates to a method for the ordering of sequence tags from at least part of genome. The method at least comprises the steps of providing a genomic clone library, generating clone aliquots from the library, generating sequence tags from the clone aliquots, and ordering the sequence tags based on the combined presence of sequence tags in the clone aliquots. In one aspect, there is provided a method for the detection of genomic variation between at least two samples comprising the comparison of the ordered sequence tags of the samples and identifying any variation between the samples.
Title: Simplified De novo physical map generation from clone libraries

5 Technical Field

[01] The present invention relates to the field of molecular biology and biotechnology. In particular, the invention relates to the field of nucleic acid detection and identification. The invention relates more in particular to a method for the ordering of sequence tags of a genome. The invention further relates to the generation of a de novo physical map of a genome, or part of a genome based on the ordered sequence tags. The invention further relates to the use of the method to identify structural variants between the genomes of multiple samples based on the difference in ordered sequence tags.

Background Art

[02] Integrated genetic and physical genome maps are extremely valuable for map-based gene isolation, comparative genome analysis and as sources of sequence-ready clones for genome sequencing projects. The effect of the availability of an integrated map of physical and genetic markers of a species for genome research is enormous. Integrated maps allow for precise and rapid gene mapping and precise mapping of microsatellite loci and SNP markers. Various methods have been developed for assembling physical maps of genomes of varying complexity. One of the better characterized approaches use restriction enzymes to generate large numbers of DNA fragments from genomic subclones (Brenner et al., Proc. Natl. Acad. Sci., 1989, 86, 8902-8906; Gregory et al., Genome Res. 1997, 7, 1162-1168; Marra et al., Genome Res. 1997, 7, 1072-1084). These fingerprints are compared to identify related clones and to assemble overlapping clones in contigs. However, the utility of fingerprinting for ordering large insert clones of a complex genome is limited, due to variation in DNA migration from gel to gel, the presence of repetitive sequences in the genome, unusual distribution of restriction sites and skewed clone representation. Most high quality physical maps of complex genomes have therefore been constructed using a combination of fingerprinting and PCR-based or hybridisation-based methods. However, one of the disadvantages of the use of fingerprinting technology is that it is based on fragment-pattern matching, which is an indirect method and error-prone in the sense that fragments with a similar mobility do not necessarily originate from a single site in the genome. For example, Nelson and Soderlund (Nucleic Acids Research 2009, e36) report that up to 11% of fingerprint bands in the High Information Content Fingerprinting (HICF) map of maize co-migrate by chance. Thus, fingerprint-based methods have a limited resolution to resolve
repeat regions and it is therefore preferred to have a method for ordering large insert clones which is directly based on the clone (i.e. genome) sequence itself.

[03] Recently methods have been described that create physical maps by generating the contigs based on actual sequence data, i.e. a more direct method. A sequence-based physical map is not only more accurate, but at the same time also contributes to the determination of the whole genome sequence of the species of interest. Using state-of-the-art high-throughput sequencing allows for the determination of complete nucleotide sequences of clones in a more efficient and cost-effective manner. One of the methods for sequencing based generation of physical map is described in WO2008/007951 in the name of Keygene NV. In WO2008/007951, a physical map is generated from a clone library. The clones are pooled and adapter-ligated restriction fragments are generated. The fragments are partly sequenced and correlated to the original clone via an identifier. The sequenced fragments are also indicated as ‘tags’. The co-presence of these tags in the various clones is used as a basis to place the clones in clone contigs, resulting in a sequence-based physical map.


[05] Current methodologies for the generation of physical maps rely on clone libraries. The preparation of clone libraries is an arduous task, involving subdividing genomic DNA into clonable elements, ligating them in a vector and inserting vector-insert constructs into hosts, typically Escherichia Coli (E. coli)). The BAC clones are plated out, colonies are individually picked and each clone is stored separately. For the generation of sequence-based physical maps, for example using the process described in WO2008/007951, the clones are pooled, typically in complex pooling systems such as 2D, 3D or 6D. Multi-dimensional pooling requires deconvolution to assign the sequence reads to the individual BAC clones. Both the generation of clone libraries as well as subsequent pooling, BAC DNA isolation, sequence template preparation, sequencing and deconvolution are laborious and time-consuming (thus costly) procedures. In case the BAC clones would not be pooled, deconvolution would not be needed but the costs for DNA isolation and sequence template preparation would be even higher. Thus, in practice a pooling-based approach to construct a sequence-based physical
map is currently the preferred method, even though the process of pooling is laborious and
time-consuming.

Summary of Invention

Technical problem
The technical problem identified in the art is that elaborate schemes are necessary for the
preparation of clone libraries, involving plating out clones, for colony picking of separate
clones and storing and maintaining them in microtiter plate format. Also the subsequent
pooling steps, DNA isolation, sequence-template preparation steps and deconvolution of the
prior art are costly and difficult to perform. The problem to be solved is to invent a robust
method for construction of a high-resolution sequence-based ordering of sequence tags and
the subsequent formation of a physical map which does not suffer from these limitations.

The solution to the problem

The present inventors have found that a new type of sequence-based ordering of
sequence tags and the subsequent formation of a physical map (compared to
WO2008/007951) can be constructed when the steps of plating out clones for individual
colony picking and multi-dimensional pooling of individual clones are omitted, and the
physical map is based directly on ordering the sequence tags, rather than ordering the clones
from which the sequence tags are derived (as in WO2008/007951). Thus, the solution
provided to the problem is to that co-retention frequencies of the sequence tags are
measured in clone aliquots of the clone library (i.e. in multiple clones simultaneously), without
making attempts to assign these tags to the individual clones. In the invention, the clones
merely serve to contribute a number of closely linked tags to a complex mixture of library
clones which are sequenced en masse and to provide a stable, lasting resource that can be
used for other purposes that are beyond the scope of the current invention.

Essentially, the solution provided combines the power and advantages of BAC
library technology, sequence tag preparation as described in WO2008/007951, and the
principle of tag ordering based on co-retention frequencies as described a.o. by Dear and
Cook (1989, and WO2103046, WO2103047 and WO2103048), but adapted to NGS
sequences and large insert clones, into a novel method for sequence-based ordering of
sequence tags and the subsequent construction of a physical map.

The present inventors realised that the short sequence stretches that can be
generated by the current Next-Generation Sequencing (NGS) technologies contain sufficient
information to order the sequenced tags and to link clones together to generate a physical
map, as also described in WO2008/007951.
The present invention is based on the insight that co-retention of sequence tags in samples of clones provides adequate information to order the sequences tags and to build a sequence-based physical map, analogous to building physical maps based on known STS markers as has been described by Dear and Cook (1989), but omitting the use of previously known markers.

**Brief Description of Drawings**

**Fig. 1**

Principle of physical mapping based on co-presence analysis of a sequence tags in different aliquots of clones. A BAC library is provided, clone aliquots are established, and sequence data are generated from fragments obtained from clone aliquots. Co-presence analysis of sequence tags is performed to determine their physical order in the genome.

**Fig. 2**

Different configurations for annealing of primers to adaptor-ligated fragments:
1) no amplification takes place, hence no annealing of primer (PPPPPP) to adaptor (AAAAAAA).
2a) the amplification primer anneals to the 5' boundary of the identifier sequence (IDIDIDID) in the adaptor.
2b) the amplification primer anneals to the 3' boundary of the identifier sequence (IDIDIDID) in the adaptor.
3) the amplification primer contains an identifier sequence (IDIDID) anneals to a degenerate position in the adaptor, while the remaining part of the primer anneals to the 5' end of the adaptor.
4) the amplification primer contains an identifier sequence (IDIDID) at the 5' end that does not match the adaptor sequence.

**Definitions**

In the following description and examples, a number of terms are used. In order to provide a clear and consistent understanding of the specification and claims, including the scope to be given such terms, the following definitions are provided. Unless otherwise defined herein, all technical and scientific terms used have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. The disclosures of all publications, patent applications, patents and other references are incorporated herein in their entirety by reference.

As used herein, the singular forms "a," "an" and "the" include plural referents unless the context clearly dictates otherwise. For example, a method for isolating "a" DNA molecule,
as used above, includes isolating a plurality of molecules (e.g. 10's, 100's, 1000's, 10's of thousands, 100's of thousands, millions, or more molecules).

Genomic library: A genomic library is a population of host bacteria, each of which carries a DNA molecule that was inserted into a cloning vector, such that the collection of cloned DNA molecules represents the entire genome of the source organism. This term also represents the collection of all of the vector molecules, each carrying a piece of the chromosomal DNA of the organism, prior to the insertion of these molecules into the host cells. The process of subdividing genomic DNA into clonable elements, ligating them into suitable vectors and inserting the vector-insert constructs into hosts is called creating a library, a clone bank or a gene bank. A complete library of host cells will contain all the genomic DNA of the source organism. Genomic libraries come in sizes: Plasmids (~15 kb), Phage (lambda)s (~25 kb). Cosmid (fossid)s (~35-45 kb). Bacterial artificial chromosomes (BAC, P-1 derived, ~50-300 kb). Yeast artificial chromosomes (YAC, ~300- >1500 kb). Human artificial chromosomes (HAC, ~>2000 kb). The probability of reaching this goal is reflected by the number of genome equivalents that the library represents. A higher number of genome equivalents contained in the library increases the chances that its coverage nears completion.

A genomic library is created by isolating the DNA molecules of an organism of interest. Typically, the DNA molecules are then partially digested by an endonuclease restriction enzyme. Sometimes, the DNA molecules are digested for different lengths of time or using combinations of restriction enzymes in order to ensure that all the DNA has been digested to manageable sizes. Alternatively, the DNA molecules of an organism are randomly sheared into the desired size range that is compatible with the cloning vector/host cell combination. Theoretically, a random sheared library has a higher likelihood of covering the organism's genome in its entirety. The digested or randomly sheared DNA molecules are separated by size, for instance using agarose electrophoresis or pulsed-field gel electrophoresis (PFGE), and a suitable range of lengths of DNA pieces are isolated and ligated into vectors. The vectors can then be taken up by suitable hosts.

Clone bank: As used herein, a clone bank (or genomic clone library) is a genomic library wherein all the clones have been separately isolated (via plating out and colony picking). Each entry of a clone bank contains one clone.

Restriction endonuclease: a restriction endonuclease or restriction enzyme is an enzyme that recognizes a specific nucleotide sequence (target site) in a double-stranded DNA molecule, and will cleave both strands of the DNA molecule at or near every target site, leaving a blunt or a staggered end.

Frequent cutters and rare cutters: Restriction enzymes typically have recognition sequences that vary in number of nucleotides from 3, 4 (such as Msel) to 6 (EcoRI) and even
The restriction enzymes used can be frequent or rare cutters. The term 'frequent' in this respect is typically used in relation to the term 'rare'. Frequent cutting endonucleases (aka frequent cutters) are restriction endonucleases that have a relatively short recognition sequence. Frequent cutters typically have 3-5 nucleotides that they recognise and subsequently cut. Thus, a frequent cutter on average cuts a DNA sequence every 64-512 nucleotides. Rare cutters are restriction endonucleases that have a relatively long recognition sequence. Rare cutters typically have 6 or more nucleotides that they recognise and subsequently cut. Thus, a rare 6-cutter on average cuts a DNA sequence every 1024 nucleotides, leading to longer fragments. It is observed again that the definition of frequent and rare is relative to each other, meaning that when a 4 bp restriction enzyme, such as Msel, is used in combination with a 5-cutter such as Avall, Avall is seen as the rare cutter and Msel as the frequent cutter.

Adaptor: short double-stranded DNA molecule with a limited number of base pairs, e.g. about 10 to about 30 base pairs in length, which are designed such that they can be ligated to the ends of restriction fragments. Adaptors are generally composed of two synthetic oligonucleotides which have nucleotide sequences which are partially complementary to each other. When mixing the two synthetic oligonucleotides in solution under appropriate conditions, they will anneal to each other forming a double-stranded structure. After annealing, one end of the adaptor molecule is designed such that it is compatible with the end of a restriction fragment and can be ligated thereto; the other end of the adaptor can be designed so that it cannot be ligated, but this need not be the case (double ligated adaptors).

Identifier: a short sequence that can be added to an adaptor or a primer or included in its sequence or otherwise used as label to provide a unique identifier. Such a sequence identifier (tag) can be a unique base sequence of varying but defined length, typically from 4-16 bp used for identifying a specific nucleic acid sample. For instance 4 bp tags allow 4\(\times 256\) = 256 different tags. Using such an identifier, the origin of a PCR sample can be determined upon further processing. In the case of combining processed products originating from different nucleic acid samples, the different nucleic acid samples are generally identified using different identifiers. Identifiers preferably differ from each other by at least two base pairs and preferably do not contain two identical consecutive bases to prevent misreads. The identifier function can sometimes be combined with other functionalities such as adapters or primers.

The term "contig" is used in connection with DNA sequence analysis, and refers to reassembled contiguous stretches of DNA derived from two or more DNA fragments having contiguous nucleotide sequences. Thus, a contig is a set of overlapping DNA fragments that provides a partial contiguous sequence of a genome. A "scaffold" is defined as a series of
contigs that are in the correct order, but are not connected in one continuous length, i.e. contain gaps. Contig maps also represent the structure of contiguous regions of a genome by specifying overlap relationships among a set of clones. For example, the term "contigs" encompasses a series of cloning vectors which are ordered in such a way as to have each sequence overlap that of its neighbours. The linked clones can then be grouped into contigs, either manually or, preferably, using appropriate computer programs such as FPC, PHRAP, CAP3 etc.

[21] Primer: in general, the term primer refers to DNA strands which can prime the synthesis of DNA. DNA polymerase cannot synthesize DNA de novo without primers: it can only extend an existing DNA strand in a reaction in which the complementary strand is used as a template to direct the order of nucleotides to be assembled. We will refer to the synthetic oligonucleotide molecules which are used in a polymerase chain reaction (PCR) as primers.

[22] DNA amplification: the term DNA amplification or, more general, amplification, will be typically used to denote the in vitro synthesis of double-stranded DNA molecules using PCR. It is noted that other amplification methods exist and they may be used in the present invention without departing from the gist.

[23] Amplicon: The product of a polynucleotide amplification reaction, namely, a population of polynucleotides that are replicated from one or more starting sequences. Amplicons may be produced by a variety of amplification reactions, including but not limited to polymerase chain reactions (PCRs), linear polymerase reactions, nucleic acid sequence-based amplification, rolling circle amplification and like reactions.

[24] High-throughput screening: High-throughput screening, often abbreviated as HTS, is a method for scientific experimentation especially relevant to the fields of biology and chemistry. Through a combination of modern robotics and other specialised laboratory hardware, it allows a researcher to effectively screen large amounts of samples simultaneously.

[25] Nucleic acid: a nucleic acid according to the present invention may include any polymer or oligomer of pyrimidine and purine bases, preferably cytosine, thymine, and uracil, and adenine and guanine, respectively (See Albert L. Lehninger, Principles of Biochemistry, at 793-800 (Worth Pub. 1982) which is herein incorporated by reference in its entirety for all purposes). The present invention contemplates any deoxyribonucleotide, ribonucleotide or peptide nucleic acid component, and any chemical variants thereof, such as methylated, hydroxymethylated or glycosylated forms of these bases, and the like. The polymers or oligomers may be heterogenous or homogenous in composition, and may be isolated from naturally occurring sources or may be artificially or synthetically produced. In addition, the nucleic acids may be DNA or RNA, or a mixture thereof, and may exist permanently or
transitionally in single-stranded or double-stranded form, including homoduplex, heteroduplex, and hybrid states.

[26] Ligation: the enzymatic reaction catalyzed by a ligase enzyme in which two double-stranded DNA molecules are covalently joined together is referred to as ligation. In general, both DNA strands are covalently joined together, but it is also possible to prevent the ligation of one of the two strands through chemical or enzymatic modification of one of the ends of the strands. In that case the covalent joining will occur in only one of the two DNA strands.

[27] Restriction fragments: the DNA molecules produced by digestion with a restriction endonuclease are referred to as restriction fragments. Any given genome (or nucleic acid, regardless of its origin) will be digested by a particular restriction endonuclease into a discrete set of restriction fragments. The DNA fragments that result from restriction endonuclease cleavage can be further used in a variety of techniques and can for instance be detected by gel electrophoresis.

[28] Sequencing: The term sequencing refers to determining the order of nucleotides (base sequences) in a nucleic acid sample, e.g. DNA or RNA. Many techniques are available such as Sanger sequencing and high-throughput sequencing technologies (also known as next-generation sequencing technologies) such as the GS FLX platform offered by Roche Applied Science, and the Genome Analyzer from Illumina, both based on pyrosequencing.

Description of Embodiments

[29] Thus, in one aspect, the invention pertains to a method for the ordering of sequence tags from at least part of genome, comprising the steps of providing a genomic clone library, generating clone aliquots from the library, generating sequences tags from the clone aliquots, ordering the sequence tags based on the combined presence of sequence tags in the clone aliquots. The invention is based on the insight that sequence tags can be ordered based on the relative presence with which they occur in samples such as clones. Parts of the sequence that are located close to each other will, in subsequent sequencing steps, have a tendency to occur more frequently together than sequence parts that are located further apart. This principle allows for the relative ordering of the sequence information and hence allows for the building of ordered set of sequence information. Such information can be used for a variety of purposes, including the building of a physical map.

[30] Thus, in one aspect the invention further pertains to the generation of a physical map of at least part of a genome comprising the steps of:

a) providing a genomic library containing at least part of the genome in a plurality of clones;
b) generating clone aliquots of the genomic library, whereby each clone aliquot contains more than one clone;
c) isolating DNA from the clone aliquots;
d) sequencing at least part of the DNA of the clone aliquot to provide sequence
tags;
e) determining the combined presence of two or more sequence tags in the clone
aliquots;
f) ordering the sequence tags based on their combined presence in the clone
aliquots, thereby generating the physical map.

Thus, in a first step a genomic library is provided. The genomic library contains at
least part of a genome of an organism of interest. The genome of interest is provided in a
plurality of clones. The number of clones that constitute a genomic library depends on (1) the
size of the genome in question and (2) the insert size tolerated by the particular cloning vector
system. For most practical purposes, the tissue source of the genomic DNA is unimportant
because each cell of the body contains virtually identical DNA. The clones can be prepared
using the conventional procedures known in the art for subdividing genomic DNA into
clonable elements, inserting them into vectors and transforming host cells. Suitable
techniques are for instance described in Singer et al., in 1997 [Nucleic Acids Research, 25(4)
781-786 (1997)] using restriction enzymes, but also methods using random fragmentation are
known in the art can be used. Typically, a genomic library contains a number of clones with
inserts that together are equivalent to several times the genome size of the investigated
genome, indicated as the number of genome equivalents. In the present invention at least
one genome equivalent is present in the genomic library, but more preferably more than three
genome equivalents are present. Even more preferably the library contains at least five
genome equivalents, more preferably at least 7, most preferably at least 8. Particularly
preferred are at least 10 genome equivalents. More genome equivalents improve the
accuracy and resolution of the physical map. Thus, the higher the number of genome
 equivalents in the library, the more reliable the physical map will be. A typical genomic library
contains many thousands of clones. For instance, the total Arabidopsis thaliana genome is 130 Mbp. A Bacterial Artificial Chromosome (BAC) has a genomic insert of -130 kbp on
average. One genome equivalent of BACs of the Arabidopsis genome thus comprises
approximately 1,000 BACs. Similarly, cucumber (Cucumis sativus) has a genome of about
360 Mbp. In that case, one genome equivalent BACs corresponds to approximately 2,800
BACs with an average insert size of 130 kbp. Similarly, the human genome is approximately
3,000 Mbp in size. One genome equivalent human BACs thus corresponds to 23,000 BACs
of 130 kbp. A 12X human BAC library thus contains about 12 times 23,000 = 276,000 BAC
clones. Hence, the number of BAC clones needed for high-resolution physical mapping can
be considerable and directly scales with genome size.
The genomic library can be subjected to the usual quality checks such as determining the average insert size, the fraction of empty clones (which should be as low as possible, preferably below 2%). The concentration/titer of the clones in the genomic library can be determined to allow steering of the complexity of the aliquots. The genomic library is divided into clone aliquots. The clone aliquots are, in fact, parts of the clone library as the ligated mixture of DNA fragments is transferred en masse into the E. coli. bacteria. Thus the clones in the library are divided in a set of aliquots that each contain a certain number of clones. Preferably, the aliquots together contain all clones of the library, but in certain embodiments it may be sufficient to have a portion of the clones of the library divided into aliquots such as at least 75%, at least 80% at least 85% and at least 90% of the clones are divided into aliquots. In case an already existing clone library is used, the clones are already plated out and separated and can be already pooled in clone pools. For the purpose of this invention, clone pools can be used, but in this embodiment the pool information (which clone is present in which pool(s)) is disregarded. Hence, in this invention the term clone aliquots is used to indicate a group of clones together, regardless of their origin (i.e. obtained via aliquots or pooling).

The clone aliquots contain at least one clone per aliquot. The (average) number of clones per aliquot may vary, depending on the size of the genome of the organism and on the average insert size. Typically, each aliquot contains around 0.5 to 0.9 genome equivalents of clones, but the method can also be performed (although less cost-efficiently), when each aliquot contains up to 1 genome equivalent clones. In practice, for most organisms this amounts to between 100 and 1000 clones per aliquot, but this depends on the on the organism’s genome size and the average clone insert size as mentioned above. The aliquots can be prepared directly from the genomic library, i.e. without the preparation of a separate clone library which comprises the steps of plating out the clones and individually picking the colonies into microtiter plates etc. In this embodiment, the prepared genomic library is distributed (aliquoted) into a number of aliquots. In an alternative embodiment, for instance using an already available clone bank stored in microtiterplates containing individual clones, the aliquots can be prepared by randomly pooling the clones in the clone bank. Either way, a aliquot contains a number of clones. In certain embodiments, there is no knowledge about which clone is present in which pool, i.e. it is irrelevant for the concept of the present invention to have or use knowledge about which clone is present in which aliquot. For the example of the human BAC library, about 800 aliquots of each 200-500 clones per aliquot would be sufficient.

The amount of DNA present in the aliquot is less than one genome equivalent, i.e. a typical range for a genome equivalent in an aliquot is in the order of 0.4 - 1.0 GE, but the most preferred amount of GE per aliquot is 0.7.
The isolation of clone DNA is generally achieved using common methods in the art such as for instance using the Q-Biogene fast DNA kit. For sequence-based technologies, it is important that DNA preps contain minimal amounts of E. coli (host cell) DNA. DNA quantification and optionally normalization may be performed to obtain equal amounts of DNA per aliquot.

The sequence of at least part of the DNA in the aliquots is determined to provide the sequence tags. To this end, the isolated DNA can be fragmented. Fragmentation of the DNA can be achieved in various ways, for instance in a more controlled way, using one or more restriction endonucleases. In principle any restriction endonuclease can be used such as blunt cutters (provide blunt ends) or staggered cutters (provide staggered ends). Restriction endonucleases may be frequent cutters (4 or 5 cutters, such as MseI or PstI) or rare cutters (6 and more cutters such as EcoRI, HindIII). Typically, restriction endonucleases are selected such that restriction fragments are obtained that are, on average, present in an amount or have a certain length distribution that is adequate for the subsequent steps. In certain embodiments, two or more restriction endonucleases can be used and in certain embodiments, combinations of rare and frequent cutters can be used. For large inserts the use of, for instance, three or more restriction endonucleases can be used advantageously. In certain embodiments, restriction endonucleases can be used that are non-palindromic. Fragmentation using at least one restriction endonuclease is preferred.

Fragmentation can also be achieved by physical techniques, i.e. techniques of a more random nature such as radiation, shearing, sonification or other random fragmentation methods. The amount or time of shearing or sonification then determines the length of the resulting fragments.

In certain embodiments, there is a preference for the use of at least two restriction endonucleases or a combination of at least one restriction endonuclease with a random fragmentation step. In this way, the size of the (restriction) fragments can be fine-tuned to the requirements of the NGS platform and/or the density of sequence tags across the genome can be controlled. It is generally not recommended to select a restriction enzyme that cuts within a very high copy number repetitive genome sequence. Furthermore, the use of a restriction enzyme in combination with random fragmentation can be used to determine the sequence of the resulting restriction fragments by paired-end sequencing.

The thus obtained fragments may be ligated to adapters. The ligation to adapters may be useful for further sequencing purposes on the various sequencing platforms. The use of adapters that are ligated to the fragments is that the adapters serve to initiate amplification of the fragments and to introduce the sequencing primers that are used in the sequencing steps. Furthermore the adapters may serve to introduce identifiers/barcodes.
To one or both ends of the restriction fragments, adapters can be ligated to provide for adapter-ligated restriction fragments. A fragment may contain the same or different adapters at each end. Typically, adapters are synthetic oligonucleotides as defined herein elsewhere. The adapters used in the present invention may contain an identifier section, in essence as defined herein elsewhere to provide for ‘tagged adapters’. In certain embodiments, the adapter contains an identifier. In certain embodiments, such an identifier may be aliquot-specific, i.e. for each aliquot, an adapter containing a unique identifier is used that unequivocally indicates the aliquot. In certain embodiments, the adapter contains a degenerate identifier section which is used in combination with a primer containing a aliquots-specific identifier. The adapter-ligated fragments can optionally be amplified using a set of primers of which at least one primer amplifies the aliquot-specific identifier at the position of the aliquot-specific or degenerate identifier in the adapter. The primer may contain (part of) the identifier, but the primer may also be complementary to a section of the adapter that is located outside the identifier, i.e. downstream in the adapter. Amplification then also amplifies the tag. See in this respect Fig 2 for various embodiments. However, as mentioned above, the amplifications step is optional, as nowadays, certain NGS platforms (e.g. the Heliscope™ produced by Helicos) are capable of performing single-molecule sequencing and therefore do not require amplification of the target molecules prior to sequencing. Hence, in such cases, an amplification step is not included because it is not required for subsequent sequencing.

In certain embodiments, the adapter-ligated fragments can be combined in larger groups, in particular when the adapters contain a aliquot-specific identifier. This combination in larger groups may aid in reducing the number of parallel amplifications of each set of adapter-ligated restriction obtained from a aliquot.

The (identifier-containing) adapter-ligated fragment can be amplified. The amplification may serve to reduce the complexity or to increase the amount the DNA available for analysis/sequencing (see above). The amplification can be performed using a set of primers that are at least partly complementary to the adapters and/or the tags/identifiers. This amplification may be independent from the amplification described herein above that introduces the unique identifiers by amplification with primers that match a degenerate identifier sequences in the adapters. In certain embodiments, the amplification may serve several purposes at a time, i.e. reduce complexity, increase DNA amount and introduce identifiers in the adapter-ligated fragments in the pools.

Part of the sequence of the (adapter-ligated) fragment is determined to provide for sequence tags. The (adapter-ligated) fragments are subjected to sequencing, preferably high throughput sequencing using an NGS platform as described herein elsewhere. During sequencing, at least part of the nucleotide sequence of the (amplified) tagged adapter-ligated fragment is determined. Preferably at least part of the sequence of the fragment (i.e.
derived from the sample genome) of the (amplified tagged adapter-ligated) fragment is determined. In certain embodiments also the sequence of the tag/identifier is determined. Preferably, a sequence of at least 10 nucleotides of the fragment is determined. In certain embodiments, at least 15, 20, 25, 30 or 35 nucleotides of the fragment (i.e. derived from the sample genome) are determined. The minimum number of nucleotides that will be determined is, again, genome- as well as sequencing platform dependent. For instance, in silico calculations on the known genome sequence of Arabidopsis have shown that, when including a 6 bp restriction site in the sequencing step, about 20 bp per fragment needs to be determined in order to ensure that the majority (> 75%) of sequences are unique in the genome. It is possible to determine the sequence of the entire fragment, but this is not an absolute necessity to obtain sufficient sequence information for the ordering of the sequence tags or the generation of a physical mapping.

[44] In the sequencing step, to provide for maximum coverage of all fragments and increased accuracy, the fragments may be sequenced with an average redundancy level (aka oversampling rate) of at least 5. This means that, on average, the sequence of a specific (optionally amplified) adapter-ligated fragment is determined at least five times. In other words: each fragment is (statistically) sequenced on average at least five times. Increased redundancy is preferred as it improves the fraction of fragments that are sampled in each pool (i.e. reducing sampling variation; see further below) and increases the accuracy of these sequences. So preferably the redundancy level is at least 7, more preferably a least 10. Increased average sequencing redundancy levels are used to compensate for a phenomenon that is known as 'sampling variation', i.e. random statistical fluctuation in sampling subsets from a large "population". In addition, a higher average sequencing redundancy level alleviates possible differences in the abundance of amplified fragments which result from differences in their amplification rates caused by length variation between fragments and differences in sequence composition.


[46] The sequencing results in sequence tags, unique stretches of sequence that are derived from the fragments of the genome under investigation. These stretches of sequence are distributed over the genome. The way in which the sequence tags are distributed over the fragments provides information on their relative position. When two sequence tags are linked, for instance located <125 kb apart, they are likely to appear be part of the same clone and
this appear in the same aliquot. Consequently, these two tags will show a certain percentage (for instance about 97-100%) co-presence or combined presence in the same aliquot. On the other hand, if they are located further apart, in the above example >125 kb they are less likely to appear in the same aliquot and show a lower percentage co-presence (for instance about 3%). Thus, the closer the physical distance between two tags in the genome, the higher their % of co-presence will be. This principle is used in determining the order of the sequence tags and hence of the clones.

[47] Thus, in the method according to the invention, the presence or absence of the sequence tags is determined in the aliquots. The presence of the sequence tags in an aliquot can be determined directly, when the aliquots are sequenced separately, or via the identifier that links the sequence tag to the aliquots from which it originates.

[48] The relative co-presence (co-presence %) of the sequence tags in the aliquots is determined. By ordering the sequence tags based on their relative co-presence, the order of the sequence tags in the aliquot can be determined. By ordering the sequences tags in a aliquot based on their relative co-presence, an ordering of sequence tags is generated for the DNA of the genome that was cloned. This ordering can provide the basis of a sequence-based physical map of the genome, without knowing exactly which clone harbours particular tags. However, since it is known which aliquot (of clones) contain particular tags, the individual clone containing particular tags may be identified by further analysis of the aliquoted ed clones via individual colony picking. This is beyond the scope of this invention and also not required for the construction of the physical map according to the method of the invention.

[49] Ordering of the sequence tags on the genome can be based on conventional ordering algorithms using the co-presence test, such as those used for HAPPY mapping and radiation hybrid (RH) mapping (Walter et al., 1994, Nature Genetics vol 7, pp 22-28). In certain embodiments, the genetic map of the genome can be used as a reference or to integrate the physical map with the genetic map. In certain embodiments one can use other markers to further supplement the data and improve the physical map. Examples of such other marker types are Sequence tagged sites (STS) markers, Simple Sequence Repeat (SSR) markers, AFLP markers, SNP markers, VNTR markers, RAPD markers or other types of genetic markers.

[50] The method for ordering the sequence tags based on co-presence resembles the HAPPY Mapping methodology. HAPPY mapping relies on the differential probability of two or more DNA sequences being physically separated. In genetic mapping, the probability of a recombination event between two genetic loci on the same chromosome is directly proportional to the distance between them. HAPPY mapping replaces recombination with fragmentation - instead of relying on recombination to separate genetic loci, the entire
genome is fragmented, for example, by radiation or mechanical shearing. If the DNA is broken on a random basis, the longer the distance between two DNA sequences, the higher the chances of it to break between the two, and vice versa. HAPPY mapping relies on the use of STS markers for the determination of co-presence, whereas the current invention, a.o. relies on the use of more reliable direct sequencing approach.

[51] In one embodiment of the present invention, two or more samples can be analysed and the ordering of their respective sequence tags can be compared. This allows for the identification of differences in the ordering of their sequence tags or of the presence or absence of sequence tags. It is advantageously if the sequence tags are obtained using the above methods that lead to principally the same result, for instance by using restriction enzymes in the fragmenting steps instead of random fragmentation, but this is not essential as in the end sequences are directly compared. The two or more samples can be from different individuals from the same species (be it human, non-human animal, plant, microorganism and the like) or of different parts of the same individual, or a sample that is to be tested for the presence or absence of a characteristic, a test sample (such as a known or suspected to contain an affliction or a disease) and a second sample that is a control sample. The two samples may be treated simultaneously and distinguished from one another by the introduction of identifiers in the analysed fragments or sequence tags.

Examples

Example 1 Description of the principle of the invention and comparison to sequence-based physical mapping using WO2008/007951

[52] A sequence-based physical map is constructed based on co-retention analysis of tag sequences in aliquots of BACs from a BAC library. A BAC library (e.g. 12X of which 6X EcoRI and 6X HindIII) is made from every sample being analyzed. However, instead of plating and colony picking individual BAC clones in 384-well plates, in aliquots of BAC clones are made by lifting the BAC library and making aliquots containing around 300 to 500 BAC clones, further referred to as instant BAC (iBAC) library. For example, in case of the melon genome of (450 Mbp genome size) and BACs with an average insert size of 125 kb, 384 BACs will comprise 48 Mbp of the genome, which is 0.106X genome equivalents. Hence, 113 aliquots provide 12X genome coverage of the melon genome (Table 1).

[53] Next, high quality DNA (low E. coli) is isolated from these aliquots containing about 384 BACs per aliquot and the terminal ends of (restriction) fragments generated from the BAC aliquots are generated to an average redundancy of 5 reads/tag (as opposed to 40 reads/tag in WO2008/007951). Next, a physical map is made by determining the co-presence of WGP tags in the different aliquots. This is the same principle used for radiation hybrid mapping and HAPPY mapping, but now on sequence tags instead of STS markers. Although
with the method of this invention it is not precisely known on which individual BAC clone a
particular tag sequence is located, the WGP tag sequence information for each aliquot
provides sufficient information to determine the relative order of these tags and thereby
building the physical map. The cost and effort-savings over of the present invention over
WO2008/007951 stems from the fact that no individual BAC clone plating and picking are
necessary, no 2D DNA pooling is needed, no library copying and individual clone storage is
needed, 20 to 40-fold less DNA preps are needed (depending on genome size and pooling
scheme used in WO2008/007951) and an estimated 8-fold less sequencing depth is needed
(Table 1). Also no deconvolution of the sequencing data to individual BACs is required.
Together, this leads to a much lower cost price for the physical map compared to the method
described in WO2008/007951.

| Table 1. Comparison features physical map produced using the current invention to the
method described in WO2008/007951 for the melon genome |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Genome Size (Mbp)</td>
<td>WO2008/007951</td>
<td>current invention</td>
</tr>
<tr>
<td>Insert size (kbp)</td>
<td>125</td>
<td>125</td>
</tr>
<tr>
<td>Coverage (GE)</td>
<td>12 X</td>
<td>12 X</td>
</tr>
<tr>
<td># BACs</td>
<td>43,200</td>
<td>43,200</td>
</tr>
<tr>
<td>Pooling/aliquoting set size (# BACs)</td>
<td>384</td>
<td>384*</td>
</tr>
<tr>
<td>Pool/Aliquot size (# BACs)</td>
<td>16 or 24</td>
<td>384*</td>
</tr>
<tr>
<td># pools/pooling (aliquots/aliquoting) set</td>
<td>40</td>
<td>1</td>
</tr>
<tr>
<td>Total # pools/aliquots = DNA extractions (2D)</td>
<td>4,500</td>
<td>113</td>
</tr>
<tr>
<td>Redundancy (#reads/tag)</td>
<td>40**</td>
<td>5</td>
</tr>
</tbody>
</table>

* aliquot size 3840 BAC may be too large for small genomes. For example 384 BACs of 125
kb = 48 Mbp, which is around 37% of 130 Mbp Arabidopsis genome. Simulation analysis can
be used to calculate optimal aliquot size in relation to genome size.

** Redundancy based on a 2-dimensional (2D) pooling design. 3D pooling requires an even
higher redundancy.
Example 2

[54] The principle of the present was investigated by aliquoting existing sequence data from an earlier experiment in *Arabidopsis thaliana* for the determination of a physical map. Tag data for aliquots of 96 BACs were simulated by grouping sequencing data from original aliquots of 16 or 24 BACs. The BAC DNA of these original aliquots consisted of EcoRI/Msel restriction fragments and thus the grouping resulted into larger sets of these restriction fragments. In total 128 aliquots were generated, each containing 96 BACs (approx. 0.1 GE) and the data consisted of approx. 0.5 M reads per aliquot each of 26 nt long.

[55] All tags were grouped by into unique sequences and analyzed on their frequency of occurrence in the panel of aliquots. A filter was applied with a minimum of 10 reads per tag on at least 2 aliquots. A maximum threshold was set for tags found in at most 40 different aliquots. This resulted in 62,173 different tags, which could further be binned into 22,284 'markers', showing a unique segregation pattern of presence and absence in the panel of aliquots. Groups of tags within the same bin were checked on their reference genome position and they all mapped consistently to the same region.

[56] Next a "genetic mapping" analysis was performed on a subset of the data to test the possibility of splitting the markers into linkage groups and to obtain the order of markers within groups. A subset of 1,610 markers, with a presence in at least 17 aliquots, was mapped into 153 linkage groups. A part of the mapping results of one of the linkage groups is given in Table 1. The position of the tags on the reference sequence was assigned and this showed that our approach is successful in mapping the vast majority of tags from a certain region (chr 3 in this case) of the genome into the proper order.

[57] Table 1. Results for Linkage Group 5 – 28 markers, 35.9 cM. Given are the Marker name, map position, number of tags (in the bin) covered by this marker, number of aliquots (from then total of 128) in which this marker is present, total number of reads of this marker and the corresponding positions of the tag on the reference genome sequence.

<table>
<thead>
<tr>
<th>Marker-ID</th>
<th>cM</th>
<th># tags</th>
<th># aliquots</th>
<th># reads</th>
<th>Chrom</th>
<th>Pos</th>
</tr>
</thead>
<tbody>
<tr>
<td>M21T001R0067</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>21</td>
<td>1130</td>
<td>Chr5</td>
</tr>
<tr>
<td>M20T002R0003</td>
<td>9.1</td>
<td>2</td>
<td>2</td>
<td>20</td>
<td>808</td>
<td>Chr3</td>
</tr>
<tr>
<td>M20T001R0037</td>
<td>9.2</td>
<td>1</td>
<td>1</td>
<td>20</td>
<td>251</td>
<td>Chr3</td>
</tr>
<tr>
<td>M21T002R0014</td>
<td>9.4</td>
<td>2</td>
<td>2</td>
<td>21</td>
<td>3140</td>
<td>Chr3</td>
</tr>
<tr>
<td>M23T001R0026</td>
<td>10.3</td>
<td>1</td>
<td>1</td>
<td>23</td>
<td>528</td>
<td>Chr3</td>
</tr>
<tr>
<td>M22T002R0004</td>
<td>10.6</td>
<td>2</td>
<td>2</td>
<td>22</td>
<td>2034</td>
<td>Chr3</td>
</tr>
<tr>
<td>M23T005R0002</td>
<td>11.8</td>
<td>5</td>
<td>5</td>
<td>23</td>
<td>3968</td>
<td>Chr3</td>
</tr>
<tr>
<td>M18T001R0052</td>
<td>11.9</td>
<td>1</td>
<td>1</td>
<td>18</td>
<td>185</td>
<td>Chr3</td>
</tr>
<tr>
<td>M19T001R0024</td>
<td>12.4</td>
<td>1</td>
<td>1</td>
<td>19</td>
<td>132</td>
<td>Chr3</td>
</tr>
<tr>
<td>M20T002R0008</td>
<td>14.2</td>
<td>2</td>
<td>2</td>
<td>20</td>
<td>1596</td>
<td>Chr3</td>
</tr>
<tr>
<td>M23T002R0007</td>
<td>14.4</td>
<td>2</td>
<td>2</td>
<td>23</td>
<td>2856</td>
<td>Chr3</td>
</tr>
<tr>
<td>M24T002R0002</td>
<td>14.4</td>
<td>2</td>
<td>2</td>
<td>24</td>
<td>2269</td>
<td>Chr3</td>
</tr>
</tbody>
</table>
Industrial Applicability

[58] Physical mapping, genome sequencing and structural genome analysis are important applications of today's genome research in crops and other genomes. The methods of the invention find utility in the generation of physical maps that can be used in (plant) breeding, in disease screening and many other applications.

Reference Signs List

[59] none

Reference to Deposited Biological Material

[60] none

Sequence Listing Free Text

[61] none

Citation List

15 Patent Literature

[63] WO2008/00795
[64] WO2008/00795
[65] WO2008/00795

20 [66] WO03/004690
[67] WO03/004690
[68] WO03/004690
[69] WO03/004690
[70] WO03/004690

25 [71] WO2005/003375,

Non Patent Literature

C L A I M S

1. Method for the ordering of sequence tags from at least part of genome, comprising the steps of providing a genomic clone library, generating clone aliquots from the library, generating sequences tags from the clone aliquots, ordering the sequence tags based on the combined presence of sequence tags in the clone aliquots.

2. Method according to clause 1, wherein the ordered sequences are used to generate a physical map of the genome.

3. Method for the generation of a physical map of at least part of a genome comprising the steps of:
   a) providing a genomic library containing at least part of the genome in a plurality of clones;
   b) generating clone aliquots of the genomic library, whereby each clone aliquots contains more than one clone;
   c) isolating DNA from the clone aliquots;
   d) sequencing at least part of the DNA of the clone aliquots to provide sequence tags;
   e) determining the combined presence of two or more sequence tags in the clone aliquots;
   f) ordering the sequence tags based on their combined presence in the clone aliquots, thereby generating the physical map.

4. Method according to clause 3, wherein the isolated DNA of step c) is subjected to the steps of
   i.) generating fragments from the DNA of the clone aliquots;
   ii.) ligating adapters to the fragments;
   iii.) optionally, providing identifiers that link the fragment to the aliquots;

5. Method according to clause 4, wherein the DNA is fragmented by restriction enzyme digestion.

6. Method according to clause 4, wherein the DNA is fragmented by random fragmentation.

8. Method according to clause 4, wherein the adapter ligated fragment is amplified from one or more primers prior to sequencing.
9. Method according to clause 4, wherein the identifier that links the fragment to the clone aliquots is provided in the adaptor.

10. Method according to clause 4, wherein the identifier that links the fragment to the aliquots is provided in the primer that is used in amplifying the adapter-ligated fragment.

11. Method for the detection of genomic variation between at least two samples comprising the comparison of the ordered sequence tags of the samples and identifying any variation between the samples.

12. Method according to clause 11, wherein the samples are from different individuals.

13. Method according to clause 11, wherein the samples are from the same individual but form different tissues or parts.

14. Method according to clauses 11 or 12, wherein a first sample is from individual 1 or a set of (aliquoted) individuals who contain(s) trait A and the second sample is from individual 2 or a set of (aliquoted) individuals who contain(s) trait B.

15. Method according to clause 12 or 13, wherein a first sample is from individual 1 or a tissue from individual 1 that is suspected to be affected by a disease, and a second sample is from individual 2 or another tissue from individual 1 that is considered a healthy control sample.
Co-presence analysis:

<table>
<thead>
<tr>
<th>SEQ TAG 1</th>
<th>SEQ TAG 2</th>
<th>SEQ TAG 3</th>
<th>SEQ TAG 4</th>
<th>SEQ TAG 5</th>
<th>SEQ TAG 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEQ TAG 1</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEQ TAG 2</td>
<td>97</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEQ TAG 3</td>
<td>91</td>
<td>98</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEQ TAG 4</td>
<td>74</td>
<td>85</td>
<td>96</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>SEQ TAG 5</td>
<td>7</td>
<td>76</td>
<td>84</td>
<td>98</td>
<td>100</td>
</tr>
<tr>
<td>SEQ TAG 6</td>
<td>3</td>
<td>10</td>
<td>70</td>
<td>84</td>
<td>97</td>
</tr>
</tbody>
</table>
**INTERNATIONAL SEARCH REPORT**

**A. CLASSIFICATION OF SUBJECT MATTER**

INV. C12Q1/68  C12N15/10
ADD.

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

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

C12Q  C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, BIOSIS, Sequence Search, EMBASE, WPI Data

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
</table>

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 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 novel or cannot be considered to involve an inventive step when the document is taken alone and the combined effect of several documents listed or the combined effect of the document listed and the invention cannot be considered to involve an inventive step when the document is taken together with the invention

  "A" document member of the same patent family

**Date of the actual completion of the international search**

29 September 2011

**Date of mailing of the international search report**

13/10/2011

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

**Authorized officer**

Mueller, Frank

Form PCT/ISA2/10 (second sheet) (April 2005)
<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>WO 02/103046 A2 (MEDICAL RES COUNCIL [GB]; DEAR PAUL H [GB]; THANGAVELU MADANE [GB]; BA) 27 December 2002 (2002-12-27) cited in the application the whole document</td>
<td>1-15</td>
</tr>
<tr>
<td>Patent document cited in search report</td>
<td>Publication date</td>
<td>Patent family member(s)</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>-----------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>WO 2007037678 A2</td>
<td>05-04-2007</td>
<td>AT 453728 T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU 2006295556 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA 2623539 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CN 101313078 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DK 1929039 T3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 1929039 A2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ES 2338459 T3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP 2009509527 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2009170713 A1</td>
</tr>
<tr>
<td>WO 2008007951 A1</td>
<td>17-01-2008</td>
<td>AT 481506 T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CN 101484589 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DK 2038425 T3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 2182079 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ES 2352987 T3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP 2009542256 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2009246780 A1</td>
</tr>
<tr>
<td>WO 02103046 A2</td>
<td>27-12-2002</td>
<td>AT 335852 T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA 2449366 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DE 60213803 T2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DK 1399590 T3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 1399590 A2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ES 2265507 T3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP 2004532047 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PT 1399590 E</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2004248153 A1</td>
</tr>
</tbody>
</table>