



(10) International Publication Number
WO 2011/071382 A1

(43) International Publication Date
16 June 2011 (16.06.2011)

PCT

- | | |
|--|---|
| (51) International Patent Classification:
<i>C12Q 1/68</i> (2006.01) | (81) Designated States (<i>unless otherwise indicated, for every kind of national protection available</i>): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW. |
| (21) International Application Number:
PCT/NL2010/050836 | (84) Designated States (<i>unless otherwise indicated, for every kind of regional protection available</i>): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG). |
| (22) International Filing Date:
9 December 2010 (09.12.2010) | Published:
— <i>with international search report (Art. 21(3))</i> |
| (25) Filing Language:
English | |
| (26) Publication Language:
English | |
| (30) Priority Data:
2003932 10 December 2009 (10.12.2009) NL
61/285,331 10 December 2009 (10.12.2009) US | |
| (71) Applicant (<i>for all designated States except US</i>): KEY-GENE N.V. [NL/NL]; P.O. Box 216, NL-6700 AE Wageningen (NL). | |
| (72) Inventors; and
(75) Inventors/Applicants (<i>for US only</i>): MICHELIS, An [BE/NL]; c/o P.O. Box 216, NL-6700 AE Wageningen (NL). VAN OEVEREN, Adriaan Jan [NL/NL]; c/o P.O. Box 216, NL-6700 AE Wageningen (NL). | |
| (74) Agent: DE LANG, R.,J. ; P.O. Box 3241, NL-2280 GE Rijswijk (NL). | |

Published:

— with international search report (Art. 21(3))

[Continued on next page]

(54) Title: POLYMORPHIC WHOLE GENOME PROFILING

(57) Abstract: The invention relates to a method for the generation of a physical map of a sample genome combined with the identification of polymorphisms by providing a physical map of a BAC library of a heterozygous sample genome based on the high throughput sequencing of tagged adapter-ligated restriction fragments of pooled BAC clones and determining polymorphisms between the tagged adapter-ligated restriction fragments.

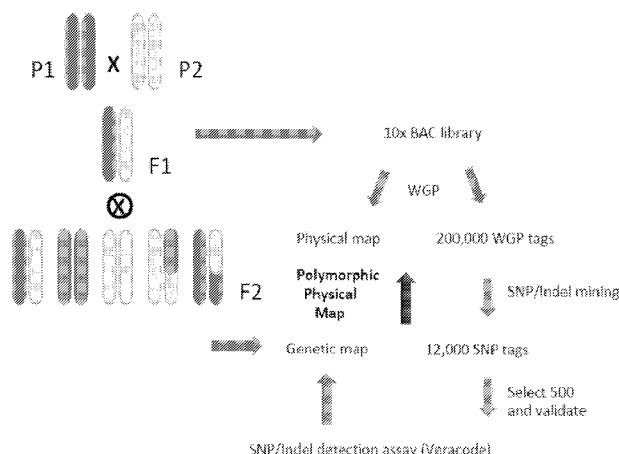


Fig 1



— *before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of*

amendments (Rule 48.2(h))

Title: Polymorphic whole genome profiling

Technical field of the invention

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. More in particular the invention relates to the generation of a physical map of a genome, or part thereof, using high-throughput sequencing technology, combined with the identification of polymorphic markers that are mapped to the physical map. The invention further relates to the development of marker assays of the discovered polymorphic markers and the use of the markers in the generation of a genetic map. More in particular, the invention relates to the identification of sequence tags, from which a physical map is created, and polymorphic sequence tags, from which a genetic map is made, and which can be integrated into a genome wide high density physical map.

Background of the invention

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). The fingerprints created from these restriction fragments are compared to identify related clones and to assemble overlapping clones in contigs. The utility of fingerprinting for ordering large insert clones of a complex genome is limited however, due to variation in DNA migration from gel to gel, the presence of repetitive DNAs, 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 hence less preferred than so-called direct methods such as sequencing.

Recently methods for high throughput sequencing have been made available that would allow for the determination of BAC clone ordering in a more efficient and cost-effective manner. Efficient methods are for instance disclosed in applicants own WO2008007951 relating to high throughput physical mapping. In this method, a physical map is generated from a combination of restriction enzyme digestion of clones in a library, pooling, restriction enzyme digestion, adapter-ligation, (selective) amplification, high-throughput sequencing and deconvolution of the resulting sequences results into BAC clone specific sets that can be used to assemble physical maps. The assembly of the clones into contigs is based on the co-presence of terminal nucleotide sequences of the sequenced fragments which can be used as sequence based anchor points for additional linkage of sequence data. More in detail, the technology disclosed in WO2008007951 is Whole Genome Profiling (WGP) , KeyGene's recently developed proprietary approach for sequence based physical mapping. Typically, a BAC library is constructed from a single homozygous individual and BAC clones are pooled in a multi-dimensional format. BAC pools are characterized by pool specific tags to allow assignment of sequences to individual BAC clones based on the coordinates in the multi-dimensional pool screening. DNA is extracted from each BAC pool and digested with restriction enzymes, for instance EcoRI and MseI. The EcoRI ends of the restriction fragments are analyzed on a next-generation sequencer such as the Illumina Genome Analyzer and in this way these relative short (20-100 basepairs) sequenced fragments, called the WGP Tags, can be assigned to individual BACs. In a next step BACs can be assembled based on overlapping WGP tag patterns using a contiging software tool such as FPC (Soderlund et al.). Typically this leads to contigs of assembled BACs, with WGP tags every 2 to 3 kilobases, about 30-40 tags per BAC clone.

Compared to other physical mapping approaches such as SNaPshot mapping, the WGP method is unique in providing sequence based anchor points instead of fragment lengths for assembly of BAC contigs. Sequence based anchors are more accurate and provide the basis for assembly of Whole Genome Shotgun data.

Definitions

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.

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 MseI) to 6 (EcoRI) and even 8 (NotI). The restriction enzymes used can be frequent and 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 64-1024 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 4096 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 MseI, is used in combination with a 5-cutter such as Avall, Avall is seen as the rare cutter and MseI as the frequent cutter.

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.

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.

Synthetic oligonucleotide: single-stranded DNA molecules having preferably from about 10 to about 50 bases, which can be synthesized chemically are referred to as synthetic oligonucleotides. In general, these synthetic DNA molecules are designed to have a unique or desired nucleotide sequence, although it is possible to synthesize families of molecules having related sequences and which have different nucleotide compositions at specific positions within the nucleotide sequence. The term synthetic oligonucleotide will be used to refer to DNA molecules having a designed or desired nucleotide sequence.

Adapters: short double-stranded DNA molecules 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. Adapters 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 adapter 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 adapter can be designed so that it cannot be ligated, but this need not be the case (double ligated adapters).

Adapter-ligated restriction fragments: restriction fragments that have been capped by adapters.

Primers: in general, the term primers refer 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.

DNA amplification: the term DNA amplification or 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.

Tagging: the term tagging refers to the addition of a sequence tag to a nucleic acid sample in order to be able to distinguish it from a second or further nucleic acid sample.

Tagging can e.g. be performed by the addition of a sequence identifier during complexity reduction or by any other means known in the art such as a separate ligation step. Such a sequence identifier can e.g. be a unique base sequence of varying but defined length uniquely used for identifying a specific nucleic acid sample. Typical examples are ZIP sequences, known in the art as commonly used tags for unique detection by hybridization (Iannone *et al.* Cytometry 39:131-140, 2000). Using nucleotide based tags, the origin of a sample, a clone or an amplified product can be determined upon further processing. In case of combining processed products originating from different nucleic acid samples, the different nucleic acid samples should be identified using different tags.

Identifier: a short sequence that can be added to an adapter or a primer or included in its sequence or otherwise used as label to provide a unique identifier (aka barcode or index). 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 $4(\text{exp}4) = 256$ different tags. Using such an identifier, the origin of a PCR sample can be determined upon further processing or fragments can be related to a clone. Also clones in a pool can be distinguished from one another using these sequence based identifiers. Thus, identifiers can be sample specific, pool specific, clone specific, amplicon specific etc. 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 and can be located at any convenient position.

Tagged library: the term tagged library refers to a library of tagged nucleic acids.

Aligning and alignment: With the term "aligning" and "alignment" is meant the comparison of two or more nucleotide sequence based on the presence of short or long stretches of identical or similar nucleotides. Several methods for alignment of nucleotide sequences are known in the art, as will be further explained below.

The term "contig" is used in connection with DNA sequence analysis, and refers to assembled 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. The term 'contig' is also used to indicate a contiguous stretch of, for instance, BACs, a "BAC-contig". A BAC contig can also be made on marker analysis, i.e. a more indirect way of sequence analysis. A "scaffold" is defined as a series of contigs that are in the correct order, but are not connected in one continuous sequence, 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.

Complexity reduction: the term complexity reduction is used to denote a method wherein the complexity of a nucleic acid sample, such as genomic DNA, is reduced by the generation or selection of a subset of the sample. This subset can be representative for the whole (i.e. complex) sample and is preferably a reproducible subset. Reproducible means in this context that when the same sample is reduced in complexity using the same method and experimental conditions, the same, or at least comparable, subset is obtained. The method used for complexity reduction may be any method for complexity reduction known in the art. Examples of methods for complexity reduction include for example AFLP® (Keygene N.V., the Netherlands; see e.g. EP 0 534 858), the methods described by Dong (see e.g. WO 03/012118, WO 00/24939), indexed linking (Unrau et al., *vide infra*), etc. The complexity reduction methods used in the present invention have in common that they are reproducible. Reproducible in the sense that when the same sample is reduced in complexity in the same manner, the same subset of the sample is obtained, as opposed to more random complexity reduction such as microdissection, random shearing, or the use of mRNA (cDNA) which

represents a portion of the genome transcribed in a selected tissue and for its reproducibility is depending on the selection of tissue, time of isolation etc..

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

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
10 hardware, it allows a researcher to effectively screen large amounts of samples simultaneously.

Artificial clone library: a population of hosts (bacteria, yeast), each of which carries a DNA molecule that was inserted into a cloning vector such that a representation of the genome of an organism is present (usually an entire genome) of clones, artificial
15 chromosome clone.

BAC: Bacterial Artificial Chromosome, a DNA construct, usually based on a functional fertility plasmid (or F-plasmid), used for transforming and cloning in bacteria, usually *E. coli*. The usual insert size is 100-350 kb but can also be 700kb. BACs are often used to sequence the genomes of organisms whereby a short piece of the organism's DNA is amplified as an
20 insert in BACs and then sequenced. Rearrangement of the sequenced parts *in silico* provides the genome sequence of the organism.

YAC: Yeast Artificial Library, the host is yeast, the insert size is larger (100-3000 kb),

Polymorphism: polymorphism refers to the presence of two or more variants of a nucleotide sequence in a population. A polymorphism may comprise one or more base
25 changes, an insertion, a repeat, or a deletion. A polymorphism includes e.g. a simple sequence repeat (SSR) and a single nucleotide polymorphism (SNP), which is a variation, occurring when a single nucleotide: adenine (A), thymine (T), cytosine (C) or guanine (G) - is altered. A variation must generally occur in at least 1% of the population to be considered a SNP. SNPs make up e.g. 90% of all human genetic variations, and occur every 100 to 300
30 bases along the human genome. Two of every three SNPs substitute Cytosine (C) with Thymine (T). Variations in the DNA sequences of e.g. humans or plants can affect how they handle diseases, bacteria, viruses, chemicals, drugs, etc.

Heterozygous: An organism is heterozygous for a particular gene when different alleles occupy the gene's position on the homologous chromosomes

35 Homozygous: An organism is homozygous for a particular gene when identical alleles are present on both homologous chromosomes

Summary of the invention

The present inventors found that 'Whole Genome Profiling' or WGP has proven to work well in providing high-quality physical maps. However, from analysis and validation on several data sets, it was observed that sometimes 'gaps' (missing WGP tags) in assembled
5 BACs occur: one overlapping BAC might not show the same set of WGP tags as another BAC covering the same region. These gaps could result from sequence errors or incomplete deconvolution due to inadequate sequencing depth, however the inventors realised that they could also have a biological background. If there would be a SNP or short indel within the WGP tag region, this would result in a polymorphic tag: either a present/absent tag or a tag
10 with a SNP or indel. Exactly these latter SNP variants provide a unique opportunity to use them as genetic markers.

The present inventors have found by using artificial chromosome libraries (particularly BAC libraries) from a heterozygous genome sample or a combination of two or more homozygous genome samples, this observed effect can be used in an efficient way to create
15 a physical map and, based on the same data set, screen the data for the presence of polymorphisms within the WGP-tags. This will significantly reduce the effort in SNP discovery and performing a large scale genotyping experiment. The addition of a (rough scale) genotyping and genetic mapping effort provides a link of BAC contigs into linkage groups. This genetic map can then be extended to a high density, high resolution map by adding all
20 SNP tags from their positions as known from the BAC contigs, resulting in an integrated physical and genetic map.

Description of the figures

25 Figure 1. Polymorphic physical mapping

Figure 2. Whole Genome Profiling approach

Figure 3. Whole Genome Profiling approach for polymorphic physical mapping

Figure 4. WGP Tags

Figure 5. SNP Tags

30 Figure 6: Various adapter-primer combinations containing identifiers to yield tagged fragments. A= adapter nucleotide, F= fragment nucleotide, P= primer nucleotide, N=degenerate nucleotide, ID= identifier nucleotides, → direction of amplification. 1.= adapter ligated fragment; 2. adapter ligated fragment containing identifier, a. amplification with primer directed against identifier, b. amplification with primer directed against adapter; 3. adapter
35 ligated fragment containing degenerate identifier section, amplification with primer directed against degenerate identifier and introducing identifier in amplified fragment. 4. adapter

ligated fragment containing no identifier, amplification with primer introducing identifier in amplified fragment.

Figure 7: Integration of genetic and physical map.

Figure 8 Alignment of BAC clones for contig 293. Indicated by the rectangular dotted box is the likely position of the polymorphic WGP tag SNP1.

Figure 9. Alignment of BAC clones for contig 307. Indicated by the rectangular dotted box is the likely position of the polymorphic WGP tag SNP2.

Detailed description of the invention

Thus in a first aspect the invention relates to a method for the generation of a physical map of a sample genome and identification of polymorphisms, comprising the steps of:

(a) providing an artificial chromosome (e.g. BAC, YAC) clone bank wherein each artificial chromosome clone contains DNA from a sample genome, wherein the sample genome is selected from the group consisting of

- a heterozygous sample genome,
- a combination of two or more homozygous sample genomes; and
- a combination of at least one heterozygous and at least one homozygous

sample genome;

(b) pooling the clones from the artificial chromosome library into pools;

(c) providing a set of fragments for each pool using restriction enzymes;

(d) ligating adapters to the fragments;

(e) determining the sequence of at least part of the adapter and part of the fragment;

(f) assigning the fragments to the corresponding clones;

(g) ordering the clones into clone-contigs thereby generating a physical map of the sample genome;

(h) aligning the sequenced fragments based on the determined sequences in step (e);

(i) determining polymorphisms between the aligned sequences of step (h).

In step (a) of the method an artificial clone bank is provided. The library can be a Bacterial Artificial Chromosome library (BAC) or based on yeast (YAC). Other libraries such as based on fosmids, cosmids, PAC, TAC or MAC are also possible. Preferred is a BAC library. The library is preferably of a high quality and preferably is a high insert size genomic library. This means that the individual BAC contains a relative large insert of the genomic DNA under investigation (typically > 100 kbp). The size of the preferred large insert is species-dependent. Throughout this application, reference can be made to BACs as

examples of artificial chromosomes. However, it is noted that the present invention is not limited thereto and that other artificial chromosomes can be used without departing from the gist of the invention. Preferably, the libraries contain at least five genome equivalents, more preferably at least 7, most preferably at least 8. Particularly preferred is at least 10. The higher the number of genome equivalents in the library, the more reliable the resulting contigs and physical map will be.

The sample genome DNA is thus selected from the group consisting of a heterozygous sample genome, a combination of two or more homozygous sample genomes; and a combination of at least one heterozygous and at least one homozygous sample genome.

The same WGP approach can be applied to a set of BAC clones from a heterozygous F1 individual. Thus, the BACs would differentiate the two sister chromosomes and polymorphisms will occur between them. It is expected that the majority of the WGP tags obtained from such "allelic" BAC clones will be identical (i.e. not containing a sequence polymorphisms) but a certain percentage, 5-10%, depending on the level of polymorphism and the length of the sequence read, will differ.

It is also possible to use two (or more), different, homozygous sample genomes. Such an approach may be efficient when for instance a BAC library of one homozygous sample is already available. Also the combination of at least one heterozygous sample genome and at least homozygous sample genome may yield a similar effect. The use of the latter two options may even result in more polymorphisms as more than two different sister chromosomes are present in the sample genomes.

To increase the likelihood that a WGP tag will contain a sequence polymorphism, it is preferred that longer sequence reads, for instance 76 nt or longer used. Obviously, the chance that a WGP tag contains a sequence polymorphism increases with increasing sequence read length. The available sequence read is typically platform dependent.

In step (b), the individual clones in the library are pooled to form pools containing a multitude of artificial chromosomes or clones. The pooling may be the simple combination of a number of individual clones into one sample (for example, 100 clones into 10 pools, each containing 10 clones), but also more elaborate pooling strategies may be used. The distribution of the clones over the pools is preferably such that each clone is present in at least two or more of the pools. Preferably, the pools contain from 10 to 10000 clones per pool, preferably from 100 to 1000, more preferably from 250 to 750. It is observed that the number of clones per pool can vary widely, and this variation is related to, for instance, the size of the genome under investigation. Typically, the maximum size of a pool or a sub-pool is governed by the ability to uniquely identify a clone in a pooling set by a set of identifiers. A typical range for a genome equivalent in a pool set is in the order of 0.2 – 0.3, and this may

again vary per genome. The pools are generated based on pooling strategies well known in the art. The skilled man is capable selecting the optimal pooling strategy based on factors such as genome size etc. The resulting pooling strategy will depend on the circumstances, and examples thereof are plate pooling, N-dimensional pooling such as 2D-pooling, 3D-pooling, 6D-pooling or complex pooling. To facilitate handling of large numbers of pools, the pools may, on their turn, be combined in super-pools (i.e. super-pools are pools of pools of clones) or divided into sub-pools. Other examples of pooling strategies and their deconvolution (i.e. the correct identification of the individual clone in a library by detection of the presence of a known associated indicator (i.e. label or identifier) of the clone in one or more pools or subpools) are for instance described in US6975943 or in Klein et al. in Genome Research, (2000), 10, 798-807. The pooling strategy is preferably such that every clone in the library is distributed in such over the pools that a unique combination of pools is made for every clone. The result thereof is that a certain combination of (sub)pools uniquely identifies a clone.

In step (c) of the method. the pools are digested with restriction endonucleases to yield restriction fragments. Each pool is preferably separately subjected to an endonuclease digest. Each pool is treated with the same (combination of) endonuclease(s). In principle any restriction endonuclease can be used. 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 genomes the use of, for instance, three or more restriction endonucleases can be used advantageously. To one or both ends of the restriction fragments, adapters are ligated in step (d) to provide for adapter-ligated restriction fragments. Typically, adapters are synthetic oligonucleotides as defined herein elsewhere. The adapters used in the present invention preferably contain an identifier section, in essence as defined herein elsewhere to provide for 'tagged adapters'. In certain embodiments, the adapter contains a pool-specific identifier, i.e. for each pool, an adapter containing a unique identifier is used that unequivocally indicates the pool. In certain embodiments, the adapter contains a degenerate identifier section which is used in combination with a primer containing a pool-specific identifier.

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

Alternatively, the adapters that are ligated do not contain an identifier or a degenerate identifier section. The adapter-ligated fragments are subsequently amplified using primers that contain identifiers (tags), for instance at their 5' end. The result is that amplified, tagged adapter-ligated fragments are obtained.

5 In this embodiment, the adapters can be the same for a plurality (or all) of pools and the amplification using tagged primers creates the distinction between the pools that can later be used in the deconvolution.

Either way, a set of tagged adapter-ligated fragments is obtained that are linked to the pool from which they originate by the presence of the tag.

10 The tagged adapter-ligated fragment can be amplified. The amplification may serve to reduce the complexity or to increase the amount the DNA available for analysis. 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 independently from the amplification described herein above that introduces the tags into the adapters. In certain
15 embodiments, the amplification may serve several purposes at a time, i.e. reduce complexity, increase DNA amount and introduce tags in the adapter-ligated fragments in the pools.

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

The adapter-ligated fragments can be amplified using a set of primers of which at least one primer amplifies the pool-specific identifier at the position of the pool-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 tag,
25 i.e. downstream in the adapter. Amplification then also amplifies the tag. See in this respect Fig 6 for various embodiments.

In step (e) part of the sequence of the tagged adapter-ligated fragment is determined. The tagged adapter-ligated fragments are subjected to sequencing, preferably high throughput sequencing as described herein elsewhere. During sequencing, at least part of the
30 nucleotide sequence of the (amplified) tagged adapter-ligated fragment is determined. Preferably at least the sequence of the pool-specific identifier and part of the fragment (i.e. derived from the sample genome) of the (amplified) tagged adapter-ligated fragment 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
35 from the sample genome) are determined. The number of nucleotides that are to be determined minimally will be, again, genome- as well as sequencing platform dependent. For instance, in plants more repetitive sequences are present, hence longer sequences (50-150

nucleotides) may to be determined for a contig of comparable quality. 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 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 for contig building of a BAC clone. In general the method of the present invention, there is a preference for longer reads, from 50 nt upward, more preferably 75 nt and more, even more preferably more than 100 nt, to increase the probability of detecting SNPs.

In the sequencing step, to provide for maximum coverage of all fragments and increased accuracy, the sequence library may be sequenced with an average redundancy level (aka oversampling rate) of at least 5. This means that, on average, the sequence is determined of at least 5 amplicons obtained from the amplification of one specific adapter-ligated fragment. 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 and 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.

It is preferred that the sequencing is performed using high-throughput sequencing methods, such as the methods disclosed in WO 03/004690, WO 03/054142, WO 2004/069849, WO 2004/070005, WO 2004/070007, and WO 2005/003375, by Seo et al. (2004) Proc. Natl. Acad. Sci. USA 101:5488-93, and technologies of Helicos, Illumina), US Genomics, etcetera, which are herein incorporated by reference.

In the following step (f), the (partly) sequenced (amplified) tagged adapter-ligated fragments are correlated to the corresponding clone, typically in silico by means of computerized methods. The (amplified) tagged adapter-ligated fragments are selected that contain identical sections of nucleotides in the restriction fragment-derived part. Subsequently the different pool-specific identifiers (tags) are identified that are present in those (amplified) tagged adapter-ligated fragments. The combination of the different pool-specific identifiers and hence the sequence of the restriction fragment can be uniquely assigned to a specific clone (a process indicated as 'deconvolution'). For example, in the case of a 3D pooling strategy (X,Y,Z), each pool in the library is uniquely addressed by a combination of 3 pool-specific identifiers with the same restriction fragment-derived section. In other words: a restriction fragment-derived section originating from a clone will be tagged with 3 different

identifiers. Unique restriction fragment-derived sections, when observed in combination with the 3 identifiers can be assigned to a single BAC clone. This can be repeated for each (amplified) tagged adapter-ligated fragment that contains other unique sections of nucleotides in the restriction fragment-derived part.

5 After assigning the fragments to the corresponding clones in step (f), the clones are combined and ordered into clone contigs in step (g) of the method. The grouping and ordering can be performed by fingerprint contiging software for this purpose such as FPC software (Soderlund et al (1997) FPC: a system for building contigs from restriction fingerprinted clones. *Comput. Appl. Biosci.*, 13:523–535.) essentially as described herein elsewhere. The
10 alignment of the clones into contigs and the corresponding order of WGP tags generates a physical map of the sample genome.

 In certain embodiments, for instance where a physical map is already available for a homozygous sample genome, the above steps can be performed independently for each genome sample. More in particular, in certain embodiments, for each heterozygous and/or
15 homozygous sample genome steps (a) – (g) are performed independently. The steps relating to the screening for polymorphisms in the subsequent steps are taken from the separate steps (e).

 In step (h) of the method, the sequenced fragments (the WGP tags) are aligned based on the determined sequences in step (e) of the method. Typically, sequenced fragments
20 comprise a sample identification tag, identifying the pool from which the BAC was derived (and hence the fragment). The sequence fragments further comprise the remains of the restriction site and further a sequence that is derived from the sample genome (sometimes indicated as the 'sequence tag'). By aligning the sequence tags, (see for example fig 4, Pool C19 and Pool R3 and fig 5, pool C19, C13, RA and RH) themselves, polymorphisms can be
25 identified between them, in this case allele 'G' for C19 and RA vs. allele 'C' for C13 and RH.

 In step (i) the polymorphisms are determined between the aligned sequences of the WGP tags of step (h).

 For most situations, a polymorphism in the WGP tag region will result in two variants, with a SNP (or Indel) somewhere in the 65 nt, non-restriction site part of a (75 nt sequence
30 read length)) WGP tag (which is now termed 'SNP Tag', i.e. a SNP Tag is a WGP tag that contains a polymorphism or an indel between its two variants, See fig 3). Given equal portions of BACs and pool samples, the two variants ('alleles') of each SNP Tag are expected to be found in a 50/50 distribution. In those cases where a polymorphism is present in the restriction site region, it will cause the WGP tag to be present in only 50% of the BACs from
35 the corresponding region in the genome. These tags are less useful to discriminate between alleles as it is not sure whether the tags are missing because the region was not covered by the BAC, or because of a polymorphism.

The low number of SNP tags will still allow the formation of contigs from overlapping BACs, as obtained in a standard WGP approach, albeit with less stringent alignment (FPC) settings. In an alternative embodiment, a binning step is executed first to combine allelic variants (and possibly sequence errors) into a single WPG tag followed by a FPC step as for a homozygous line (i.e. with more stringent settings). In parallel or subsequent to the contig building step, an additional analysis will be done to identify "SNP tags" and their position on the contigs or physical map.

In a further aspect of the invention, the discovered polymorphism and SNPs are converted into conventional SNP assays using conventional technology. Such SNP tags can due to the existence of sequence information be converted to PCR-assays, invader assays, Golden Gate assays and the like. The developed assay can be validated on the sample genomes(s) used to generate the SNPs.

The discovered SNP tags (and underlying SNPs) are both physical markers (as their position on the physical map is known) and genetic markers, as they differ between sample genomes. This allows to couple genetic and physical maps (See Fig 7).

The thus discovered SNPs can be followed in subsequent crosses, for instance in the offspring of crosses between parents of which at least one of the parents was used to create the physical map. In this manner, the SNPs can serve as genetic markers that are linked to the physical map. At the same time, the BACs used to generate the physical map can be anchored to the genetic map, resulting in a high resolution map based on a scaffold of genetic markers (SNP tags) and supplemented with WGP tags. It is further possible to also link genetic markers obtained via other ways to the map to further complete the high resolution map.

Thus, a further aspect of the invention relates to a method for the generation of a linked genetic map of genetic markers (SNP tags) and a physical map (of WGP tags) comprising

- providing two parents, wherein at least one of the parents has been used as a sample genome in the generation of the physical map that provided the SNP tags of step (i);
- crossing the parents
- screening the offspring for the presence (or absence) of the SNP tags;
- providing a genetic map from the SNP tags
- coupling the genetic map tags to the physical map.

The genetic linkage map is constructed using the markers discovered using the method of the invention. The map can be constructed by observing how frequently pairs of SNP Tag markers are inherited together after selfing an F1 population and analyzing the F2 individuals (about 100). The thus obtained genetic map shows the relative locations of these

SNPs along the chromosome, thus providing the link with the position of the SNP tags on the BAC contigs of the physical map. Typically, it would suffice to use a single SNP assay of 384 validated SNPs (Veracode), derived from about 500 candidates.

5 Example

This example is to provide an illustration for the polymorphic Whole Genome Profiling (pWGP) concept described herein. The invention is aimed to perform physical mapping and SNP discovery at the same time, by executing a WGP project on a BAC library derived from a heterozygote individual or a combination of two polymorphic individuals. In this example data have been used from both *Arabidopsis thaliana* Columbia and Landsberg *erecta* ecotypes.

Data

Existing WGP Columbia (Col) data were used and combined with data from a new WGP project on Landsberg *erecta* (Ler). Where the Col data consisted of 16 plates (6,144 BACs), the Ler set consisted of 24 plates (9,216 BACs – see Table 1). The Ler BACs were pooled in a 3D manner using an H1Quad pooling setup with 22 pools (termed 1 SuperPool = SP) per plate. The BAC pools were processed with the regular WGP protocol of EcoRI/MseI restriction ligation, amplification and subsequent sequencing on the Illumina GAI. The 24 SPs were sequenced using only two lanes. Deconvolution and subsequent analysis also conformed to the WGP protocol.

Results

WGP map

Results in terms of numbers of reads and tags are given in Table 1.

Table 1. Deconvolution results WGP on *Arabidopsis* Col and Ler

	Arabidopsis		
	Colombia	Landsberg	Combined
genome size (Mbp)	130	130	130
WGP tag length (incl. RE site)	26	26	26
nr BACs tested	6,144	9,216	15,360
GE BACs tested (125 kb inserts)	5.9	8.9	14.8
nr OK reads generated (M)	28.2	62.9	91
nr deconvolvable reads (M)	12.1	28.9	45
% deconvolvable reads	43%	46%	50%
all deconvolvable tags	183,366	344,955	528,321
nr unique tags	65,734	62,423	84,710

nr tagged BACs	4,599	7550	12,149
average nr tags/BAC	40	46	43
average nr reads/tag	66	84	86

The combined set of Col and Ler BACs was subjected to mapping using FPC. As a large fraction of the clones clustered to one single contig with low stringent setting, the cutoff had to be decreased to 1.0×10^{-20} . After an additional DQ step (step size = 5) 378 contigs were obtained, which encompassed 10,299 BACs (85% of all deconvoluted BACs). The combined contigs covered an estimated 132 Mbp (Table 2).

Table 2. FPC results for both the original Col WGP map and the combined Col+Ler WGP data.

FPC	Col	Col+Ler
Cutoff	1.0E-06	1.0E-20
nr contigs	273	378
nr singletons	551	1850
nr BACs in contig	4,048	10,299
% BACs in contig (nr BACs in contig/nr BACs tested)	66%	67%
% BACs in contig (nr BACs in contig/nr tagged BACs)	88%	85%
average BACs/contig	14.8	27.2
N50 BACs/contig	25	49
average contig size (Mbp)	0.369	0.347
N50 contig size (Mbp)	0.548	0.608
genome coverage (Mbp)	100.8	131.3
% genome coverage	78%	101%

SNP discovery

In the next step, WGP tags were identified which were specific to either the Col or the Ler data. When a single Col tag matches a single Ler tag with only one nucleotide difference, than such a tag is a putative SNP marker. Two of such candidates and corresponding BAC clones are identified. They are presented in Table 3.

Table 3. Two examples of SNPs between Col and Ler (bold underlined characters indicate the SNP)

Ctg293	BAC IDs
SNP1 GAATTCAGCTGTGAACATAACCG <u>AAA</u> Col	AT.E102.G22, AT.E110.N08, AT.E115.P18, AT.E112.D16,

GAATTCAGCTGTGAACATAACCGGA A Ler AT.E022.F04, AT.E013.A12,
 AT.E009.H24, AT.E010.E18,
 AT.E001.L24, AT.E005.A07,
 AT.E015.E24, AT.E018.C01,
 AT.E004.D16, AT.E003.L16,

Ctg307

SNP2	GAATTCAGGAGAAATAGAGTG <u>G</u> ATAT	Col	AT.E110.I20, AT.E102.E22, AT.E113.D18, AT.E101.J24,
	GAATTCAGGAGAAATAGAGTG <u>A</u> ATAT	Ler	AT.E021.P20, AT.E003.M03, AT.E015.B21, AT.E004.K13,

The BAC clones corresponding to a single SNP tag all map to a single region on the contig as is visualized in Figures 8 and 9 for contig 293 and contig 307 resp.

- 5 From the data it is concluded that a physical map can be built from the combined WGP data of two different genotypes and that SNPs can be detected in this combined data set between genotype specific WGP tags, mapping to the same location on the physical map

CLAIMS

1. Method for the generation of a physical map of a sample genome and identification of polymorphisms, comprising the steps of:

- 5 (a) providing an artificial chromosome (e.g. BAC, YAC) clone bank wherein each artificial chromosome clone contains DNA from (part of) a sample genome, wherein the sample genome is selected from the group consisting of a heterozygous sample genome, a combination of two or more homozygous sample genomes or a combination of at least one heterozygous and at least one
- 10 homozygous sample genome;
- (b) pooling the clones from the artificial chromosome library into pools;
- (c) providing a set of fragments for each pool using restriction enzymes;
- (d) ligating adapters to the fragments;
- (e) determining the sequence of at least part of the adapter and part of the
- 15 fragment;
- (f) assigning the fragments to the corresponding clones;
- (g) ordering the clones into clone-contigs thereby generating a physical map of the sample genome;
- (h) aligning the sequenced fragments based on the determined
- 20 sequences in step (e);
- (i) determining polymorphisms between the aligned sequences of step (h).

2. Method according to claim 1, wherein for each heterozygous and/or homozygous sample genome steps (a) – (g) are performed independently and whereby

25 in step (h) the sequences determined in step (e) are taken from the separate steps (e).

3. Method according to claim 1, wherein the adapter contains an identifier or a degenerate identifier section.

30 4. Method according to claims 1-3, wherein the adapter-ligated fragments are amplified using

- a primer that amplifies at least the identifier and part of the fragment; or
- a primer that contains a section that is complementary to the degenerate section in the adapter and introduces an identifier in the amplified fragment; or
- 35 - a primer that is complementary to at least part of the adapter and provides an identifier in the amplified adapter-ligated fragment.

5. Method according to claim 1, wherein the sequence of at least consecutive 50 nucleotides are determined from the tagged adapter and part of the fragment.

5 6. Method according to claim 1, wherein the sequence of at least consecutive 75 nucleotides are determined from the tagged adapter and part of the fragment.

7. Method according to claim 1, wherein the sequence of at least consecutive 100 nucleotides are determined from the tagged adapter and part of the fragment.

10 8. Method according to claim 1, wherein the determined polymorphisms are converted to marker assays.

9. Method for the generation of a linked genetic map and a physical map comprising

- 15
- providing two parental lines, or a heterozygous line wherein at least one of the lines has been used as a sample genome in the generation of the physical map that provided the polymorphisms of step (i) of claim 1;
 - crossing the parents or selfing the heterozygote;
 - screening the offspring for the presence (or absence) of the
- 20
- polymorphisms and/or markers associated with the polymorphisms;
 - providing a genetic map from the polymorphisms;
 - coupling the genetic map tags to the physical map.

Fig 1

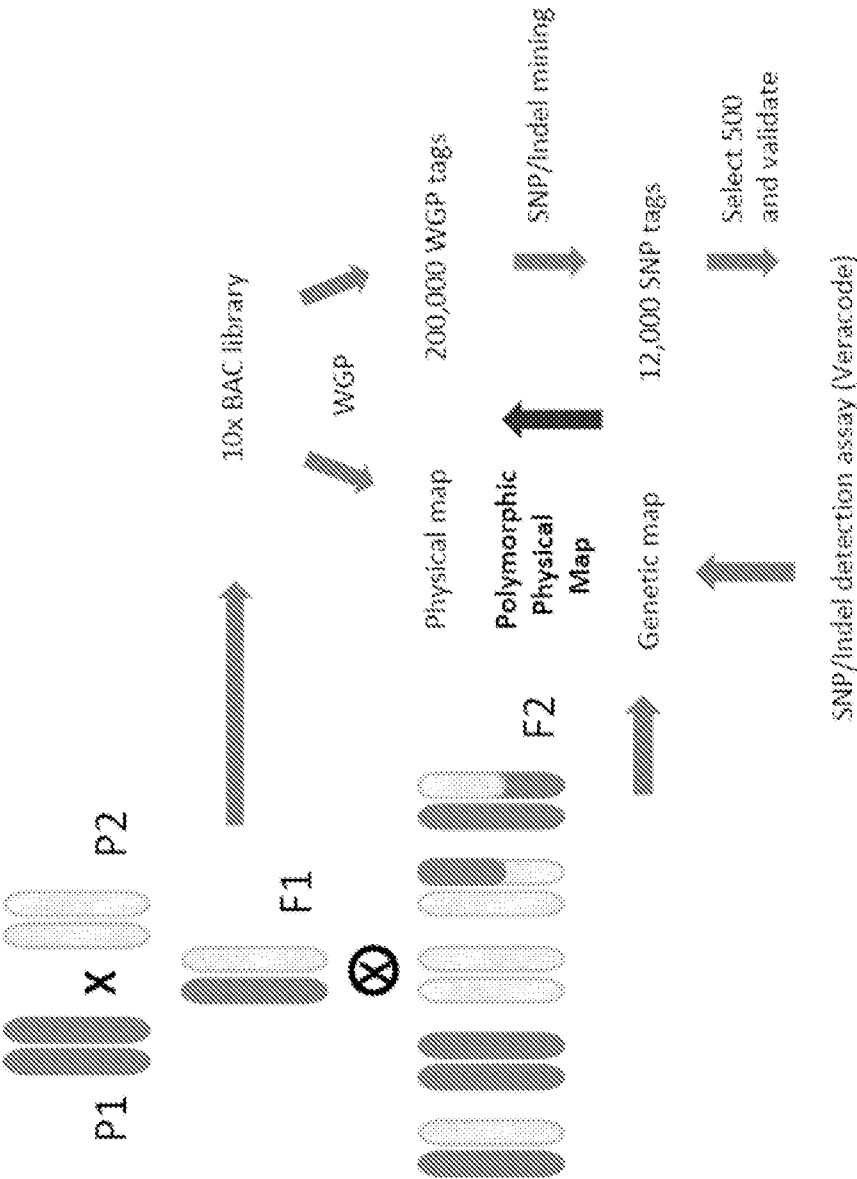


Fig 2

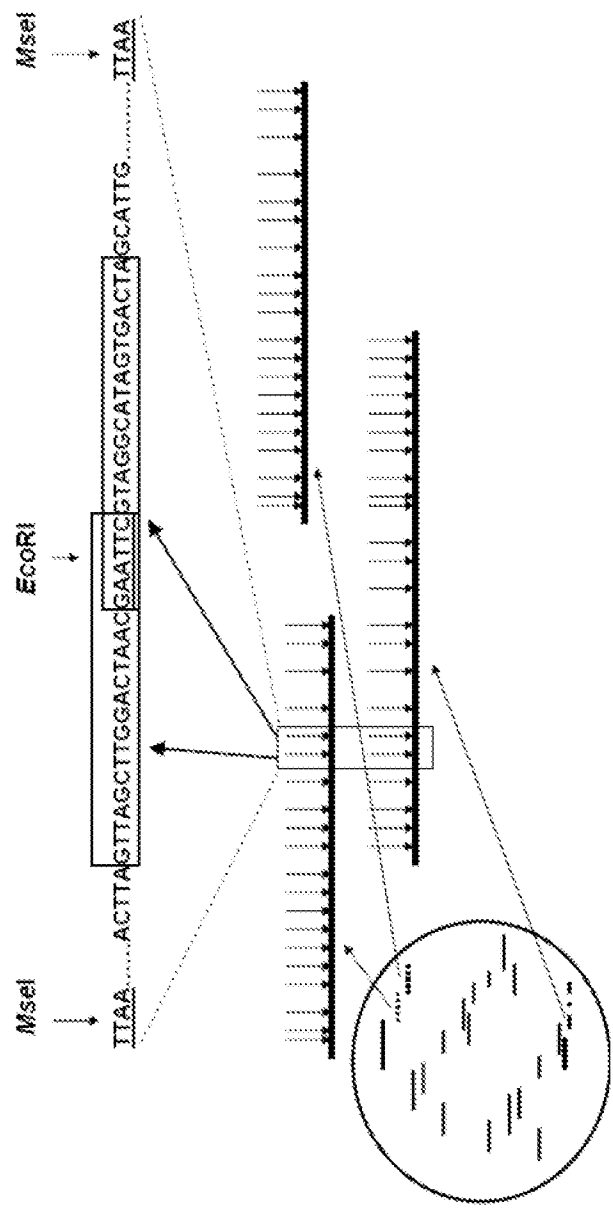


Fig 3

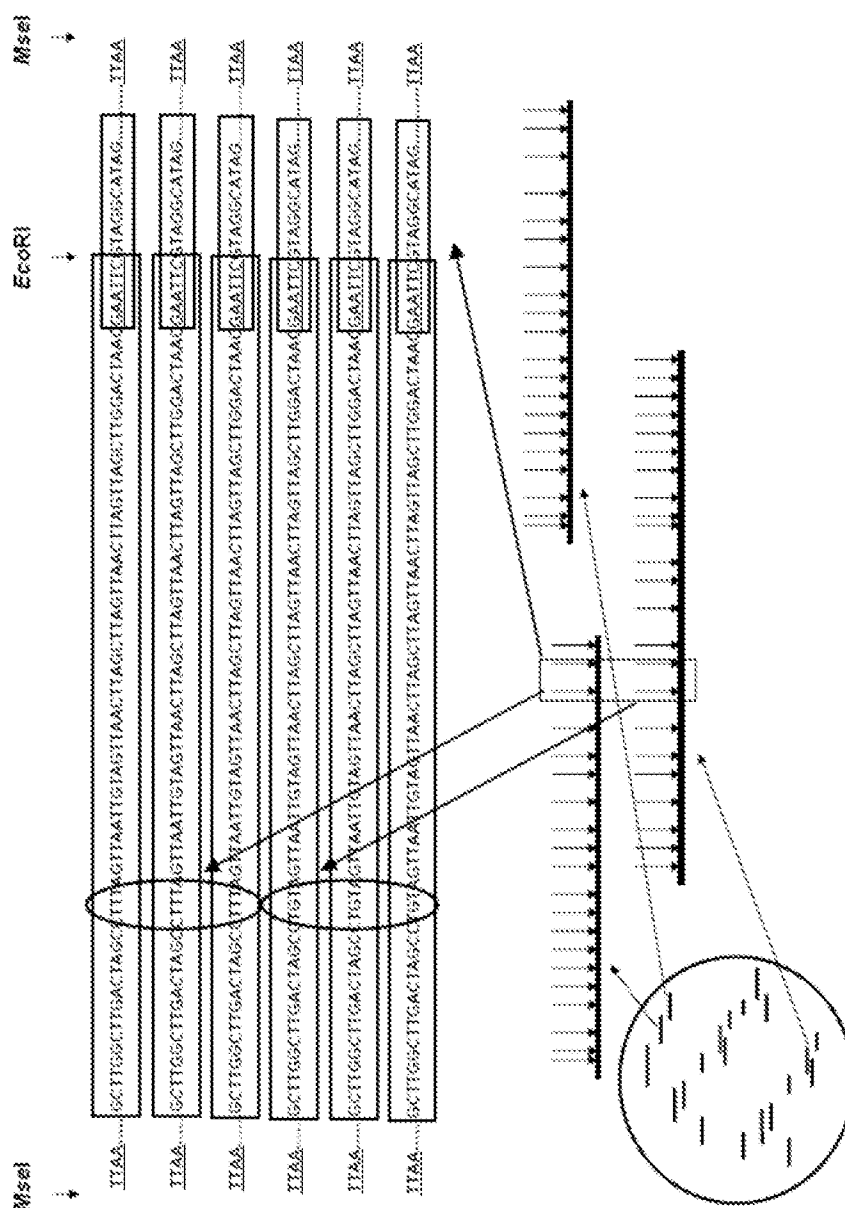


Fig 4

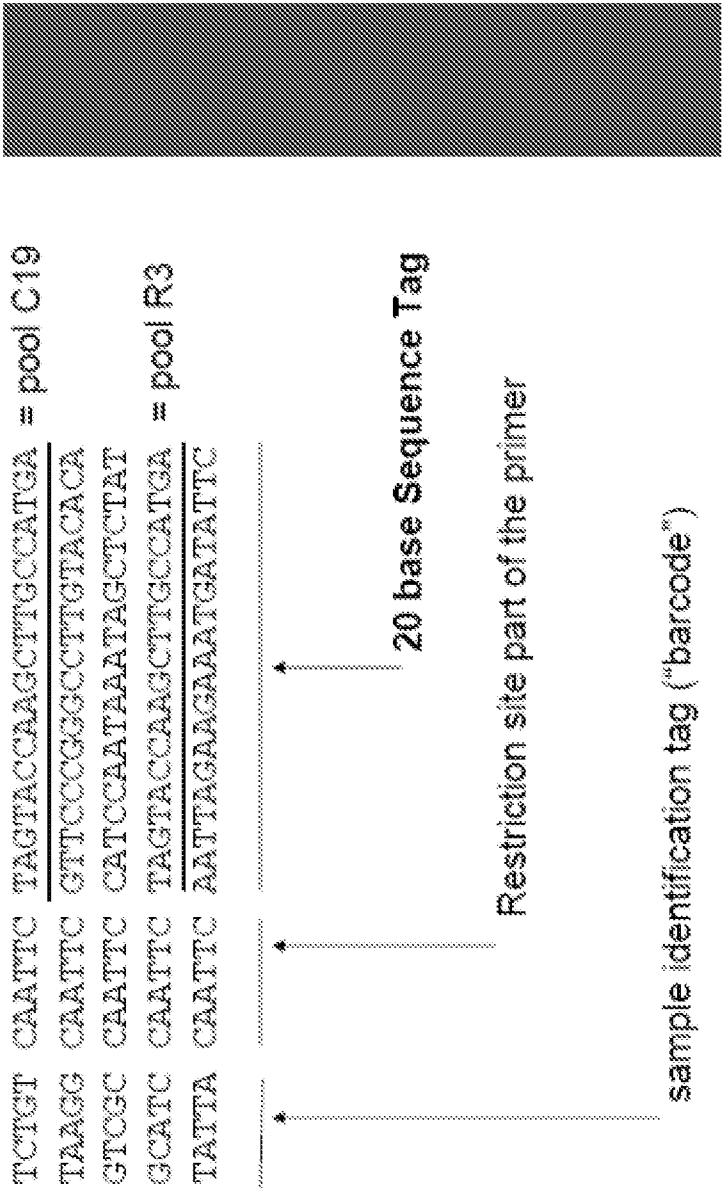


Fig 5

5

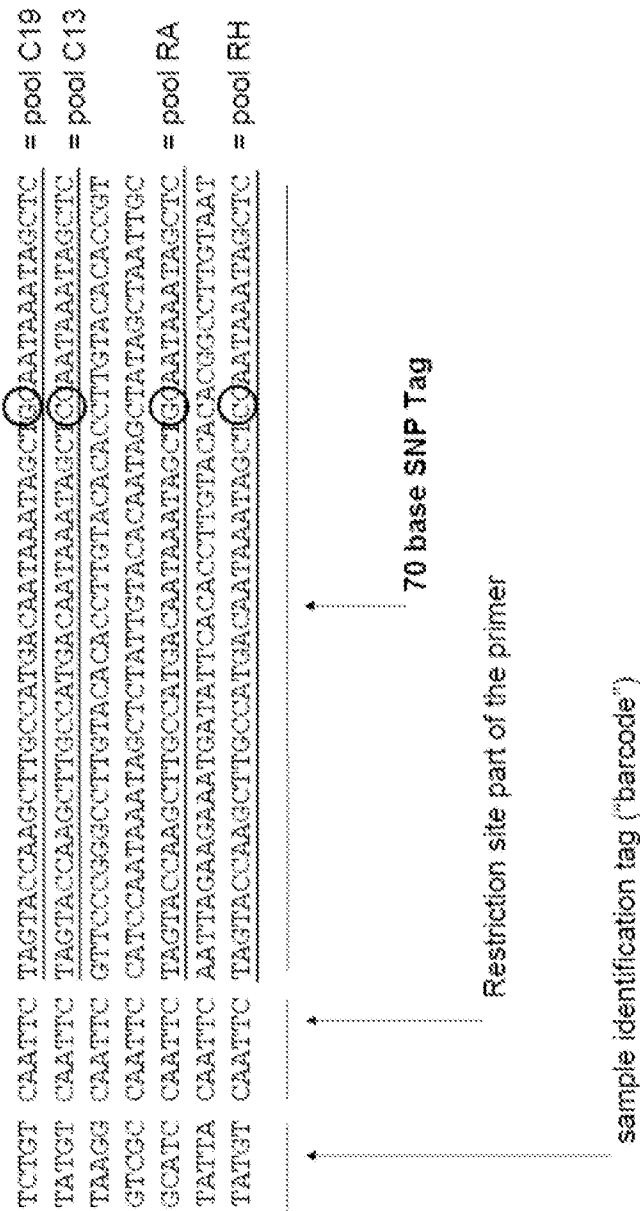


Fig 6

5

1. AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAFFFFFFFFFFFFFFFF

10

2. AAAAAAIDIDIIDIDAIAAAAAAAAAAAAAAAAAAFFFFFFFFFFFFFFFF

a. PPPPPP→

b. PPPPPP→

3. AAAAAANNNNNNAIAAAAAAAAAAAAAAAAAAFFFFFFFFFFFFFFFF

15

PPPPIDIDID→

4. AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAFFFFFFFFFFFFFFFF

20

IDIDIDPPPP→

Fig 7

5

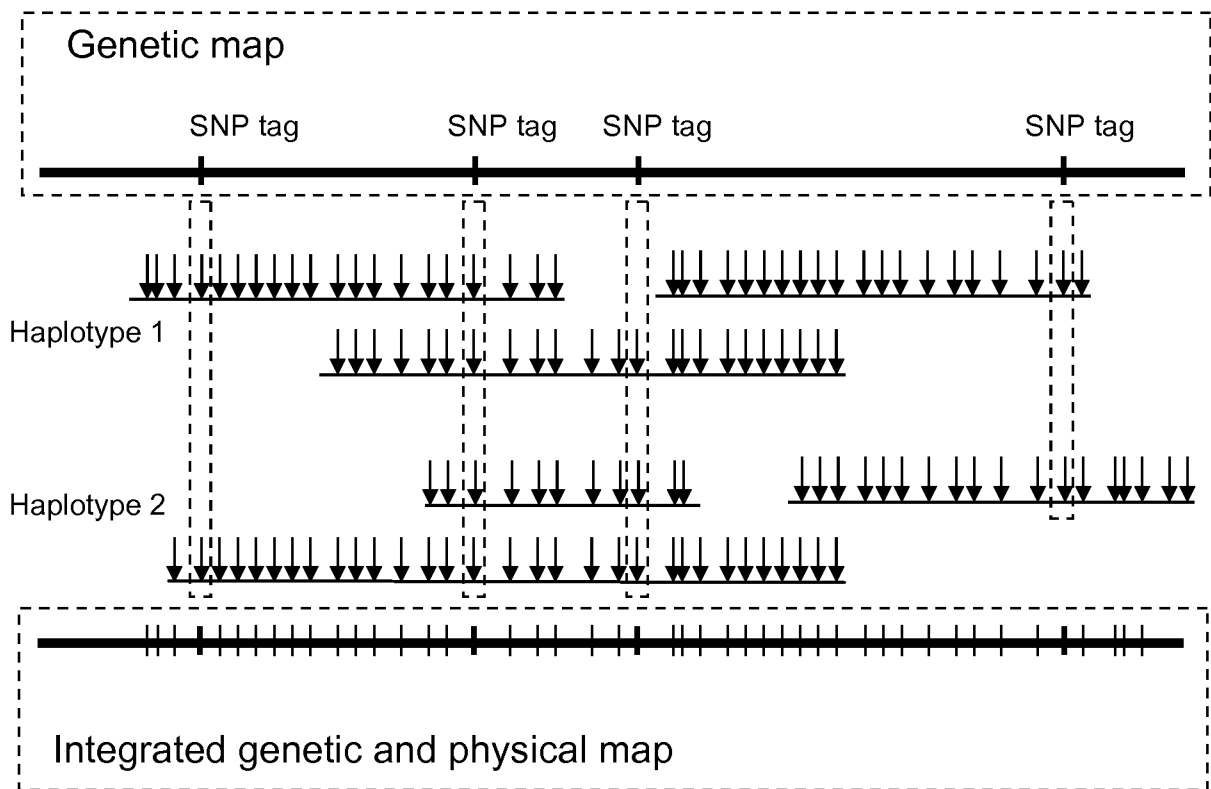
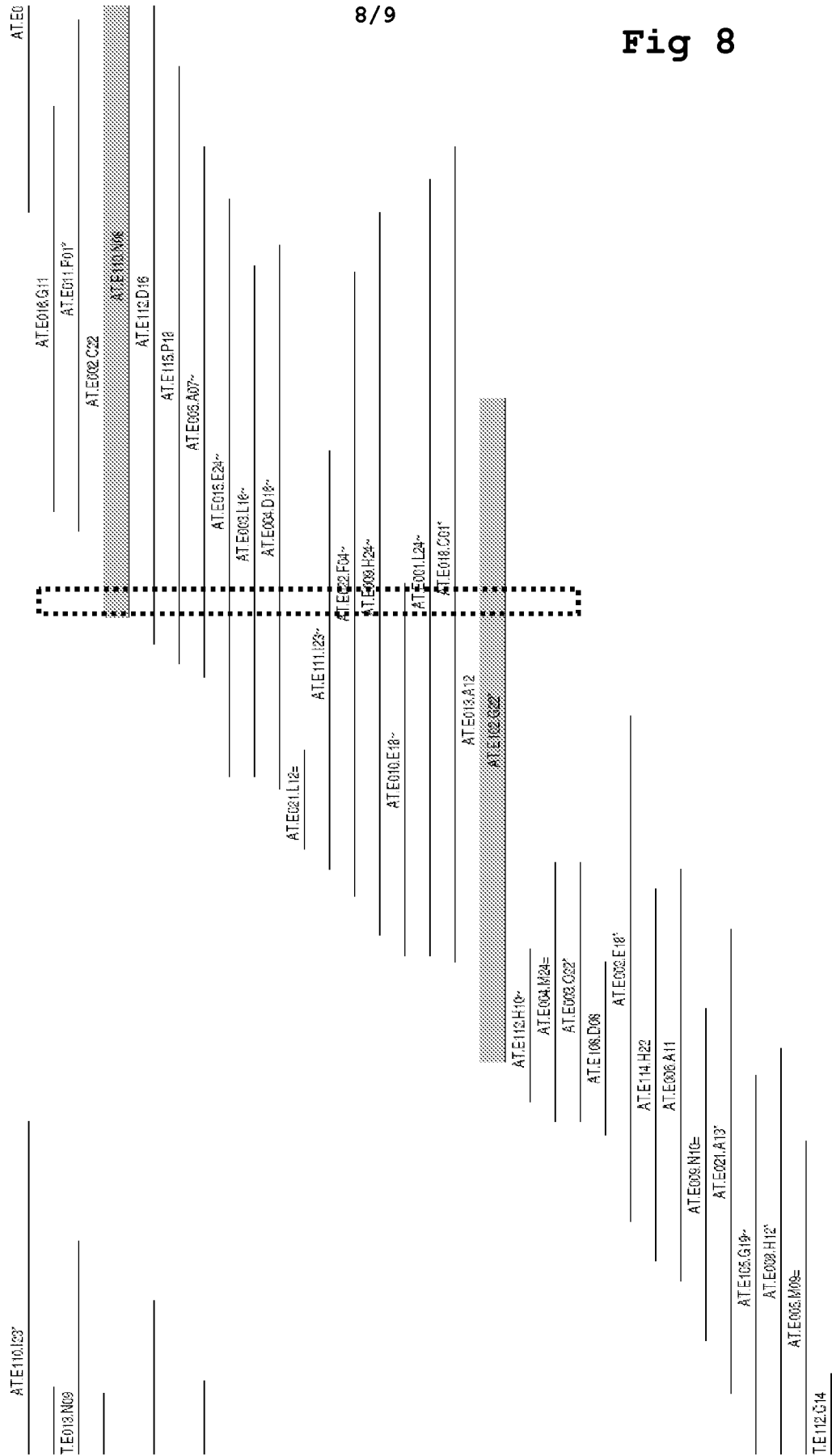


Fig 8



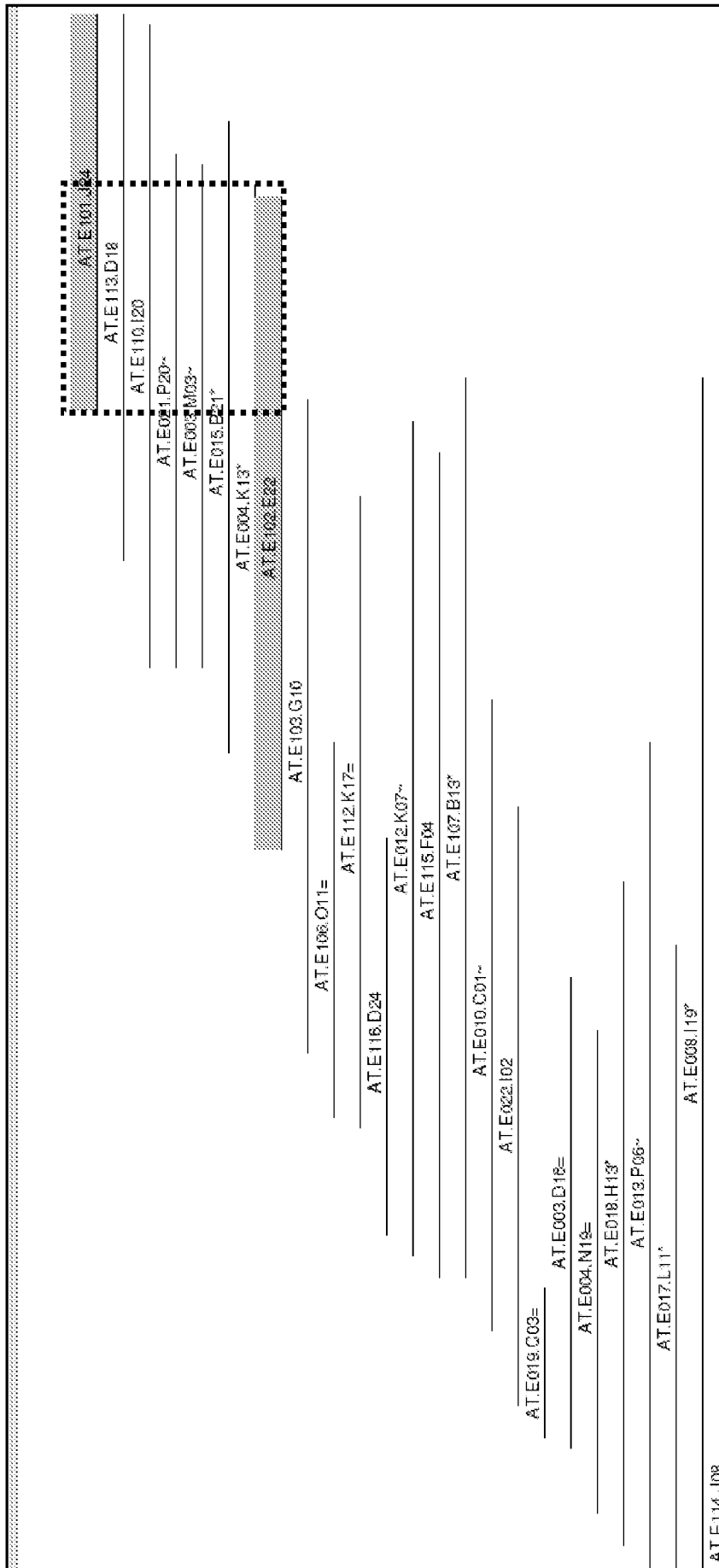


Fig 9

INTERNATIONAL SEARCH REPORT

International application No
PCT/NL2010/050836

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12Q1/68
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

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, CAB Data, Sequence Search, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2008/007951 A1 (KEYGENE NV [NL]; VAN EIJK MICHAEL JOSEPHUS THER [NL]; JESSE TACO PETER) 17 January 2008 (2008-01-17) cited in the application the whole document -----	1-9
A	WO 2006/137734 A1 (KEYGENE NV [NL]; VAN EIJK MICHAEL JOSEPHUS THER [NL]; SOERENSEN ANKER) 28 December 2006 (2006-12-28) the whole document -----	1-9
A	WO 2004/063323 A2 (KEYGENE NV [NL]; JESSE TACO PETER [NL]) 29 July 2004 (2004-07-29) the whole document ----- -/-	1-9



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 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

31 March 2011

Date of mailing of the international search report

08/04/2011

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

Hornig, Horst

INTERNATIONAL SEARCH REPORT

International application No
PCT/NL2010/050836

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 03/027311 A2 (BAYLOR COLLEGE MEDICINE [US]; GIBBS RICHARD [US]; BRADLEY ALLAN [GB];) 3 April 2003 (2003-04-03) cited in the application the whole document	1-9
A	----- KLEIN P E ET AL: "A high-throughput AFLP-based method for constructing integrated genetic and physical maps: Progress toward a sorghum genome map", GENOME RESEARCH, COLD SPRING HARBOR LABORATORY PRESS, WOODBURY, NY, US LNKD-DOI:10.1101/GR.10.6.789, vol. 10, no. 6, 1 June 2000 (2000-06-01), pages 789-807, XP002240094, ISSN: 1088-9051 cited in the application the whole document	1-9
A	----- CAI W-W ET AL: "A clone-array pooled shotgun strategy for sequencing large genomics", GENOME RESEARCH, COLD SPRING HARBOR LABORATORY PRESS, WOODBURY, NY, US LNKD-DOI:10.1101/GR.198101, vol. 11, 1 January 2001 (2001-01-01), pages 1619-1623, XP002967818, ISSN: 1088-9051 the whole document	1-9
A	----- EP 0 969 102 A2 (KEYGENE NV [NL]) 5 January 2000 (2000-01-05) the whole document	1-9
A	----- EP 0 534 858 A1 (KEYGENE NV [NL]) 31 March 1993 (1993-03-31) cited in the application the whole document -----	1-9

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/NL2010/050836

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2008007951	A1	17-01-2008	AT 481506 T 15-10-2010
		CN 101484589 A 15-07-2009	
		DK 2038425 T3 06-12-2010	
		EP 2038425 A1 25-03-2009	
		EP 2182079 A1 05-05-2010	
		EP 2275576 A1 19-01-2011	
		ES 2352987 T3 24-02-2011	
		JP 2009542256 T 03-12-2009	
		US 2009246780 A1 01-10-2009	
WO 2006137734	A1	28-12-2006	AT 465274 T 15-05-2010
		AU 2006259990 A1 28-12-2006	
		CA 2613248 A1 28-12-2006	
		CN 101278058 A 01-10-2008	
		EP 1910563 A1 16-04-2008	
		ES 2344802 T3 07-09-2010	
		JP 2008546405 T 25-12-2008	
		US 2009142758 A1 04-06-2009	
WO 2004063323	A2	29-07-2004	EP 1581661 A2 05-10-2005
		JP 2006516894 T 13-07-2006	
		US 2006246445 A1 02-11-2006	
WO 03027311	A2	03-04-2003	AU 2002334672 A1 07-04-2003
		US 2003092007 A1 15-05-2003	
EP 0969102	A2	05-01-2000	AT 191510 T 15-04-2000
		AT 382094 T 15-01-2008	
		AU 672760 B2 17-10-1996	
		CA 2119557 A1 01-04-1993	
		CZ 9400669 A3 15-12-1994	
		CZ 291877 B6 18-06-2003	
		CZ 298187 B6 18-07-2007	
		DE 69230873 D1 11-05-2000	
		DE 69230873 T2 09-11-2000	
		DE 69233719 T2 02-01-2009	
		DK 0534858 T3 11-09-2000	
		DK 0969102 T3 13-05-2008	
		WO 9306239 A1 01-04-1993	
		EP 0534858 A1 31-03-1993	
		ES 2147550 T3 16-09-2000	
		ES 2299229 T3 16-05-2008	
		FI 941360 A 24-05-1994	
		FI 20031526 A 17-10-2003	
		GR 3033895 T3 30-11-2000	
		HU 68504 A2 28-06-1995	
EP 0969102	A2		IL 103267 A 25-07-2004
		JP 6510668 T 01-12-1994	
		JP 3236295 B2 10-12-2001	
		JP 4088403 B2 21-05-2008	
		JP 2001061486 A 13-03-2001	
		JP 4385072 B2 16-12-2009	
		JP 2008131947 A 12-06-2008	
		NO 941064 A 20-05-1994	
		PT 534858 E 29-09-2000	
		PT 969102 E 25-03-2008	
		RU 2182176 C2 10-05-2002	

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/NL2010/050836

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
		US 6045994 A	04-04-2000
		ZA 9207323 A	30-08-1993

EP 0534858	A1 31-03-1993	AT 191510 T	15-04-2000
		AT 382094 T	15-01-2008
		AU 672760 B2	17-10-1996
		CA 2119557 A1	01-04-1993
		CZ 9400669 A3	15-12-1994
		CZ 291877 B6	18-06-2003
		CZ 298187 B6	18-07-2007
		DE 69230873 D1	11-05-2000
		DE 69230873 T2	09-11-2000
		DE 69233719 T2	02-01-2009
		DK 0534858 T3	11-09-2000
		DK 0969102 T3	13-05-2008
		WO 9306239 A1	01-04-1993
		EP 0969102 A2	05-01-2000
		ES 2147550 T3	16-09-2000
		ES 2299229 T3	16-05-2008
		FI 941360 A	24-05-1994
		FI 20031526 A	17-10-2003
		GR 3033895 T3	30-11-2000
		HU 68504 A2	28-06-1995
		IL 103267 A	25-07-2004
		JP 6510668 T	01-12-1994
		JP 3236295 B2	10-12-2001
		JP 4088403 B2	21-05-2008
		JP 2001061486 A	13-03-2001
		JP 4385072 B2	16-12-2009
		JP 2008131947 A	12-06-2008
		NO 941064 A	20-05-1994
		PT 534858 E	29-09-2000
		PT 969102 E	25-03-2008
		RU 2182176 C2	10-05-2002
		US 6045994 A	04-04-2000
EP 0534858	A1	ZA 9207323 A	30-08-1993
