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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 February 2002 (21.02.2002)

PCT

(10) International Publication Number
WO 02/14546 A1

(51) International Patent Classification[7]: C12Q 1/68

(21) International Application Number: PCT/IB01/00189

(22) International Filing Date: 15 February 2001 (15.02.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PCT/IB00/01127 16 August 2000 (16.08.2000) IB

(71) Applicant and
(72) Inventor: **FRITZSCHE, Markus** [CH/CH]; Gartenstrasse 10, CH-8134 Adliswil (CH).

(74) Agent: **E. BLUM & CO.**; Vorderberg 11, CH-8044 Zürich (CH).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

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

Published:
— *with international search report*

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

WO 02/14546 A1

(54) Title: USE OF MICROBIAL DNA SEQUENCES FOR THE IDENTIFICATION OF HUMAN DISEASES

(57) Abstract: The use of DNA sequences comprising a fragment of a nucleic acid encoding a microbial virulence factor as means for the identification of diseases or a genetic predisposition thereof as well as its use for the development of disease animal models is disclosed.

Use of microbial DNA sequences for the identification of human diseases

Cross Reference to Related Application

This application claims the priority of PCT patent application IB00/01127, filed August 16, 2000, the disclosure of which is incorporated herein by reference in its entirety.

Technical Field

The present invention relates to the use of a DNA sequence comprising a fragment of a nucleic acid encoding a microbial virulence factor as means for the identification of a disease or a genetic predisposition thereof as well as its use for the development of disease animal models.

Background Art

The functional sequences of higher eukaryotes consist of genetic modules of at least two kinds. Modules of coding sequence are combined in many ways to produce proteins, whereas modules of non-coding sequences regulate the expression of genes. Some of the mutations created duplicates of entire genes, which have then evolved new functions, while others altered the expression of old genes by exposing them through gene shuffling to new regulatory sequences. By these means the human genome as a whole has evolved to its present day complexity. Since the 3' untranslated region (3'UTR) and especially the polyadenylation signal within the 3' UTR regulate the translation and expression of a gene, the entire gamut of molecular perturbations can be accomplished with the 3'UTR as a primary target.

There is therefore a need for molecular tools allowing the detection of diseases or a predisposition thereof caused by mutations within the 3' UTR.

Disclosure of the Invention

In the scope of the present invention it was now found that there are microbial DNA insertions in the 3' UTR of human genes which are associated with human diseases.

Hence it is an object of the present invention to provide the use of DNA sequences comprising a fragment of a nucleic acid sequence encoding a putative microbial virulence factor as means for the identification of diseases caused by bacterial mutations or a genetic predisposition thereof. The virulence factor stems preferably from a intracellular microorgansim and is located on a linear or circular chromosome or a plasmid, more preferably said virulence factor stems from a microorganism which is selected from the group consisting of Borrelia species, Chlamydia sp., Escherichia sp., Plasmodium sp. and Rickettsia. Even more preferably said nucleic acid encoding a virulence factor is selected from the group consisting of Seq. Id. No. 1 to Seq. Id. No. 17. Virulence factors stemmming from non-intracellular microorganisms which are part of a cluster shared by intracellular microonganisms are as well suitable for the use in the present invention.

Another object of the present invention is a method for the identification of a disease or a genetic predisposition thereof, which comprises in a tissue or blood sample of a subject or in a fetal neuro-graft a mutation within a nucleic acid sequence selected from the group consisting of Seq. Id. No. 1 to Seq. Id. No. 17 and said sequence is inserted in a gene of said subject. Preferably said sequence is inserted in the 3'UTR of said gene and said mutation is found in the polyadenylation signal of said gene and said mutation preferably affects the expression of the protein encoded by said gene.

Another object of the present invention are transgenic non-human animals, which comprise in their ge-

nome a partial or complete inactive endogenous gene which is selected from the group consisting of cannabinoid receptor 1 gene, MAP 2C gene, apolipoprotein E gene, presenelin 2 gene, integral membrane protein 2B gene, alpha synuclein gene, oligophrenin 1 gene and myotonin protein kinase gene. The gene is inactivated due to at least one mutation in its 3' untranslated region (3' UTR) and said mutation leads to an inhibition or suppression of protein expression. The term mutation as used herein encompasses any nucleotide change, insertion or deletion independent of their length that influences the activity, expression or regulation of a gene.

Although in the context of the present invention any mutation in the 3'UTR region leading to an inhibition or suppression of protein expression e.g. CB1 protein expression, can be used, preferred are mutations located in the sequence following the polyadenylation signal, more preferably in the polyadenylation signal sequence of said gene. Any mutation in the polyadenylation signal sequence leading to an inactivation of said signal can be used. The mutation can e.g.be caused by a sequence of the same or a different microbial species e.g. by gene conversion or recombination.

The polyadenylation signal in eukaryotes has the following conserved sequence: AATAAA.

For the purpose of the present invention any non-human mammal can be used. Preferred are rodents e.g. mice or rats, which preferably harbor a homozygous or heterozygous CB1 gene inactivation in their genome.

Another object of the present invention is the use of the transgenic animals of the present invention for the identification of compounds that have an effect on the activity, expression or regulation of the gene encoded protein. The animals of the present invention allow the identification of compounds which have a direct or indirect effect on CB1 protein, MAP 2C protein, Apolipoprotein E, presenilin 2 protein, integral membrane

protein 2B, alpha synuclein protein, oligophrenin 1 protein and myotonin protein kinase activity, expression or regulation. The compound may consist of a multiplicity of compounds e.g. as obtained from combinatorial chemical libraries. A compound identified by the use of the animals of the present invention can serve as a lead compound for the development of medicaments to treat e.g. schizophrenia or other forms of dementia.

A further object of the present invention is a method for screening for compounds that have an effect on the activity, expression or regulation of proteins selected form the group consisting of CB1 protein, MAP 2C protein, Apolipoprotein E, presenilin 2 protein, integral membrane protein 2B, alpha synuclein protein, oligophrenin 1 protein and myotonin protein kinase. Said method comprises introducing a compound in an transgenic animal of the present invention, preferably in a mouse or a rat, and monitoring e.g. behavioural changes in said animal. Mice and rats are preferred animals as their observable behavioural changes are considered relevant to clinical phenomenology. A transgenic animal, which harbors one of the above mentioned genes in its genome that is inactivated by a mutation in the sequence following the polyadenylation signal or in the polyadenylation signal is preferably used in a screening method of the present invention since such an animal reflects the situation found in human subjects suffering e.g. from schizophrenia, other forms of dementias and other neurological disorders,.

In a further aspect the invention provides a use of a transgenic non-human animal whose genome comprises a disruption of the endogenous CB1 gene for the identification of compounds that have an effect on the activity, expression or regulation of CB1 protein. Any animal comprising a non-functional CB1 gene in its genome is suitable for such a use, preferably a mouse or a rat.

Brief Description of the Drawings

The invention will be better understood and objects other than those set forth above will become apparent when consideration is given to the following detailed description thereof. Such description makes reference to the annexed drawings, wherein:

Figure 1 shows genetic exposure triggering gene conversion and further infectious recombination,

Figure 2a shows genetic exposure leading to multiple translocations into the human genome and

Figure 2b shows genetic exposure leading to multiple translocations into the human genome.

Modes for Carrying Out the Invention

A gene construct for the production of a transgenic animal of the present invention, which comprises in its genome a partially or completely inactivated gene, can be prepared using standard genetic engineering technologies known in the art, such as described in Maniatis et al., Molecular cloning: A Laboratory Manual, Cold Spring Laboratory, Cold Springs Harbor, N.Y. The starting material for said construct can be a portion of e.g. the genomic or cDNA CB1 nucleotide sequence. Introduction of the wanted mutation in e.g. the CB1 sequence can be done by methods known to a person skilled in the art e.g. by site directed mutagenesis. The term mutation as used herein encompasses any nucleotide change, insertion or deletion independent of their length that influence the activity, expression or regulation of a gene.

A transgenic animal in accordance with the present invention can be made using generally known methods in the field. See, for example, Manipulating the Mouse Embryo, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986), Knock-out mouse models used to study neurobiological systems, Critical Reviews in

Neurobiology, 13 (2), 103-149, (1999). Recombinase dependent knockouts can also be generated, e.g. by homologous recombination to insert recombinase target sequences flanking portions of an endogenous gene, such that tissue specific and/or temporal control of inactivation of an allele can be controlled.

The transgenic animals of the present invenion are suitable to identify lead compounds that can serve as a starting point for the development of a medicament to treat e.g. schizophrenia, Alzheimer disease, Parkinson disease, Myopathy or other forms of dementia. Screening for such useful compounds typically involves administering the candidate compound to an animal and monitoring behavioural changes in said animal. Said behavioural changes include for example changes in locomotor activity, stereotypical behaviour, enhanced spatial memory and disruption of working memory. Also biochemical or molecularbiological assays can be used in a screening test for the identification of active compounds. Such tests include for example looking for increased or decreased expression and/or stability and/or activity of CB1 protein or increased or decreased levels and/or activity of messengers such as dynorphin, substance P in various tissues of the animals at different time points. At the end of the test the animals are preferably sacrificed to determine for example the effect of the administerd test compound on the expression level and/or expression sites of the CB1 protein in the brain by for example in situ techniques.

There are at least three different categories of compounds that can be screened by a screening test of the present invention: chemical libraries, natural product libraries and combinatorial libraries. Chemical libries consist of structural analogs of known compounds. Natural product libraries are collections of microorganism, animals, plants or marine organisms which are used to create mixtures for screening by for example fermenta-

tion and extraction of broths from soil, plant or marine microorganisms or extraction of plants or marine organisms. Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or organic compounds as a mixture. They are relatively easy to prepare by traditional synthesis methods, PCR or cloning.

In the above described screening systems not only transgenic animals of the present invention but any transgenic animal comprising a non-functional gene encoding a protein selected from the group consisting of CB1 protein, MAP 2C protein, Apolipoprotein E, presenilin 2 protein, integral membrane protein 2B, alpha synuclein protein, oligophrenin 1 protein and myotonin protein kinase.in its genome can be used. The non-functional gene can preferably be produced by introduction of at least one mutation in the coding-sequence and/or 5' or 3'regulatory sequences or in an intron of the gene.

Microbial DNA sequence insertions found in several genes can be used to identify prenatal infectious diseases, bacterial mutations and other diseases, preferably human diseases, e.g. dementias or a predisposition thereof. Genes with such insertions are e.g. genes encoding the cannabinoid receptor, the microtubule associated protein 2C, apolipoprotein B, presenilin 2, amyloid precursor protein, integral membrane protein 2B, notch 3, microtubule associated protein tau, alpha synuclein and myotonin protein kinase. Said genes are known to be associated with the development or a predisposition for diseases like schizophrenia, Parkinson disease, Alzheimer disease, myopathy, frontotemporal lobe dementia, hereditary multi-infarct dementia, autosomal dominant Parkinson Lewy-Body dementia, familial British dementia and primary X-linked mental retardation. Said microbial DNA sequences are as well suitable for the detection of a predisposition or a genetic variation for diseases, the pathological manifestation of which is triggerd by medicaments or drugs like e.g. cannabis. For the above mentioned use se-

quences encoding microbial virulence factors and harboring at least one mutation affecting expression, activity or regulation of proteins encoded by said genes are preferred.

As already mentioned above the present invention also concerns methods for the identification of a disease caused by bacterial mutations or a genetic predisposition thereof. Said methods comprise detecting the presence in a tissue-or blood sample of a subject a mutation within a nucleic acid sequence selected from the group consisting of Seq. Id. No.1 to Seq. Id. No. 17 and said sequence is part of a gene of said subject. The method of the present invention is e.g. suitable for the identification of one of the above mentioned diseases.

Suitable methods for the identification of said diseases or a genetic predisposition thereof in humans are for example PCR techniques, DNA or gene chips, hybridisation techniques and Ligases chain reaction (LCR).

In a preferred embodiment said method comprises the following steps:

1. Blood or Tissue sample
2. DNA extraction
3. Amplification of at least one Sequence selected from the group consisting of Seq. Id. No. 1 to 17 using flanking oligonucleotide primers
4. Analysis of amplification products by se quencing

The sequencing step allows the identification of mutations present in the insertion sequence of interest. The design of the oligonucleotide primers is known to those skilled in the art and can be done using standard software.

Examples

Example 1

Microbial sequences in 3'UTR polyadenylation regions of schizophrenia and dementia genes harbour microbial virulence factors and plasmids

DNA from *Borrelia burgdorferi*, Chlamydia and other intracellular microbes has inserted into ancestral 3' polyadenylation sites of the following candidate neurological genes: The central cannabinoid receptor gene (CB1), Alzheimer disease and other dementia genes, Parkinson disease and myotonic dystrophy (see Seq. Id. No. 1 to 17 and Table 1A to 1E). Most insertions originate from microbial virulence factors, transposable elements and plasmids.

CB1 and 5HT1E contain related insertions from *B.burgdorferi*. The human CB1 gene is located at 6q14, which has been reported as a candidate region for schizophrenia involving a translocation break-point co-segregating with schizophrenia, immediately adjacent to the 5-hydroxytryptamine (5HT1E) gene. The spirochaetal insertion into 5HT1E originated from a *B.burgdorferi* virulence factor, the flagellar basal-body rod protein (fbrp), responsible for chemotaxis, locomotion and a syringe mechanism for injection into cells. Through infectious recombination, *B.burgdorferi* introduced another nucleotide sequence containing p115 and a polyadenylation site into our ancestral genome. As a result of this recombination, p115 overlaps with fbrp. The fbrp section on the CB1 gene originates from an ancient 5HT1E receptor already containing the spirochaetal insertion and not from a direct transposition of *B.burgdorferi* onto 6q14, because the first three nucleotides of the overlapping sequence (see Table 1A) are on 5HT1E, and not on fbrp of *B. burgdorferi*. Gene conversion is the most likely explanation since CB1 and human 5HT1E, on which fbrp is absent, are both located in tandem on 6q14.

The spirochaetal recombination with p115 then led to the introduction of the poly-A signal (AATAAA) which codes for the translation and genetic expression of CB1. Being located on an important 3' regulatory code, whose mutational rate runs slower than the molecular clock of silent point mutations in protein coding regions, microbial sequences are thus better protected from mutations and hence stabilised.

B. burgdorferi and *Chlamydia pneumoniae* have been detected in post mortem Alzheimer brains. Within the polyadenylation signal of the candidate gene apolipoprotein E (Apo E) epsilon 4 allele, which is a risk factor for schizophrenia, Alzheimer's disease, and chlamydial infection, the possible molecular basis for this - an overlapping sequence of microbial virulence factors originating from the linear chromosome of *B.burgdorferi* (including two copies on plasmids) and the human pathogenic species *Chlamydia pneumoniae* (see Table 1B) were identified.

The origin of new traits through lateral microbial gene transfer into mammals is still in its infancy and counter-current to the established views of understanding evolution through mutation and selection of existing sequences. Instead, multiple infectious recombinations and genetic exposure of preinserted templates has led to a further spread of Borrelia DNA all over the human genome. Intriguingly, by a single frame-shift mutation within the original translocase, genomic decay has apparently incapacitated *B. burgdorferi* to translocate further genetic templates into its hosts. Resulting from the same process of genomic decay in plasmids, or alternatively, from genetically induced variability by virulence factor operons, osp A and B, nucleotide dissimilarity between *B. burgdorferi* and its pre-inserted human templates have occurred. In the case of further infectious recombination with Osp A, this might trigger mismatch repair mutations within the poly-A signal and a ge-

netic knock-out of CB1. The transition of the parasite from arthropod vector to human host is accompanied by significant changes in gene expression, which has practical relevance for prevention, as well. The recently approved vaccine against Lyme disease consists of an immunogenic Osp A, that is expressed by the spirochaetal parasite while in the tick gut, turned off on entry into humans and then expressed again at a late chronic stage. After attachment of an infected tick and initiation of a blood meal, anti-OspA antibodies enter the tick gut and mediate killing of *B. burgdorferi*.

Example 2

Lateral gene transfer of B.burgdorferi into 5HT1E, gene conversion and homologous recombination with CB1

On the human CB1 gene, located within a candidate region for schizophrenia at 6q14 immediately adjacent to the 5HT1E (5-hydroxytryptamine) gene, a nucleotide sequence of (p115) originating from *B.burgdorferi* could be identified (Seq. Id. No. 1). Another ancestral spirochaetal inclusion on 5HT1E originates from a *B.burgdorferi* virulence factor, the flagellar basal-body rod protein (fbrp). It can still be found within the serotonin receptor 1E gene (5HTIE) on the mouse (Mus musculus) and rat (Rattus norvegicus). During phylogeny, multiple recombinations between the spirochaetal fbrp and its pre-inserted fbrp templates have exposed the identical sequences on the complementary strand on the double helix, including adjacent non-microbial nucleotides, to further recombination all over the human genome and a clustering of microbial virulence factors from Chlamydia muridarum iron binding protein, Plasmodium falciparum rhoptry and Staphylococcus aureus penicilllin binding protein, a specific gene conversion occured (see Fig. 1 and Fig. 2). Through infectious recombination,

B.burgdorferi subsequently introduced p115 into our ancestral genome resulting in a genetic overlap with fbrp. Containing the poly-A signal (AATAAA), p115 thus introduced the code for the translation and actual genetic expression of CB1.

Example 3

No point mutations within the polyadenylation signal (AATAAA)

With the introduction and natural selection of the inserted polyadenylation site, lateral gene transfer have effectively influenced the genetic expression of CB1. Presenting save havens for microbial DNA there are no point mutations on this important signal, for a change in the signal would disrupt the genetic expression of CB1. Suppose the mutually advantageous sequences on the polyadenylation signals recombine again with slightly dissimilar spirochaetal strands, the subsequent mismatch repair mutations will be deleterious for both hosts and parasites. Being located on an important 3' regulatory code, whose mutational rate runs slower than the molecular clock of silent point mutations in protein coding regions, microbial sequences are thus better protected from mutations and hence stabilised.

Example 4

DNA of B.burgdorferi underlies the schizophrenic genotype

That lateral gene transfer has influenced the evolution of higher eukaryotes, such as mammals is counter-current to established views. *B. burgdorferi* appears to be excluded from the benefits of the extensive lateral gene transfer between micro-organisms, however, as an intracellular parasite with an incomplete genome *B. burgdorferi* has a direct access to host genes, which it

exploits for replication, and not to be recognised as foreign, the spirochete depends on its own sequences within the human genome. Through molecular mimicry of fbpr antigens within the host's CD45 leukocyte defence system and a structure homologous to nucleoprotein, *B. burgdorferi* might, for example, dispose of a protective shield at the DNA and protein level, respectively.

Dissimilarity and mismatch mutations between *B. burgdorferi* and its pre-inserted human templates may nevertheless occur, resulting from genetically induced variability by virulence factor operons, osp A and B, or, alternatively, from a genomic decay in plasmids (Casjens et al., Mol. Microbiol. 2000, 35, 490-516) borrelia re-infection causing putative mutations in AATAAA. Intriguingly, by a single frame-shift mutation within the original translocase, the same process of genomic decay has apparently incapacitated *B. burgdorferi* to translocate further genetic templates into its hosts (Casjens et al., Mol. Microbiol. 2000, 35, 490-516). Whereas point-mutations within the poly-A signal would impair the genetic expression, a change in the adenine content of polyadenylation tail (Alberts et al., Molecular Biology of the Cell, 1994, Garland Publishing), would alternatively enhance or reduce the ribosomal translation of CB1. Several mutations at an early blastular stage can lead to different (chimeral) expressions of CB1 (and perhaps other genes) and account for the reported continuum of major psychoses between schizophrenia and bipolar manic-depression.

Dissimilarity and infectious recombination between protein coding sequences of the human chromosome associated protein hCAP and borrelia P115 (63% identities with P115) could, on the other hand, account for the altered leukocyte chromatin ultra-structure reported in schizophrenic patients.

Genetic epidemiology which has provided consistent evidence over many years that schizophrenia has a

genetic component and that this genetic component is complex and polygenic (Riley and McGuffin, Am J Med Genet 2000, 97, 23-44) does not challenge the main tenet of the present disclosure. The many reported coincident regions for schizophrenia and the sequences of B. burgdorferi within the human genome just reflect two different aspects of the same phenomenon - infectious recombination (see Fig. 1 and Fig. 2). Figure 1 shows that multiple recombinations between the spirochaetal fbrp and its pre-inserted fbrp template on ancestral 5HT1E exposed the complementary strand on the double helix, including adjacent non-microbial nucleotides, to further recombination with both microbial and ancestral DNA. This has led to a gene conversion from 5HT1E onto CB1, which are both located adjacent to each other on the 6q14 candidate region for schizophrenia. Since the first three nucleotides (att) can still be found on 5HT1E of the mouse (Mus musculus) and rat (Rattus norvegicus), but not on fbrp of Borrelia burgdorferi, the spirochaetal template on the CB1 gene originates from ancient 5HT1E already containing the spirochaetal inclusion, and not from a direct transposition of B. burgdorferi onto 6q14. Observe the 'loophole' mutation, inserting an additional adenine into CB1. Homologous recombination and mismatch-repair mutation between B. burgdorferi p115 and the borrelia template fbrp subsequently introduced the polyadenylation signal AATAAA, which now encodes the ribosomal translation and genetic expression of CB1.

Figures 2a and 2b show that infectious recombinations between B. burgdorferi fbrp and its pre-inserted fbrp template on ancestral 5HT1E repeatedly exposed the complementary strand on the double helix, including adjacent non-microbial nucleotides, to further recombinations with ancestral DNA. This has occurred before and after the point mutation from adenine to thymin A-T within the pre-inserted fbrp on 5HT1E. In comparison to the adjacent non-exposed control sequences of the same

length (32 base pairs), genetic exposure has thus given rise to a high number of translocations within the human genome. Most translocations are, in addition to chromosome X, located on 6q reflecting their original spread from ancestral 5HT1E on nearby 6q14. Multiple translocations into 6q21, 6p21 and Xq28 also point to candidate regions of major psychoses, that correlate with the highest lod-score for schizophrenia at 6q21, the candidate region for schizophrenia at 6p22 and the candidate region for bipolar manic depression at Xq28.

The phylogenetic trace implies that such recombinations have occured over and over again, and that novel mutations within pre-inserted spirochaetal templates are likely to occur. The excess of winter-spring births in sporadic schizophrenia, which follows the spring-autumn season of tick born infection, and its geographical overlap of the tick vector with areas of increased risk for sporadic schizophrenia (Brown, Schizophr Bull 1994, 20, 755-75) suggests *B. burgdorferi* as the prime infectious candidate. Only a few other microorganisms express nucleotide sequences for transplacental penetration, and fetal infection, which are homologous to those of their host and thereby not viewed as foreign. With only an incomplete genome at its disposal, *B. burgdorferi* exploits and therefore interferes with the genetic machinery of its host cell for replication (Fraser et al., Nature 1997, 390, 580-86). If in analogy to the neurotropic murine leukemia virus, *B. burgdorferi* infects oligo-cellular blastulas in early pregnancy, a horizontal transfection of germ line cells would lead to a vertical transmission of the disease. This, in fact, occurs during the transmission of *B. burgdorferi* from the adult ixoid tick into eggs and larvae, which can thus become infectious without prior blood meal, and sporadic prenatal *B. burgdorferi* infection have indeed been reported in human beings (Steere in Principles and Practice of Infectious Diseases. Philadelphia: Churchill Living-

stone, 2000, 2504-2518). By adding genetically vulnerable cases to the population human germ-line transfection by B. burgdorferi would thus explain the continued presence of schizophrenia at high prevalence, despite the fact that the disease confers reduced procreational fitness and fertility.

The association between level of cannabis consumption and development of schizophrenia during a 15-year follow-up was studied in a cohort of 45,570 Swedish conscripts. The relative risk for schizophrenia among high consumers of cannabis (use on more than fifty occasions) was 6.0 (95% confidence interval 4.0-8.9) compared with non-users. Persistence of the association after allowance for other psychiatric illness and social background indicated that cannabis is an independent risk factor for schizophrenia (Andreasson et al., Lancet 1987 Dec 26; 2 (8574:1483-6).

Example 5

Significant CB1-D1 protein homology versus CB1-D2 nucleotide homology

There are significant homologies, which cannot be reduced to lateral gene transfer. CB1 shows significant homology with the dopamine D1 - a type I G-protein coupled receptor - at the amino-acid level (Score = 71.0 bits (171), Expect = 2e-11, Identities = 75/320 (23%), Positives = 137/320 (42%), Gaps = 50/320 (15%) (Online Mendelian Inheritance in Man, OMIM 2000, Center for Medical Genetics, John Hopkins University (Baltimore, MD) and National Center for Biotechnology Information, (Bethesda, National Library of Medicine, 2000). Other homologous proteins comprise G-protein coupled receptors that receive information from outside the cell, such as olfaction or vision. However, owing to one specific sequence of 19 homologous base pairs at the nucleotide level, it is the dopamine 2 receptor (D2) - a

type II G-protein coupled receptor - which shows significant DNA homology to CB1 (Score = 37.2 bits (19), Identities = 19/19 (100%). The circumscribed homology between CB1 and the D2 receptor, which is an important pharmacological target for the treatment of schizophrenia, encodes the seventh transmembrane loop which is known for its inhibitory-mode (i-mode) of metabotropic action. Apart from a silent point mutation from GTG to GTC in the rat D2 receptor gene (Rattus norvegicus), which has occurred after the phylogenetic rat mouse divergence 35 million years ago, this homologous nucleotide sequence can be found an all sequenced primate (Macaca mulatta, Cercopithecus aethiops, H.sapiens; OMIM, 2000) and rodent (Mus musculus) D2 receptor genes.

An early prenatal event interfering with neuronal migration from inner to outer cortical laminae in mid pregnancy most likely underlies the consistent pattern of cellular disarray observed in schizophrenic brains. If the expression of higher levels of CB1 within outer cortical compared to inter-cortical layers (Glass et al., 1997, Neuroscience, 77, 299-318) resulted from a CB1 mediated migration of neurons, the inter-cortical disarray could be explained by a knock-out of CB1, whose metabotropic i-mode has recently been reported to be crucial for cellular migration (Song and Zhong, Journal of Pharmacology and Experimental Therapeutics, 2000, 294, 204-209). The distribution of CB1, furthermore, exactly mirrors the macro-anatomic regions mainly affected in schizophrenia (Schultz and Andreasen, 1999, Lancet, 353, 1425-30).

Through a dysinhibition of the i-mode, spatial memories (hippocampal long-term potentiation) of CB1 knock-out mice are enhanced (Bohme et al., 2000, Neuroscience, 95, 5-7) and goal directed, temporal memories decreased. This mnemonic effect, however, not only parallels the pattern of spatio-temporal distortions in schizophrenia. Despite the fact that patients with Alz-

heimer's dementia and tertiary neurosyphilis do hallucinate, they apparently do not remember, or reconnect their hallucinations with a fixed delusion. The difference between dementia praecox (schizophrenia) and syphilitic dementia is that without memory hallucinations are lost, and that the thoughts of schizophrenics are flooded with fixed hallucinations, expanding into overt delusions. Furthermore, CB1 knock-out mice do show reduced exploratory, goal-directed behaviours(Steiner et al. Proc. Natl. Acad. Sci. USA, 2000, 96, 5786-5790); symptoms that appear to be among the most robust indices in schizophrenia. Without time-bridging working memory, no creative speech and no logical thinking would be possible. A disruption of CB1, which reaches its highest levels in the left-hemispheric area of Wernicke, would thus account for the impairment of goal directed behaviour and speech - the highest form of sequential behaviour in man - being more stereotyped in schizophrenic patients than in healthy persons.

Example 6

The phylogenetic traces of microbial insertions into the human genome were investigated and the genetic mechanism were analysed, by which bacterial virulence factors and mobile elements from intracellular parasites could disrupt candidate genes for schizophrenia and dementia. This was done by co-incident DNA homology BLAST searches between neurotropic microorganisms, the central cannabinoid receptor CB1, and other known dementia genes, whose complete sequences with 3'UTR and polyadenylation signals are entered on Gene-Bank databases. Several such genes have now been characterised including those of ApoE4 Alzheimer Disease type II (AD2) and presenilin 2 (AD 4). Positional cloning and sequencing has also been carried out in other dementias including Familial British Dementia (with mutations in the gene for integral membra-

ne protein 2A), Hereditary Multi-infarct Dementia (Notch-3 gene), primary X-linked mental retardation (oligophrenin 1 gene), Frontotemporal Lobe Dementia, Autosomal Dominant Parkinson Lewy-Body Dementia and Familial Parkinson disease type I (with mutations in the alpha synuclein gene). The genomic facilities used are accessible at Online Mendelian Inheritance of Man (OMIM) and very much recommended to anybody interested in applied medical genetics.

While there are shown and described presently preferred embodiments of the invention, it is to be distinctly understood that the invention is not limited thereto but may be otherwise variously embodied and practiced within the scope of the following claims.

Genes & Events	Organism	Sequence	Seq. Id. No.

Table 1A

SCHIZOPHRENIA

CB1

1st & 2nd borrelia integrations into ancestral 5HT-CB1 and introduction of poly-A signal

	Borrelia (fbrp) (357)	CATTTCTTCAACTAAATTAACATT(334)5′	18
5HT1E/		AttCATTTCTTCAACTAAATTAACATT	19
CB1(anc.1)		AttCATTTCTTCAACTAAATTAACATT	
6q14	Borrelia (P115)(4133)	GATTCAAATAAAAATTCTAAATTACCAT(4160)3′	1

Chlamydia integration into ancestral rodent CB1

	Chlamydia (8221)	TTACCTGGACTCAAATAAAAGT(8242)3′	20
CB1(anc.2)	(ibp)	GAATCAAATAAAAATTCTAGATTACCATgaagaacata	22
CB1(rat)	(5430)	TTACCTGGAATCAAATAAAAGTTCTAGATTATCACg (5465)3′	21

Chlamydia muridarum integration into ancestral primate CB1

	Chlamydia (8219)	AGTTACCTGGACTCAAATAAAA(8240)3′	2
CB1(anc.2)	(ibp)	GAATCAAATAAAAATTCTAGATTACCATgaagaacata	22
CB1(human)/6q14	(5484)	AGTTACCTGGAATCAAATAAAAATTCTAGATTACCATgaagaacata(5530)	47

	Chlamydia	(1063)AGTTTAATAAAGATT (1049)5′	4
Apo E (anc.)		GTTTAATAAAAATT	50
Apo E		(4611) AGTTTAATAAAGATTca(4627)3′	27

borrelia infection - predicted mutations by virulence factors

	Borrelia (acrB)	(8106) TTTAAAAAAGATTCA (8120)3′	28
	Borrelia (comp.loc)	(11560) TTTAAAAAAGATTCA (11546)5′	28

PRESENILIN 2 (PS2) - Alzheimer's disease type 4 (AD4)

Borrelia burgdorferi (sgp) stability governing protein integrated within presenilin 2 and possible predicted mutation by Borrelia reinfection

	Borrelia (sgp)	(10599)ATACTAATATCAATAA (10614)3′	5
PS 2		(1651) ATACTAATATCAATAAa(1667)3′	29
	Borrelia (sgp) reinfection	(10599)ATACTAATATCAATAAT(10615)3′	30

AMYLOID PRECURSOR PROTEIN (APP) Alzheimer's disease type 1 (protease nexin-II)

No Borrelia but has Plasmodium falciparum inclusion and Salmonella typhimurium transposable plasmid

APP		(3547) TTTTCATGTAAATAAATACATTCT(3570)3′	31
	Salmonella plasmid traJ	(425) TGTAAATAAATACATTCT(442)3′	48
	Plasmodium chromosome 3	(79642) TTTTCATGTAAATAAATA (79625)5′	6

	(hypothetical protein)	(141360)	TTTCAGGTAAATAAATA	(141344)5′	32

FAMILIAL BRITISH DEMENTIA (integral membrane protein 2B gene) Borrelia burgdorferi linear plasmids lp25, lp36 and Plasmodium falciparum major merozoite surface & receptor binding protein

FBD		(1753)	GATTTTTTCTTTAAATAAAAATAAGT	(1778)3′	33
	Borrelia lp25 plasmid	(18048)	TTTAAATAAAAATAAGT	(18032)5′	7
	Borrelia lp 36 plasmid	(9312)	TTTAAATAAAAATAAG	(9327)3′	51
	Plasmodium (mmsp)	(1083)	TTTTTTTTTTAAATAAAAATA	(1103)3′	34
	Plasmodium (pfemp1)	(4682)	GATTTTTTCTTTAGATAAAAATAAG	(4658)5′	8

Oligophrenin 1 (OPHN1)

Borrelia burgdorferi inclusion from tryptophanyl-t-RNA synthetase

OPHN 1	(6648)	CAAATAAAGTAGTAAAAGA	(6666) 3‘	56
Borrelia (trsa)	(101)	CAAATAAAGTAGTAAAAGA	(83) 5‘	17

HEREDITARY MULTI-INFARCT DEMENTIA (NOTCH 3 gene)

No Borrelia or plasmodium, but has Vibrio cholera virulenence factor and Clostridium insertion

DHMI		(8048)	CCTAATAAAGGAATAGTTAAC	(8068)3'	35
	Vibrio (ntno)	(4641)	AATAAAGGAATAGTTAA	(4657)3'	52
	Clostridium (ctfA)	(2544)	CCTAATAAAGGAATAG	(2559)3'	53

FRONTOTEMPORAL LOBE DEMENTIA (gene for microtubule associated protein tau)

No Borrelia or plasmodium, but has Staphylococcus aureus antibiotic resistance plasmid

MAP Tau	(2276)GCTAGTAATAAAATAT(2291)3′		36
Staph. Plasmid pS194	(2276)GCTAGTAATAAAATAT(2291)3′		36

PARKINSON DISEASE

AUTOSOMAL DOMINANT LEWY BODY (PDLBD) (alpha synuclein gene)

Borrelia burgdorferi inclusion from linear plasmid lp36

PDLBD

2nd poly A

signal	(1521) ACAATAAATAATATTC(1536)3′		37
Borrelia lp36	(13825) ACAATAAATAATATTC(13810)5′		9

PARKINSON DISEASE, FAMILIAL TYPE I (PD1) (alpha synuclein gene)

Borrelia burgdorferi linear plasmid lp17, Plasmodium falciparum plasmid and Escherichia coli pilus protein

PD - 1

3rd poly A signal	(952) TAATAATAAAAATCATGCTT(971)3'		49
Borrelia lp17 plasmid	(9683) TAATAATAAAAATCAT (9668)5'		10
Plasmodium plasmid	(10913) AATAATAAAAATCATG (10928)3'		12
Escherichia (pilus protein)	(860) AATAATAAAAATCATGCTT(878)3'		11

Table 1C

MYOPATHY

MYOTONIC DYSTROPHY, 3' UTR & TRIPLET REPEAT

Borrelia burgdorferi and Chlamydia muridarum integration into human myotonin protein kinase (Mt-PK)

	Borrelia (fbrp)	(194)	TCCGGAATAAAAGGCCCT(177)5'	13
	(Mt-PK, anc.)		TCCGGAATAAAAGGCCCT	38
5	(Mt-PK)	(2462)	TCGCGAATAAAAGGCCCT(2479)3'	39
	Chlamydia	(3551)	GCGAATAAAAGGCCCT(3536)5'	14

Table 1 D

GLOBIN GENES - CONTROLS FOR MICROBIAL INSERTIONS

10 β-globin (chromosome 11)

β-globin chain versus β-thalassaemia

Poly-A cleavage site (1736 - 1737)

Normal β-globin	(1711)	AATAAAAAACATTTATtttcattgca	atgATGTATTTAAATta(1753)...	40
Rickettsia	(232700)	AATAAAAAACATTTAT(232685)5'		15
15 Borrelia (oppAIV)	(1143)		ATGTATTTAAAT (1132) 5'	41

normal β- globin	(1754)	tTTCTGAATATTTTACTAAAAA(1775)3'		42
Borrelia (ospC)	(364)	TTCTGAAGATTTTACTAAAAA(384)3'		43
β-Thalassaemia	(1748)	AATAAGAAACATTTATTTtcattgca	(1773)3'	44
Plasmodium falciparum	(470)	ATAAGAAACATTTATTT	(486)3'	45

α-globin (chromosome 16)

α-globin chain (NM_000558) in lower case letters versus Plasmodium bergei (L21708) within 3' α-globin in upper case letters 46

α-1 and α-2

```
  1 actcttctgg tccccacaga ctcagagaga acccaccatg GTGCTGTCTC CTGCCGACAA
 61 GACCAACGTC AAGGCCGCCT GGGGTAAGGT CGGCGCGCAC GCTGGCGAGT ATGGTGCGGA
121 GGCCCTGGAG AGGATGTTCC TGTCCTTCCC CACCACCAAG ACCTACTTCC CGCACTTCGA
181 CCTGAGCCAC GGCTCTGCCC AGGTTAAGGG CCACGGCAAG AAGGTGGCCG ACGCCCTGAC
241 CAACGCCGTG GCGCACGTGG ACGACATGCC CAACGCGCTG TCCGCCCTGA GCGACCTGCA
301 CGCGCACAAG CTTCGGGTGG ACCCGGTCAA CTTCAAGCTC CTAAGCCACT GCCTGCTGGT
361 GACCCTGGCC GCCCACCTCC CCGCCGAGTT CACCCCTGCG GTGCACGCCT CCCTGGACAA
421 GTTCCTGGCT TCTGTGAGCA CCGTGCTGAC CTCCAAATAC CGTTAAGCtg gagcctcggt
481 ggccatgctt cttgcccctt gggcctcccc ccagcccctc ctcccccttcc tgcacccgta
541 ccccccgTGGT CTTTGAATAA AGTCTGAGTg ggcggc 3'
```

Table 1E

VIRULENCE FACTORS AND OTHER MICROBIAL INSERTIONS

Borrelia burgdorferi (p115): related to human chromosome associated protein responsible for DNA and intracellular movement.

Borrelia b. (fbrp): flagellar basal rod protein for extracellular movement,chemotaxis, syringe mechanism for cell injection.

Borrelia b. (ospA): outer surface protein A, antigen of genetically induced variation.

Borrelia b. (plasmid lp28-1): pseudogene on linear plasmid size 28, group 1, being the result of genomic decay.

Borrelia b. (lip P):lipoprotein P (homologous sequences on plasmids, i.e. lp38), anti-genetic surface protein.

Borrelia b. (acrB):acriflavine resistance protein.

Borrelia b. (comp.loc):competence locus with multiple homologous copies throughout genome.

Borrelia b. (sgp): stability governing protein for stabilisation of membrane.

Borrelia b. (oppAIV): oligopeptide permease; immediately adjacent to 3'poly-A cleavage site of (globin.

Borrelia b. (ospC): outer surface protein C; 18 base pair distant from 3' poly-A cleavage site of (globin. Target of Borrelia vaccine.

Borrelia b. (pus): pseudo uridylate synthase

Borrelia b. (slipp): surface lipoprotein p27

Borrelia b. (trsA): tryptophanyl t-RNA synthetase

Borrelia garini (ospA): outer surface protein A. Chlamydia muridarum (ibp) within CB1: iron binding protein to overcome host barriers of low iron levels.

Chlamydia m. within Mt-Pk: phosphocarrier protein.

Chlamydia pneumonae: hypothetical protein.

Clostridium beijerinckii (ctfA):small subunit of coenzyme A transferase.

Escherichia coli pilus protein: responsible for cellular adherence and infection.

Plasmodium bergei within (globin series: phosphoprotein mRNA.

Plasmodium falciparum within CB1: rhoptry associated protein (264)ATCAAATAAAAGTTCTA(280)3' for erythrocyte penetration.

Plasmodium f. within amyloid precursor protein: several sequences form chromosome 3 including hypothetical proteins.

Plasmodium f. (mmsp):major merozoite surface protein, expressed during sexual stage.

Plasmodium f. (pfemp1): Plasmodium falciparum-encoded protein on the surface of infected erythrocytes mediates receptor binding.

Plasmodium f. in thalassaemia (globin:RNA polymerase III.

Rickettsia prowazekii in (globin: proline-betaine transporter for the reduction of osmotic stress in host environment.

Vibrio cholerae (ntno): Na+-translocating NADH-ubiquinone oxidoreductase enzyme complex involved in flagella rotation.

Microbial virulence factors within the 3' genetic hotspot of human disease. Base pairs originating form the lateral gene transfer of microbial nucleotides are indicated in upper case, and non-microbial nucleotides in lower case letters. There are normally no point mutations on the polyadenylation signal, whose non-redundant code protects the microbial inclusions from mutations. Homologous recombination between B. burgdorferi p115 and the borrelia template fbrp originating from ancestral 5HT1E has introduced the polyadenylation signal AATAAA into CB1. Note that recombinational mismatch-repair has inserted an additional adenine into CB1. A point mutation from A to G has through a reduction of the adenine content led to a shortening of the 3' polyadenylation tail of the rat. Within human CB1 and rat CB1, almost identical, but independent insertions of C. muridarium nucleotides must have occurred twice. This reoccurrence emphasises the attraction CB1 and its polyadenylation signal exerts on microbial DNA to recombine. In the case of dissimilarity (i.e. AATATA) mismatch repair mutations might, in analogy to the genetic knock-out of globin in thalassaemia, result in a knock out of CB1 and other neurological candidate genes.

3rd borrelia infection - predicted mutations by virulence factor or plasmid of human CB1 poly-A signal

Borrelia burgdorferi (ospA)	(41)	ATAATAATTCTAAATTA	(25)5′	24
Borrelia garinii (ospA)	(229)	ATAATAATTCTAAATTA	(213)5′	23
Borrelia (plasmids;i.e.1p28-1)	(7553)	AATATAAATTCTATAT	(7568)3′	25

Microtubule associated protein 2C (MAP2C)

Borrelia burgdorferi inclusion from pseudouridylate synthetase

MAP2C	(3634)	AAACTCAGAAAATAAAATGT	3‘ (3653)	54
Borrelia (pus)	(2645)	AAACTCAGAAAATAAAATGT	5‘ (2626)	16
Borrelia (slipp) reinfection	(40665)	CAGAAAATAAAAT	5‘ (40653)	55

Table 1B

DEMENTIA

APOLIPOPROTEIN E (Apo E) Alzheimer's diease type 2 (AD2)

Borrelia burgdorferi surface protein integrated within ancestral ApoEe4

Borrelia (lipP)	(11)	GTTTAATAAAAATT	(24) 3'	3
Apo E (anc.)		GTTTAATAAAAATT		26

Chlamydia pneumoniae integration into ancestral ApoEe4

Claims

1. Use of a DNA sequence comprising a fragment of a nucleic acid encoding a putative microbial virulence factor as means for the identification of a disease caused by mutations or a genetic predisposition thereof.

2. Use of claim 1 wherein said virulence factor is located on a linear or cirular chromosome or a plasmid.

3. Use of claim 1 or 2 wherein said virulence factor stems from a intracellular microorganism.

4. Use of claim 1 or 2 wherein said virulence factor stems from a non-intracellular pathogen and is part of a cluster shared by intracellular microorganisms.

5. Use of claim 1 wherein said microorganism is selected from the group consisting of Borrelia species, Chlamydia species, Escherichia sp., Plasmodium species and Rickettsia species.

6. Use of anyone of claims 1 to 5 wherein said fragment is selected from the group consisting of Seq. Id. No. 1 to Seq. Id. No. 17.

7. Use of anyone of claims 1 to 4 wherein said sequence comprises a mutation, either caused by by the same or a different species, preferably within the polyadenylation signal sequence.

8. Use of anyone of claims 1 to 7 wherein said disease is a human disease.

9. Use of claim 8 wherein said human disease is selected from the group consisting of schizophrenia, Alzheimer disease, Parkinson disease, Myopathy and other forms of dementias.

10. Use of claim 8 wherein said human disease constitutes a predisposition or a genetic variation, the pathological manifestation of which is triggered by medicaments or drugs.

11. Use of claim 10 wherein said drug is cannabis.

12. Use of claim 11 wherein said pathological manifestation comprises any form of dementia, schizophrenia, or related psychatric disorders.

13. A method for the identification of a disease or a genetic predisposition thereof, which comprises detecting the presence in a tissue-or blood sample of a subject a mutation within a nucleic acid sequence selected from the group consisting of Seq. Id. No.1 to Seq. Id. No. 17 and said sequence is part of a gene of said subject.

14. The method of claim 13 wherein said tissue sample is a foetal graft for neurotransplantation.

15. The method according to claim 13 or 14, wherein said sequence is inserted in the 3'UTR of said gene.

16. The method according to anyone of claims 13 to 15, wherein said mutation is found in the polyadenylation signal of said gene.

17. The method according to anyone of claims 13 to 16, wherein said mutation affects the expression of the protein encoded by said gene.

18. The method according to anyone of claims 13 to 17, wherein said gene is selected from the group consisting of Cannabinoid receptor 1 gene, MAP 2C gene, apolipoprotein E gene, presenilin 2 gene, integral membrane protein 2B gene, , alpha synuclein gene, oligophrenin 1 gene and myotonin protein kinase gene.

19. A transgenic non human animal whose genome comprises a partially or completely inactivated endogenous gene as defined in claim 18, wherein said inactivation is due to at least one mutation in its 3' untranslated region, said mutation leading to inhibition or suppression of the subsequent gene translation.

20. The transgenic non-human animal of claim 19, wherein the mutation is located in the nucleic acid

sequence following the polyadenylation signal, more preferably in the polyadenylation sequence of said gene.

21. The transgenic non-human animal of claim 20, wherein said mutation is a point mutation.

22. The transgenic non-human animal according to anyone of claims 19 to 21, wherein said animal is a mammal, in particular a rodent.

23. The transgenic non-human animal of claim 22, wherein said animal is a mouse or a rat.

24. The transgenic non-human animal according to anyone of claims 19 to 23, wherein said inactivation is a homozygous or a heterozygous inactivation.

25. Use of a transgenic non-human animal according to anyone of claims 19 to 24 for the identification of compounds that have an effect on the activity, expression or regulation of the translated protein.

26. A method of screening compounds that have an effect on the activity, expression or regulation of a protein encoded by a gene according to claim 18 comprising introducing a compound in an animal according to anyone of claims 19 to 24 and monitoring behavioural changes in said animal as compared to a control animal.

27. Use of a transgenic non-human animal whose genome comprises a non-functional endogenous CB1 gene for the identification of compounds that have an effect on the activity, expression or regulation of CB1 protein.

28. A DNA and/or RNA chip comprising at least one of the nucleic acid sequences selected from the group consisting of Seq. Id. No. 1 to Seq. Id. No. 17.

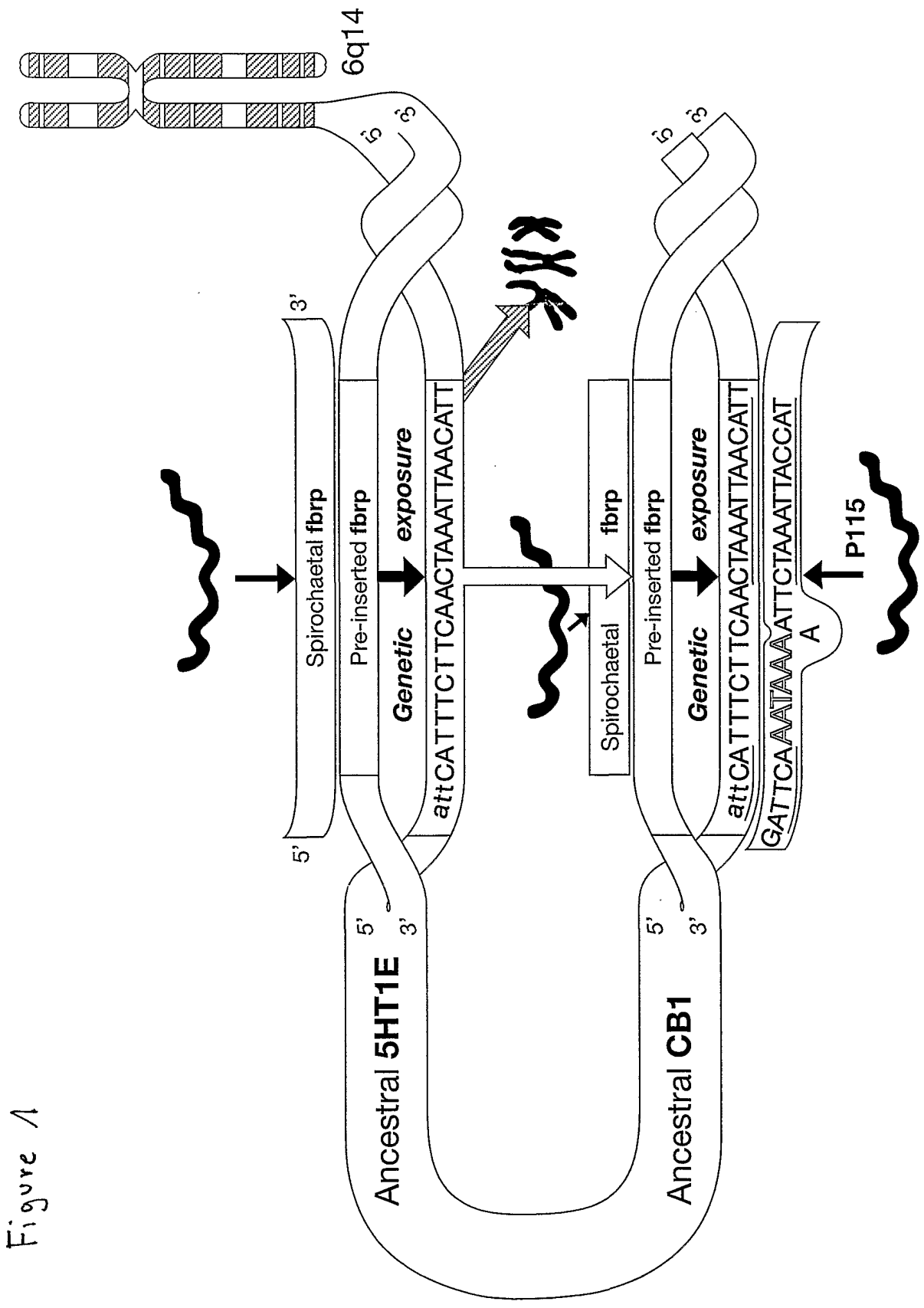
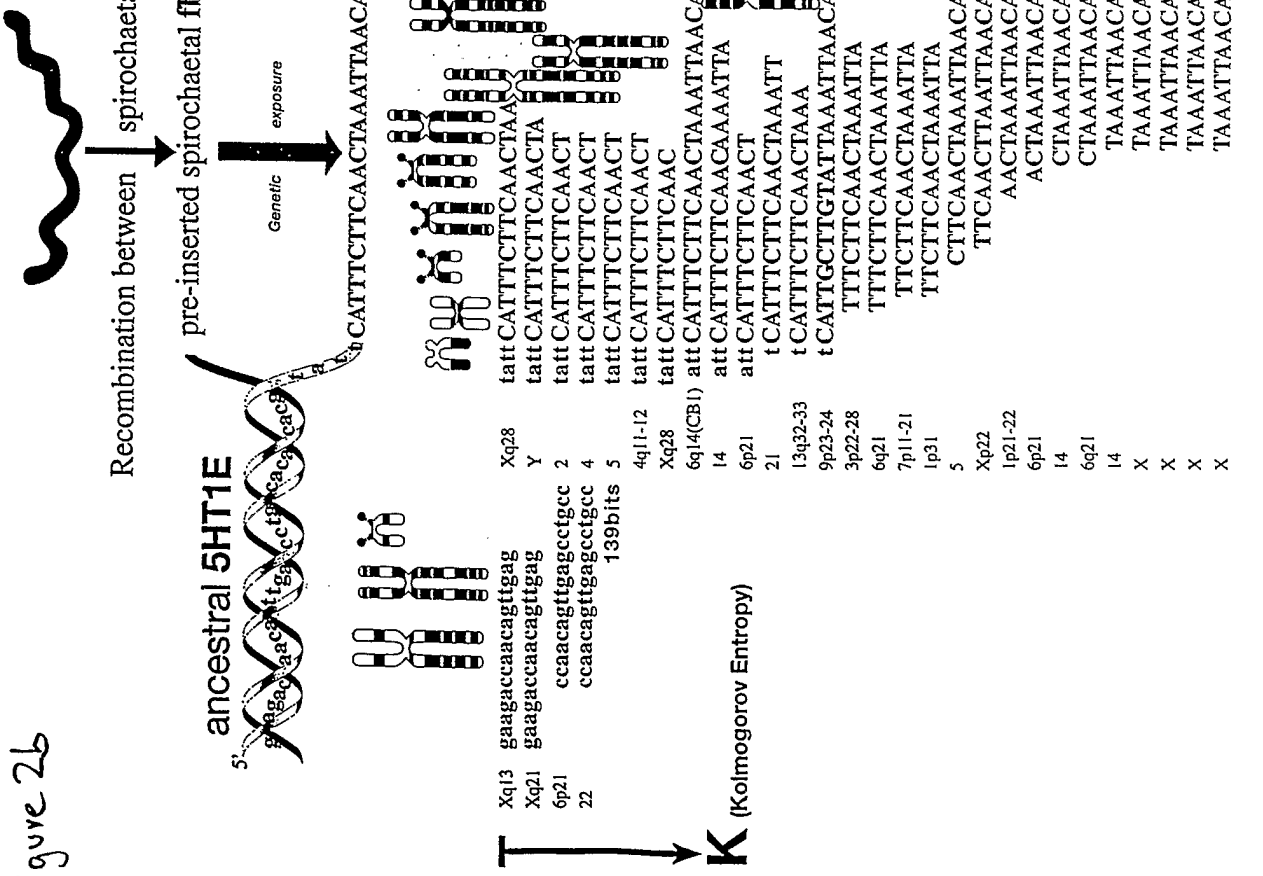

6q14
5'
3'
Spirochaetal fbrp
Pre-inserted fbrp
Genetic exposure
attCATTTCTTCAACTAAATTAACATT
Ancestral 5HT1E
fbrp
Spirochaetal
Pre-inserted fbrp
Genetic exposure
attCATTTCTTCAACTAAATTAACATT
GATTCAAATAAAATTCTAAATTACCAT
A
P115
Ancestral CB1

Figure 1

Figure 2a


Recombination between spirochaetal fbrp and
pre-inserted spirochaetal fbrp
ancestral 5HT1E
Genetic exposure
5'
3'
tCATTTCTTCAACTATATTAACATTt
Xq13 gaagaccaacagttgag
Xq21 gaagaccaacagttgag
6p21 ccaacagttgagcctgcc
22 ccaacagttgagcctgcc
139bits
Xq28 tattCATTTCTTCAACTA
Y tattCATTTCTTCAACTA
2 tattCATTTCTTCAACT
5 tattCATTTCTTCAACT
4p16 tattCATTTCTTCAACTTTATT
Xq27 tattCATTTCTTCAACAATAT
14 attCATTTCTTCAACTA
19 attCATTTCTTCAAGTATATT
Xp11 ttCATTTCTTCAACTATA
7 CATTTCTTCAACTATATTAAC
Xq23 TTTCTTCAACTATATT
13 TTTCTTCAACTATATT
21q22 TCTTCAACTATATTAACATT t
8p11 TCAACTATATTAACATT ttaa
X AACTATATTAACATT tt
21q21 ACTATATTAACATT ttaa
14q31 ACTATATTAACATT tta
7q31 CTATATTAACATT ttaa
Xq28 CTATATTAACATT ttaa
21q21 CTATATTAACATT ttaa
6q25-26 TATATTAACATT ttaa t
6q21 TATATTAACATT ttaa
6p11-12 TATATTAACATT ttaa
749bits
13q33 atttttttaaacgcatc
32bits
T
K
(Kolmogorov Entropy)

Figure 2b


Recombination between spirochaetal fbrp and
pre-inserted spirochaetal fbrp
ancestral 5HT1E
Genetic exposure
5'
3'
CATTTCTTCAACTAAATTAACATTt
Xq13 gaagaccaacagttgag Xq28 tatt CATTTCTTCAACTAA 13q33 atttttttaaacgcatc
Xq21 gaagaccaacagttgag Y tatt CATTTCTTCAACTA 32bits
6p21 ccaacagttgagcctgcc 2 tatt CATTTCTTCAACT
22 ccaacagttgagcctgcc 4 tatt CATTTCTTCAACT
139bits 5 tatt CATTTCTTCAACT
4q11-12 tatt CATTTCTTCAAC
Xq28 att CATTTCTTCAACTAAATTAACATT
6q14(CB1) att CATTTCTTCAACAAAATTA
14 att CATTTCTTCAACT
6p21 t CATTTCTTCAACTAAATT
21 t CATTTCTTCAACTAAA
13q32-33 t CATTGCTTGTATTAAATTAACATTttaa
9p23-24 TTTCTTCAACTAAATTA
3p22-28 TTTCTTCAACTAAATTA
6q21 TTCTTCAACTAAATTA
7p11-21 TTCTTCAACTAAATTA
1p31 CTTCAACTAAATTAACA
5 TTCAACTTAATTAACATTtta
Xp22 AACTAAATTAACATTtt
1p21-22 ACTAAATTAACATTtt
6p21 CTAAATTAACATTttaattaa
14 CTAAATTAACATTttaa
6q21 TAAATTAACATTttaa
14 TAAATTAACATTttaa
X TAAATTAACATTttaa
X TAAATTAACATTttaa
X TAAATTAACATTttaa
X TAAATTAACATTttaatt
957bits
K (Kolmogorov Entropy)

SEQUENCE LISTING

<110> Fritzsche, Markus

<120> Use of microbial DNA sequences for the identification of diseases

<130> Seq. Listing to file 04037PC1

<140>
<141>

<150> PCT/IB00/01127
<151> 2000-08-16

<160> 56

<170> PatentIn Ver. 2.1

<210> 1
<211> 28
<212> DNA
<213> Borrelia burgdorferi

<400> 1
gattcaaata aaaattctaa attaccat 28

<210> 2
<211> 22
<212> DNA
<213> Chlamydia pneumoniae

<400> 2
agttacctgg actcaaataa aa 22

<210> 3
<211> 14
<212> DNA
<213> Borrelia burgdorferi

<400> 3
gtttaataaa aatt 14

<210> 4
<211> 15

<212> DNA
<213> Chlamydia pneumoniae

<400> 4
agtttaataa agatt 15

<210> 5
<211> 16
<212> DNA
<213> Borrelia burgdorferi

<400> 5
atactaatat caataa 16

<210> 6
<211> 18
<212> DNA
<213> Plasmodium falciparum

<400> 6
ttttcatgta aataaata 18

<210> 7
<211> 17
<212> DNA
<213> Borrelia burgdorferi

<400> 7
tttaaataaa aataagt 17

<210> 8
<211> 25
<212> DNA
<213> Plasmodium falciparum

<400> 8
gatttttct ttagataaaa ataag 25

<210> 9
<211> 16
<212> DNA
<213> Borrelia burgdorferi

<400> 9
acaataaata atattc 16

<210> 10
<211> 16
<212> DNA
<213> Borrelia burgdorferi

<400> 10
taataataaa aatcat 16

<210> 11
<211> 19
<212> DNA
<213> Escherichia coli

<400> 11
aataataaaa atcatgctt 19

<210> 12
<211> 16
<212> DNA
<213> Plasmodium falciparum

<400> 12
aataataaaa atcatg 16

<210> 13
<211> 18
<212> DNA
<213> Borrelia burgdorferi

<400> 13
tccggaataa aaggccct 18

<210> 14
<211> 16
<212> DNA
<213> Chlamydia muridarum

<400> 14
gcgaataaaa ggccct 16

<210> 15
<211> 16
<212> DNA
<213> Rickettsia prowazekii

<400> 15
aataaaaaac atttat 16

<210> 16
<211> 20
<212> DNA
<213> Borrelia burgdorferi

<400> 16
aaactcagaa aataaaatgt 20

<210> 17
<211> 20
<212> DNA
<213> Borrelia burgdorferi

<400> 17
caaaataaag tagtaaaaga 20

<210> 18
<211> 24
<212> DNA
<213> Borrelia burgdorferi

<400> 18
catttcttca actaaattaa catt 24

<210> 19
<211> 27
<212> DNA
<213> Homo sapiens

<400> 19
attcatttct tcaactaaat taacatt 27

<210> 20
<211> 22

<212> DNA
<213> Chlamydia sp.

<400> 20
ttacctggac tcaaataaaa gt 22

<210> 21
<211> 36
<212> DNA
<213> Rattus norvegicus

<400> 21
ttacctggaa tcaaataaaa gttctagatt atcacg 36

<210> 22
<211> 38
<212> DNA
<213> ancestor rodent

<400> 22
gaatcaaata aaaattctag attaccatga agaacata 38

<210> 23
<211> 17
<212> DNA
<213> Borrelia garinii

<400> 23
ataataattc taaatta 17

<210> 24
<211> 17
<212> DNA
<213> Borrelia burgdorferi

<400> 24
ataataattc taaatta 17

<210> 25
<211> 16
<212> DNA
<213> Borrelia sp.

<400> 25
aatataaatt ctatat 16

<210> 26
<211> 14
<212> DNA
<213> ancestor human

<400> 26
gtttaataaa aatt 14

<210> 27
<211> 17
<212> DNA
<213> Homo sapiens

<400> 27
agtttaataa agattca 17

<210> 28
<211> 15
<212> DNA
<213> Borrelia sp.

<400> 28
tttaaaaaag attca 15

<210> 29
<211> 17
<212> DNA
<213> Homo sapiens

<400> 29
atactaatat caataaa 17

<210> 30
<211> 17
<212> DNA
<213> Borrelia burgdorferi

<400> 30
atactaatat caataat 17

<210> 31
<211> 24
<212> DNA
<213> Homo sapiens

<400> 31
ttttcatgta aataaataca ttct 24

<210> 32
<211> 17
<212> DNA
<213> Plasmodium falciparum

<400> 32
tttcaggtaa ataaata 17

<210> 33
<211> 26
<212> DNA
<213> Homo sapiens

<400> 33
gatttttct ttaaataaaa ataagt 26

<210> 34
<211> 21
<212> DNA
<213> Plasmodium falciparum

<400> 34
ttttttttt aaataaaaat a 21

<210> 35
<211> 21
<212> DNA
<213> Homo sapiens

<400> 35
cctaataaag gaatagttaa c 21

<210> 36
<211> 16

<212> DNA
<213> Homo sapiens

<400> 36
gctagtaata aaatat 16

<210> 37
<211> 16
<212> DNA
<213> Homo sapiens

<400> 37
acaataaata atattc 16

<210> 38
<211> 18
<212> DNA
<213> Homo sapiens

<400> 38
tccggaataa aaggccct 18

<210> 39
<211> 18
<212> DNA
<213> Homo sapiens

<400> 39
tcgcgaataa aaggccct 18

<210> 40
<211> 43
<212> DNA
<213> Homo sapiens

<400> 40
aataaaaaac atttattttc attgcaatga tgtatttaaa tta 43

<210> 41
<211> 12
<212> DNA
<213> Borrelia sp.

<400> 41
atgtatttaa at 12

<210> 42
<211> 22
<212> DNA
<213> Homo sapiens

<400> 42
tttctgaata ttttactaaa aa 22

<210> 43
<211> 21
<212> DNA
<213> Borrelia sp.

<400> 43
ttctgaagat tttactaaaa a 21

<210> 44
<211> 26
<212> DNA
<213> Homo sapiens

<400> 44
aataagaaac atttattttc attgca 26

<210> 45
<211> 17
<212> DNA
<213> Plasmodium falciparum

<400> 45
ataagaaaca tttattt 17

<210> 46
<211> 576
<212> DNA
<213> Homo sapiens

<400> 46
actcttctgg tccccacaga ctcagagaga acccaccatg gtgctgtctc ctgccgacaa 60
gaccaacgtc aaggccgcct ggggtaaggt cggcgcgcac gctggcgagt atggtgcgga 120

ggccctggag aggatgttcc tgtccttccc caccaccaag acctacttcc cgcacttcga 180
cctgagccac ggctctgccc aggttaaggg ccacggcaag aaggtggccg acgccctgac 240
caacgccgtg gcgcacgtgg acgacatgcc caacgcgctg tccgccctga gcgacctgca 300
cgcgcacaag cttcgggtgg acccggtcaa cttcaagctc ctaagccact gcctgctggt 360
gaccctggcc gcccacctcc ccgccgagtt cacccctgcg gtgcacgcct ccctggacaa 420
gttcctggct tctgtgagca ccgtgctgac ctccaaatac cgttaagctg gagcctcggt 480
ggccatgctt cttgcccctt gggcctcccc ccagcccctc ctcccccttcc tgcacccgta 540
cccccgtggt ctttgaataa agtctgagtg ggcggc 576

<210> 47
<211> 47
<212> DNA
<213> Homo sapiens

<400> 47
agttacctgg aatcaaataa aaattctaga ttaccatgaa gaacata 47

<210> 48
<211> 18
<212> DNA
<213> Salmonella typhimurium

<400> 48
tgtaaataaa tacattct 18

<210> 49
<211> 20
<212> DNA
<213> Homo sapiens

<400> 49
taataataaa aatcatgctt 20

<210> 50
<211> 14
<212> DNA
<213> ancestor human

<400> 50
gtttaataaa aatt 14

<210> 51
<211> 16

<212> DNA
<213> Borrelia sp.

<400> 51
tttaaataaa aataag 16

<210> 52
<211> 17
<212> DNA
<213> Vibrio cholerae

<400> 52
aataaaggaa tagttaa 17

<210> 53
<211> 16
<212> DNA
<213> Clostridium sp.

<400> 53
cctaataaag gaatag 16

<210> 54
<211> 20
<212> DNA
<213> Homo sapiens

<400> 54
aaactcagaa aataaaatgt 20

<210> 55
<211> 12
<212> DNA
<213> Borrelia burgdorferi

<400> 55
cagaagtaaa at 12

<210> 56
<211> 19
<212> DNA
<213> Homo sapiens

<400> 56
caaataaagt agtaaaaga 19

INTERNATIONAL SEARCH REPORT

Interi ial Application No
PCT/IB 01/00189

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12Q1/68

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12Q

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

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, MEDLINE, EMBASE, BIOSIS, EMBL, CAB Data, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	C.M. FRASER ET AL.: "Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi et al." NATURE, vol. 390, 11 December 1997 (1997-12-11), pages 580-586, XP002125758	28
Y	Sequence Accession number: AE01118	1-18

X	WO 98 58943 A (DOUGHERTY BRIAN A ;HUMAN GENOME SCIENCES INC (US); FRASER CLAIRE () 30 December 1998 (1998-12-30)	28
Y	the whole document	1-18

	-/--	

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

[X] Patent family members are listed in annex.

° Special categories of cited documents :

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

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 September 2001	09. 10. 2001

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Montero Lopez, B

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

INTERNATIONAL SEARCH REPORT

Interi ial Application No
PCT/IB 01/00189

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	HOEHE M R ET AL: "GENETIC AND PHYSICAL MAPPING OF THE HUMAN CANNABINOID RECEPTOR GENE TO CHROMOSOME 6Q14-Q15" THE NEW BIOLOGIST,PHILADELPHIA,PA,US, vol. 3, no. 9, September 1991 (1991-09), pages 880-885, XP000991359 ISSN: 1043-4674 the whole document ---	1-18
Y	GADZICKI D ET AL: "A FREQUENT POLYMORPHISM IN THE CODING EXON OF THE HUMAN CANNABINOIDRECEPTOR (CNR1) GENE" MOLECULAR AND CELLULAR PROBES,ACADEMIC PRESS, LONDON,GB, vol. 13, no. 4, 1999, pages 321-323, XP000982865 ISSN: 0890-8508 the whole document ---	1-18
A	N. KALCHEVA ET AL.: "Survey for CAG repeat polymorphisms in the human MAP-2 gene." PSYCHIATRIC GENETICS, vol. 9, no. 1, March 1999 (1999-03), pages 43-46, XP001004458 the whole document ---	
A	MORROW J A ET AL: "Functional characterization of apolipoprotein E isoforms overexpressed in Escherichia coli." PROTEIN EXPRESSION AND PURIFICATION, (1999 JUL) 16 (2) 224-30., XP001002071 the whole document ---	
A	NAKAI T. ET AL: "Membrane topology of Alzheimer 's disease-related presenilin 1. Evidence for the existence of a molecular species with a seven membrane-spanning and one membrane-embedded structure." JOURNAL OF BIOLOGICAL CHEMISTRY, (13 AUG 1999) 274/33 (23647-23658)., XP001002063 the whole document ---	
A	K.A. CONWAY ET AL.: "Fibrils formed in vitro from alpha-synuclein and two mutant forms linked to Parkinson's disease are typical amyloid." BIOCHEMISTRY, vol. 39, no. 10, 14 March 2000 (2000-03-14), pages 2552-2563, XP001002067 the whole document -----	

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB 01/00189

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

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

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

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

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

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

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

As a result of the prior review under R. 40.2(e) PCT, all additional fees are to be refunded.

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [X] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

1-18, 28 and partially 19-27

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[X] The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1 to 18, 28 and partially 19-27

Use of a DNA sequence comprising a fragment of a nucleic acid encoding a putative microbial virulence factor as a means for identification of a disease caused by mutations or a genetic predisposition, where the fragment is selected from SEQ ID NO: 1-17 and method for the identification of a disease or a genetic predisposition which comprises detecting the presence in a tissue-or blood sample of a subject a mutation within a nucleic acid sequence selected from SEQ ID NO: 1-17 and where said sequence is part of a gene of said subject. DNA and/or RNA chip comprising a nucleic acid sequence selected from SEQ ID NO:1-17. Transgenic non human animal whose genome comprises a partially or completely inactiveted gene selected from Cannabinoid receptor 1 gene, MAP 2C gene, apolipoprotein E gene, presenilin 2 gene, integral membrane protein 2B gene, alpha synuclein gene, oligophrenin 1 gene and myotonin protein kinase gene mutated in a sequence selected from SEQ ID NO:1-17; use thereof for the identification of compounds that have an effect on the activity, expression or regulation of the translated protein.

2. Claims: partially 19-27

Transgenic non human animal whose genome comprises a partially or completely inactiveted gene selected from Cannabinoid receptor 1 gene, MAP 2C gene, apolipoprotein E gene, presenilin 2 gene, integral membrane protein 2B gene, alpha synuclein gene, oligophrenin 1 gene and myotonin protein kinase gene other than by a mutation in a sequence selected among SEQ ID NO:1-17; use thereof for the identification of compounds that have an effect on the activity, expression or regulation of the translated protein.

INTERNATIONAL SEARCH REPORT

Inter al Application No
PCT/IB 01/00189

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9858943 A	30-12-1998	AU 8151898 A	04-01-1999
		AU 8153498 A	04-01-1999
		EP 1009859 A1	21-06-2000
		EP 1012157 A1	28-06-2000
		WO 9859071 A1	30-12-1998
		WO 9858943 A1	30-12-1998

专利名称(译)	利用微生物DNA序列鉴定人类疾病		
公开(公告)号	EP1309722A1	公开(公告)日	2003-05-14
申请号	EP2001902591	申请日	2001-02-15
[标]申请(专利权)人(译)	Fritzsche也MARKUS		
申请(专利权)人(译)	Fritzsche也，MARKUS		
当前申请(专利权)人(译)	Fritzsche也，MARKUS		
[标]发明人	FRITZSCHE MARKUS		
发明人	FRITZSCHE, MARKUS		
IPC分类号	G01N33/50 C12M1/00 C12N5/10 C12N15/09 C12Q1/68 C12Q1/6883 G01N33/15 G01N33/53 G01N33/566		
CPC分类号	C12Q1/6883 A01K2217/075 A01K2227/105 A01K2267/03 C12Q2600/156		
优先权	PCT/IB2000/001127 2000-08-16 WO		
外部链接	Espacenet		

摘要(译)

公开了使用包含编码微生物毒力因子的核酸片段的DNA序列作为鉴定疾病或其遗传易感性的手段以及其用于开发疾病动物模型的用途。