



(86) Date de dépôt PCT/PCT Filing Date: 2006/05/11
(87) Date publication PCT/PCT Publication Date: 2006/11/16
(85) Entrée phase nationale/National Entry: 2008/11/12
(86) N° demande PCT/PCT Application No.: DK 2006/000254
(87) N° publication PCT/PCT Publication No.: 2006/119775
(30) Priorité/Priority: 2005/05/11 (DK PA 2005 00680)

(51) Cl.Int./Int.Cl. *C12N 15/12* (2006.01),
A61K 31/7088 (2006.01), *A61K 48/00* (2006.01),
C07H 21/00 (2006.01), *C07H 21/04* (2006.01),
C07K 14/47 (2006.01), *C07K 16/18* (2006.01),
C12N 15/11 (2006.01), *C12N 15/85* (2006.01),
C12Q 1/68 (2006.01)

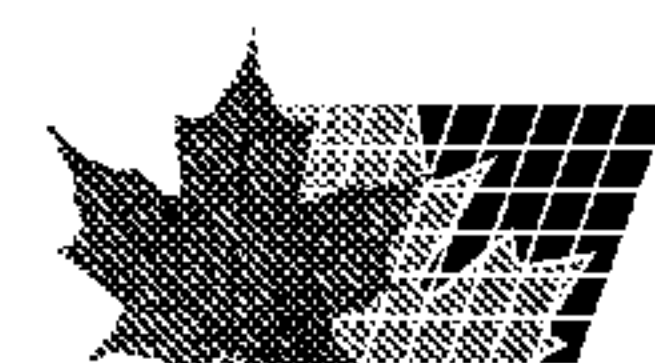
(71) Demandeurs/Applicants:
AARHUS UNIVERSITET, DK;
REGION MIDTJYLLAND, DK

(72) Inventeurs/Inventors:
BORGLUM, ANDERS, DK;
SEVERINSEN, JACOB, DK;
MORS, OLE, DK;

(54) Titre : PROCEDE DE DIAGNOSTIC ET DE TRAITEMENT D'UNE MALADIE MENTALE
(54) Title: METHOD FOR DIAGNOSIS AND TREATMENT OF A MENTAL DISEASE

(57) **Abrégé/Abstract:**

The present invention relates to association of one or more polymorphisms located in the human NHP2L1 , PACSIN2, SERHL, PIPPIN, BRD1 , EP300, FAM19A5 and/or GPR24 genes to the occurrence of schizophrenia and/or bipolar disorder. The invention relates both to methods for diagnosing a predisposition to said diseases and for treating subjects having said diseases.



(72) **Inventeurs(suite)/Inventors(continued):** MUIR, WALTER, GB; BLACKWOOD, DOUGLAS, GB;
EWALD, HENRIK (DECEASED), DK

(74) **Agent:** MACPHERSON LESLIE & TYERMAN LLP

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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
16 November 2006 (16.11.2006)

PCT

(10) International Publication Number
WO 2006/119775 A3

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(74) Agent: **HØIBERG A/S**; St. Kongensgade 59A, DK-1264
Copenhagen K (DK).

(21) International Application Number:
PCT/DK2006/000254

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(22) International Filing Date: 11 May 2006 (11.05.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PA 2005 00680 11 May 2005 (11.05.2005) DK

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(71) Applicants (*for all designated States except US*):
AARHUS UNIVERSITET [DK/DK]; Nordre Ringgade 1, DK-8000 Århus C (DK). **REGION MIDTJYLLAND** [DK/DK]; Skottenborg 26, DK-8800 Viborg (DK).

(72) Inventor: **EWALD, Henrik**; c/o Trustee Accountant, SØRENSEN, Peter, Smedevej 1B, DK-9500 Hobro.

Published:

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

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **BØRGLUM, Anders** [DK/DK]; Thunøgade 1E, 1.sal, DK-8000 Århus C (DK). **SEVERINSEN, Jacob** [DK/DK]; Snogebæksvej 13, vær, 27, DK-8210 Århus C (DK). **MORS, Ole** [DK/DK]; Permelillevej 7, DK-8240 Risskov (DK). **MUIR, Walter** [GB/GB]; The Chestnuts, Eshiels, Peebles, EH45 8NA (GB). **BLACKWOOD, Douglas** [GB/GB]; 7 Pitsligo Road, Edinburgh EH10 4RY (GB).

(88) Date of publication of the international search report:
28 June 2007

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.

(54) Title: METHOD FOR DIAGNOSIS AND TREATMENT, OF A MENTAL DISEASE

(57) Abstract: The present invention relates to association of one or more polymorphisms located in the human NHP2L1, PAC-SIN2, SERHL, PIPPIN, BRD1, EP300, FAM19A5 and/or GPR24 genes to the occurrence of schizophrenia and/or bipolar disorder. The invention relates both to methods for diagnosing a predisposition to said diseases and for treating subjects having said diseases.



WO 2006/119775 A3

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 2
CONTENANT LES PAGES 1 À 250

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 2
CONTAINING PAGES 1 TO 250

NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

Method for diagnosis and treatment of a mental disease**Field of invention**

5

The present invention relates to association of one or more polymorphisms located in the human NHP2L1, PACSIN2, SERHL, PIPPIN, BRD1, EP300, FAM19A5 and/or GPR24 genes to the occurrence of schizophrenia and/or bipolar disorder. The invention relates both to methods for diagnosing a predisposition to said diseases and for treating subjects having said diseases.

10

Background of invention***Polymorphisms***

DNA polymorphisms provide an efficient way to study the association of genes and diseases by analysis of linkage and linkage disequilibrium. With the sequencing of the human genome a myriad of hitherto unknown genetic polymorphisms have been detected. Most common among these are the single nucleotide polymorphisms, also called SNPs, of which are counted now to several millions. Other examples of genetic polymorphisms are variable number of tandem repeat polymorphisms, insertions, deletions and block modifications. Tandem repeats often have multiple different alleles (variants), whereas the other groups of polymorphisms usually just have two alleles. Some of these genetic polymorphisms probably play a direct role in the biology of the individuals, including the risk of developing disease, but the virtue of the majority is that they can serve as markers for the surrounding DNA.

25

The association of an allele of one sequence polymorphism with particular alleles of other sequence polymorphisms in the surrounding DNA has two origins, known in the genetic field as linkage and linkage disequilibrium, respectively. Linkage arises because large parts of chromosomes are passed unchanged from parents to offspring, so that minor regions of a chromosome tend to flow unchanged from one generation to the next and also to be similar in different branches of the same family. Linkage is gradually eroded by recombination occurring in the cells of the germline, but typically operates over multiple generations and distances of a number of million bases in the DNA.

35

Linkage disequilibrium deals with whole populations and has its origin in the (distant) forefather in whose DNA a new sequence polymorphism arose. The immediate surroundings in the DNA of the forefather will tend to stay with the new allele for many generations. Recombination and changes in the composition of the population will again erode the association, but the new allele and the alleles of any other polymorphism nearby will often be partly associated among unrelated humans even today. A crude estimate suggests that alleles of sequence polymorphisms with distances less than 10000 bases in the DNA will have tended to stay together since modern man arose. Linkage disequilibrium in limited populations, for instance Europeans, often extends over longer distances, e.g. over more than 1,000,000 bases. This can be the result of newer mutations, but can also be a consequence of one or more "bottlenecks" with small population sizes and considerable inbreeding in the history of the current population. Two obvious possibilities for "bottlenecks" in Europeans are the exodus from Africa and the repopulation of Europe after the last ice age.

Schizophrenia and bipolar disorder

The importance of genetics in the etiology of bipolar disorder (BPD) and schizophrenia (SCH) are by now an accepted fact and confirmed through family, twin and adoption studies (McGuffin et al., 1995; Potash and DePaulo, 2000). Overlapping between these diseases has long been noted clinically and a shared genetic heritability is probable according to genetic epidemiology studies and linkage studies (Berrettini, 2000; Potash et al., 2003a) Linkage studies have identified several shared susceptibility loci (on 18p11; 13q31; 10p14; and 22q11-13) which makes it plausible that there are partial shared genetic risks, and susceptibility genes. On chromosome 22q12-13 implication in schizophrenia has been suggested by weak linkage (Gill et al., 1996; Pulver et al., 1994) and a recent meta-analysis including 20 genome scans in schizophrenics has confirmed the locus 22q12-13 as a susceptibility locus. In addition significant linkage to BPD on 22q12-13 was evidenced by Kelso et al (Kelsoe et al., 2001). In a meta-analysis including 11 BPD and 18 schizophrenic genome scans 22q12-13 were one of two regions that showed strong evidence for harboring a susceptibility loci for both BPD and SCH (Badner and Gershon, 2002; Potash et al., 2003b). However, there were obtained no evidence that polymorphism of DNA in this locus is associated with the diseases.

Genes

Several genes with known or unknown biological function have been mapped to chromosome 22q13.

5 1. BRD1

The gene was originally described in connection with study of the genes involved in oncogenic transformation in human acute leukemias and mapped to 22q13 by fluorescence in situ hybridization (McCullagh et al. (1999)).

10 In human acute leukemias the MLL gene is reciprocally translocated with one of a number of different partner genes. The precise mechanism of oncogenic transformation is unclear since most of the partner genes encode unrelated proteins. However, 2 partner genes, AF10 and AF17, are related through the presence of a cysteine-rich region and a leucine zipper. McCullagh et al. (1999) suggested that the
15 identification of other proteins having these structures might aid understanding of their role in normal and leukemic cells. They reported the cloning of a novel human gene that encodes a 1,058-amino acid protein containing a cysteine-rich region related to that of AF10 and AF17. Overall, the protein was most closely related to the BR140 protein, and was therefore designated BRL (BR140-like gene). Northern blot
20 analysis revealed that a 4.6-kb BRL transcript is expressed at high levels in testis and in several cell lines. A monoclonal antibody raised to a BRL peptide sequence confirmed its widespread expression as a 120-kD protein and demonstrated localization to the nucleus of spermatocytes.

25 The gene encodes a protein of unknown function. The protein contains a bromodomain, a sequence motif often found in transcriptional co-activators, and localizes to the nucleus in testis and several other cell types.

2. NHP2L1

30 In the course of sequencing cDNA clones from fetal brain cDNA libraries, Saito et al. (1996) isolated a gene that encodes a protein highly homologous to NHP2 from *Saccharomyces cerevisiae* (Kolodrubetz and Burgum, 1991). NHP2 is a high-mobility group (HMG)-like protein which is located in the nucleus, although it shows weak homology to some ribosomal proteins. The cDNA cloned by Saito et al. (1996),
35 symbolized NHP2L1, was expressed in all human tissues examined and was localized to 12q24.3 by fluorescence in situ hybridization.

Originally named because of its sequence similarity to the *Saccharomyces cerevisiae* NHP2 (non-histone protein 2), this protein appears to be a highly conserved nuclear protein that is a component of the [U4/U6.U5] tri-snRNP. It binds to the 5' stem-loop
5 of U4 snRNA. It is suggested that the protein may play a role in the late stage of spliceosome assembly. The protein is related to the L7AE family of ribosomal proteins. The protein has ubiquitous tissue expression. Two transcript variants encoding the same protein have been found for human NHP2L1 gene.

10 3. PACSIN2

Human PACSIN2 gene was mapped to 22q13 (Sumoy et al. (2001)).

PACSIN family members, such as mouse Pacsin1 and chicken FAP52, are cytoplasmic adapter proteins with a common arrangement of domains and conserved
15 regions, including a CDC15 N-terminal domain, which contains a RAEYL motif and a coiled-coil region, and a C-terminal SH3 domain. By searching an EST database for novel members of the PACSIN family, Ritter et al. (1999) identified ESTs encoding human PACSIN2 and mouse Pacsin2. The complete human PACSIN2 coding sequence, obtained from 2 overlapping retina ESTs, encodes a deduced 486-amino
20 acid protein that shares 93.6% sequence identity with mouse Pacsin2. The PACSIN2 proteins contain a CDC15 N-terminal domain, a C-terminal SH3 domain, 3 conserved regions specific to the PACSIN family, and 3 asn-pro-phe (NPF) motifs, which potentially bind to EH domains. The PACSIN2 proteins share high sequence similarity with chicken FAP52 and mouse Pacsin1. However, compared to these
25 proteins, PACSIN2 proteins have a 41-amino acid insertion, which contains 1 NPF motif. In contrast to the restricted neural expression of Pacsin1 protein (Plomann et al., 1998), Northern blot analysis detected Pacsin2 transcript expression in all tissues examined, with the highest levels in brain, heart, skeletal muscle, and ovary. Immunofluorescence microscopy revealed a broad, vesicle-like cytoplasmic
30 distribution of recombinant Pacsin2 expressed in fibroblasts that appeared to partially overlap with the distributions of both the actin filament and microtubule networks. Ritter et al. (1999) suggested that PACSIN2 protein may participate in the organization of the actin cytoskeleton and the regulation of vesicular traffic.

35 Human PACSIN2 can homo- and hetero-aggregate with other PACSINs, bind dynamin 1, synaptojanin, synapsin 1 and the neural Wiskott-Aldrich syndrome protein

(N-WASP), is phosphorylated by casein kinase 2 (CK2) and protein kinase C (PKC). Vesicle-like cytoplasmic distribution of the protein suggests its possible role in vesicle formation and transport.

5

4. SERHL

Sadusky et al. (2001) cloned Serhl by subtractive hybridization of transcripts upregulated in passively stretched mouse skeletal muscle. The deduced 311-amino acid protein contains a putative serine hydrolase active center, prenylation and N-myristoylation sites, several putative phosphorylation sites, an N-glycosylation site, and a C-terminal peroxisomal targeting sequence. Northern blot analysis detected transcripts of 1.4 and 2.4 kb expressed in all mouse tissues examined. Highest levels of both transcripts were detected in kidney, and lowest levels were detected in skeletal muscle and small intestine. Brain predominantly expressed the 2.4-kb form. Western blot analysis detected a Serhl protein with an apparent molecular mass of about 35 kD in mouse skeletal and cardiac muscle, brain, and kidney. Immunolocalization detected Serhl expressed in small perinuclear vesicles in mononucleated mouse myoblasts and recently fused multinucleated myotubes. By genomic sequence analysis Sadusky et al. (2001) mapped the SERHL gene to 22q13.2.

5. PIPPIN

The gene encodes an RNA-binding protein, which is highly enriched in the brain, contains two putative double stranded RNA-binding domains, a cold-shock domain, binds with high specificity to the transcripts that encode H1 and H3.3. histone variants (Castiglia et al. 1996, Nastasi et al., 1999; Nastasi et al, 2000). It has been shown that PIPPIN inhibits translation of H1(0) and H3.3 mRNA in cell-free system. Based on this finding it has been suggested that PIPPIN down-regulates histone variant expression in the developing ret brain. The PIPPIN gene is mapped to 22q13.2

6. EP300

Eckner et al. (1994) mapped the p300 gene, symbolized EP300, to 22q13.2.

EP300 encodes the adenovirus E1A-associated cellular p300 transcriptional co-activator protein. The growth-controlling functions of the adenovirus E1A oncoprotein

depend on its ability to interact with a set of cellular proteins. Among these are the retinoblastoma protein, p107, p130, and p300. Eckner et al. (1994) isolated a cDNA encoding full-length human p300. p300 contains 3 cysteine- and histidine-rich regions of which the most carboxy-terminal region interacts specifically with E1A. In
5 its center, p300 contains a bromodomain, a hallmark of certain transcriptional coactivators. p300 and CREB-binding protein (CREBBP, or CBP) are highly related in primary structure (Arany et al., 1994). Several protein motifs such as a bromodomain, a KIX domain, and 3 regions rich in cys/his residues are well conserved between these 2 proteins. Lin et al. (2001) identified a compactly folded
10 46-residue domain in CBP and p300, the IRF3-binding domain (IBID), and determined its structure by nuclear magnetic resonance spectroscopy. IBID has a helical framework containing an apparently flexible polyglutamine loop that participates in ligand binding. Spectroscopic data indicated that induced folding accompanies association of IBID with its partners, which exhibit no evident sequence
15 similarities. IBID is an important contributor to signal integration by CBP and p300.

EP300 (p300) is a multifunctional protein:

- like CPB it can stimulate transcription through activation of CREB. This EP300 activity is specifically inhibited by the adenovirus oncoprotein E1A;
- 20 - EP300 has also been identified as a co-activator of HIF1A (hypoxia-inducible factor 1 alpha), and thus plays a role in the stimulation of hypoxia-induced genes such as VEGF;
- EP300/CREBBP and IRF3 are components of DRAF1 (double-stranded RNA-activated factor-1), a positive regulator of interferon-stimulated gene
25 transcription that functions as a direct response to viral infection (Weaver et al. 1998);
- the formation of a complex between transcription factors STAT3 and SMAD1, bridged by p300, is involved in the cooperative signaling of cytokines LIF and BMP2 and the subsequent induction of astrocytes from neuronal progenitors
30 (Nakashima et al. 1999);
- it play a role in DNA repair synthesis and other DNA metabolic events through its interaction with proliferating cell nuclear antigen (PCNA) and with flap endonuclease-1 (FEN1) (Hasan et al. 2001);
- it has a regulatory role for protein acetylation in base mismatch repair and
35 maintaining genomic stability (Tini et al. 2002);

- has a role in the mechanism for circadian phase control (Etchegaray et al. 2003);
- generation of the polyubiquitinated forms of p53 that are targeted for proteasome degradation requires the intrinsic ubiquitin ligase activities of MDM2 and p300 (Grossman et al. 2003);

The EP300 protein is a histone acetyltransferase that regulates transcription via chromatin remodeling and is important in the processes of cell proliferation and differentiation. A role for EP300 in cancer had been implied by the fact that it is targeted by viral oncoproteins (Arany et al., 1995), it is fused to MLL in leukemia (Ida et al., 1997), and 2 missense sequence alterations in EP300 were identified in epithelial malignancies (Muraoka et al., 1996). Gayther et al. (2000) described EP300 mutations that predicted a truncated protein in 6 (3%) of 193 epithelial cancers analyzed. Of these 6 mutations, 2 were in primary tumors (a colorectal cancer and a breast cancer) and 4 were in cancer cell lines (colorectal, breast, and pancreatic). In addition, they identified a somatic in-frame insertion in a primary breast cancer and missense alterations in a primary colorectal cancer and 2 cell lines (breast and pancreatic). Inactivation of the second allele was demonstrated in 5 of the 6 cases with truncating mutations and in 2 other cases. The data showed that EP300 is mutated in epithelial cancers and provided the first evidence that it behaves as a classic tumor suppressor gene.

7. GPR24

The protein encoded by GPR24 gene is a member of the G protein-coupled receptor family 1, an integral plasma membrane protein which binds melanin-concentrating hormone. The encoded protein can inhibit cAMP accumulation and stimulate intracellular calcium flux, and is probably involved in the neuronal regulation of food consumption. Although structurally similar to somatostatin receptors, the protein does not seem to bind somatostatin. The protein encoded by the GPR24 gene is also termed in scientific literature as SLC1, MCH1R, MCHR1.

The somatostatin receptors (SSTRs) are a family of G protein-coupled receptors which bind to somatostatin peptides. Kolakowski et al. (1996) identified an EST sequence with significant homology to the SSTRs that does not bind somatostatin. They cloned the gene corresponding to this EST from a human genomic library. Sequencing of a genomic clone, which they termed SLC1, revealed a full-length and

intronless open reading frame encoding a 402-amino acid protein. Its transmembrane regions are approximately 40% identical to other members of the SSTR family, including several residues considered to form the ligand-binding pocket of SSTRs. Northern blot analysis detected a single 2.4-kb transcript with greatest abundance in the human brain, particularly in the frontal cortex and hypothalamus. These regions are associated with emotion, memory, and sensory perception. Kolakowski et al. (1996) expressed the SLC1 receptor in COS-7 cells and found that it does not bind to somatostatin peptides. They identified a polymorphic CA repeat in the 5-prime untranslated region of this gene, and they used fluorescence in situ hybridization to map the gene to 22q13.3.

Chambers et al. (1999) used a reverse pharmacology approach to identify the natural cognate ligand for SLC1. They expressed the receptor in HEK293 cells and screened against a large library of known bioactive substances, including over 500 naturally occurring or putative neuropeptides. In this screen, melanin-concentrating hormone (MCH) was the only substance to produce a robust, dose-dependent ($EC_{50} = 3.72$ nM), transient elevation of intracellular calcium in HEK293 cells transiently transfected with SLC1.

MCH can act as a functional antagonist of alpha-melanocyte-stimulating hormone (alpha-MSH) in a diverse range of animal species and physiologic roles. In mammals, MCH is orexigenic and alpha-MSH is anorexigenic. Chambers et al. (1999) tested alpha-MSH at up to 10-mM concentration and no agonistic or antagonistic interaction with SLC1 was observed. Together with the observation that MCH could not displace alpha-MSH from melanocortin receptors, these results supported the idea that the functional, mutually antagonistic effects of alpha-MSH and MCH are mediated by their interaction at separate receptors. Using in situ hybridization, Chambers et al. (1999) demonstrated that SLC1 is widely and strongly expressed in the rat brain. There were clear mRNA signals in the olfactory tubercle, cerebral cortex, substantia nigra, basal forebrain, CA1, CA2, and CA3 fields of the hippocampus, amygdala, and various other nuclei in the hypothalamus, thalamus, midbrain, and hindbrain. There were also strong signals in the ventromedial and dorsomedial nuclei of the hypothalamus, areas widely recognized as being involved in feeding behavior.

Saito et al. (1999) identified SLC1 as an MCH hormone receptor using a different

technique and found similar EC-50 and brain distribution for MCHR1 (SLC1).

Borowsky et al. (2002) showed that the selective high-affinity MCH1R (MCHR1) receptor antagonist SNAP-7941 inhibited food intake stimulated by central
5 administration of MHC, reduced consumption of palatable food, and after chronic administration to rats with diet-induced obesity, resulted in a marked, sustained decrease in body weight. Borowsky et al. (2002) also showed that SNAP-7941 produced effects similar to clinically used antidepressants and anxiolytics in 3 animal models of depression/anxiety: the rat forced-swim test, rat social interaction, and
10 guinea pig maternal-separation vocalization tests. Given these observations, the authors concluded that an MCH1R antagonist may be useful not only in the management of obesity but also as a treatment for depression and/or anxiety.

Genetic markers of mental diseases

15 A number of studies have been performed in the attempt to identify the genes responsible for development of mental diseases such as SCH and BPD. Among these quite a few studies focused on the finding a correlation of the polymorphism of genomic sequences with development and inheritance of mental disorders such as SCH and BPD, which would be useful in disease diagnosis and prognosis (Skibinska
20 et al. 2004;; Xi et al. 2004; Sinibaldi et al. 2004; Li et al.,2004). However, there were so far identified no reliable genetic markers in association with the diseases so far.

Summary of invention

25

The authors of the present invention for the first time herein

1) described a strong correlation between polymorphism(s) of selected genes mapped to chromosome 22q13 such as the NHP2L1, PACSIN2, SERHL, PIPPIN, BRD1, EP300, FAM19A5 and/or GPR24 genes, and a predisposition to
30 SCH and/or BPD, proposed to use said polymorphism(s) for diagnosis/prognosis of a predisposition of an individual to SCH and/or BPD, and suggested said polymorphisms, said genes, their gene products and related pathways/interacting genes and gene products as targets for medical treatment of SCH and/or BPD and in the search for new drug candidates for medical treatment of SCH and/or
35 BPD;

- 2) described the association of specific haplotypes of the identified polymorphisms with a predisposition to SCH and/or BPD and proposed to use haplotype analysis as a mean in diagnosis/prognosis of a predisposition to SCH and/or BPD and the SNPs of said haplotypes as targets for medical treatment of SCH and/or BPD and in the search for new drug candidates for medical treatment of SCH and/or BPD;
- 3) described single nucleotide polymorphisms (SNPs) mapped to the genes of chromosome 22q13 which are in linkage disequilibrium with the polymorphisms of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 genes associated with SCH and/or BPD, and proposed to use the latter SNPs as diagnostic/prognostic markers of a predisposition to SCH and/or BPD and targets for medical treatment SCH and/or BPD and in the search for new drug candidates for medical treatment of SCH and/or BPD;
- 4) described new etiologic factors, such as transcription and translation products of the above genes containing polymorphism(s), associated with SCH and/or BPD, and provided methods for diagnosis, prognosis and treatment of SCH and/or BPD comprising a step of identification such factor in a sample from a patient and/or a step of modulating biological activity of such factor a compound described herein;
- 5) provided a method of determining a predisposition/no predisposition to/protection against SCH and/or BPD and a method for diagnosis of said diseases, said methods comprising determining a polymorphism of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 genes and/or a polymorphism of another gene of chromosome 22q13, said polymorphisms being in linkage disequilibrium;
- 6) provided a method for gene therapy treatment of SCH and/or BPD, said method comprising providing a gene therapy vector comprising a DNA sequence of the invention, e.g. a DNA sequence including the protective allele of a polymorphism being associated with SCH and/or BPD;
- 7) provided a method for identification of new drug candidate compounds for the treatment of SCH and/or BPD, said method comprising screening compounds for a capability of inhibiting/modulating biological activity of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 genes and/or products of said genes, said activity being associated with manifestation/predisposition of/to SCH and/or BPD;
- 8) provided diagnostic/prognostic kits for diagnosis/prognosis of SCH and/or BPD;

9) provided a method for estimating the likelihood of developing SCH and/or BPD in an individual comprising a step of determining the allele of an SNP of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 genes, said SNP being associated with a predisposition of an individual to SCH and/or BPD.

5

Accordingly, in the first aspect the invention relates to a method for determining a predisposition/no predisposition to a mental disease in a subject comprising determining in a biological sample isolated from said subject one or more polymorphisms in the chromosome regions containing the NHP2L1, PACSIN2,
10 SERHL, PIPPIN, BRD1, EP300, FAM19A5 and/or GPR24 genes, or in a translational or transcriptional product from said regions, said polymorphism being indicative of said predisposition/no predisposition.

The present inventors have discovered that polymorphisms, such as SNPs, identified
15 in the coding and/or non-coding regions of the NHP2L1, PACSIN2, SERHL, PIPPIN, BRD1, EP300, FAM19A5 and/or GPR24 genes are strongly associated with the presence or absence of some mental diseases such as SCH and BPD. Thus, detecting the presence or absence of one or more alleles of the SNPs of the present invention amounts to determining a predisposition to having or not having a mental
20 disease. It thus follows that determining the presence of a specific allele amounts to determining a predisposition to having/not having a mental disease. According to the invention the strength of the association between the presence/absence of a specific SNP in the above genes and the diseases is very strong.

25 Diagnosis of individuals for genetic predisposition to mental diseases, such as SCH and/or BPD is very important so that they can be given the best treatment and adapt their lifestyle according to their genetic predisposition.

The authors of the present invention performed haplotype analysis of the identified
30 SNPs and found out that the coincidence of some haplotypes in association with a particular disease is higher than the coincidence of another haplotype and the disease. Thus, the invention also relates to specific haplotypes of the identified SNPs. Moreover, it is expected that with the information made available by the inventors, more polymorphisms in the NHP2L1, PACSIN2, SERHL, PIPPIN, BRD1,
35 EP300, FAM19A5 and/or GPR24 genes will be found predisposing to mental diseases, such as SCH and/or BPD. Therefore, all polymorphisms being in linkage

disequilibrium with the identified in the present invention SNPs in the chromosome regions adjusted to the genes of the invention are included in the scope of the protection as diagnostic markers of the predisposition to a mental disease, in particular SCH and/or BPD. Moreover, some particular SNPs found in other genes
5 than the genes of the invention and described herein as a part of specific haplotypes associated with a predisposition to SCH and/or BPD are also concerned by the invention as genetic markers of said predisposition.

The inventors have defined the alleles of the described SNPs, which when expressed
10 in an individual have either promotional or no effect on developing a mental disease (the susceptibility or protective allele correspondingly) The application thus also provides a method for estimating the likelihood for development of SCH and/or BPD by analysing the allele variance of the SNPs associated with SCH and/or BPD present in an individual.

15 In a further aspect the invention relates to isolated oligonucleotide sequences comprising at least 10 contiguous nucleotides being 100% identical to a subsequence of the NHP2L1, PACSIN2, SERHL, PIPPIN, BRD1, EP300, FAM19A5 and/or GPR24 genes comprising or adjacent to a polymorphism of the invention, said
20 polymorphism or mutation being associated to a mental disease.

As the present inventors have determined that the NHP2L1, PACSIN2, SERHL, PIPPIN, BRD1, EP300 and GPR24 genes are etiological factors in mental diseases it is important to be able to detect and correct or suppress any polymorphism in the
25 genes which is correlated to these diseases. The isolated oligonucleotides may be used as probes for detection of the polymorphisms and/or as primer pairs for amplification of a target nucleotide sequence and/or as part of a gene therapy vector for administration to a patient suffering from mental diseases, such as SCH and/or BPD.

30 In a further aspect the invention relates to a kit for predicting an increased risk of a subject of developing mental diseases, such as SCH and/or BPD or for other diagnostic and classification purposes of SCH and/or BPD comprising at least one probe comprising at least two nucleic acid sequences as defined above.

35

These kits which may further comprise buffers and primers and reagents can be used for diagnosing the polymorphisms and mutations which correlate to mental diseases of the invention.

5 The invention also relates to variant proteins comprising variants which correspond to the identified in the application polymorphisms of the corresponding genes. These variant proteins may also be used for diagnosis of the described herein mental diseases.

10 According to a further aspect the invention relates to antibodies capable of selectively binding to the variant proteins as defined above with a different (such as lower or higher) binding affinity than when binding to the polypeptide having the amino acid sequence of wild type protein.

15 These antibodies may be used in diagnosing individuals with the polymorphisms. It is also envisaged that such specific antibodies may be used for treating patients carrying the variant protein.

In further aspects the present invention relates to methods of treating patients
20 suffering from mental disorders, in particular SCH and/or BPD. Among the therapeutic methods, one method relates to a method of treating SCH and/or BPD in a subject being diagnosed as having a predisposition according to the invention, comprising administering to said subject a therapeutically effective amount of a gene therapy vector. The invention also relates to a gene therapy vector itself, said vector
25 being capable of altering the polymorphism in cells of a subject being diagnosed as having a predisposition according to the invention, or being capable of correcting, suppressing, supporting or changing the expression of the NHP2L1, PACSIN2, SERHL, PIPPIN, BRD1, EP300, FAM19A5 and/or GPR24 genes in cells of a subject suffering from said diseases.

30

With the advent of gene therapy it has become possible to suppress and/or to eliminate the effects of a polymorphism by administering to a subject a gene therapy vector which either alters the polymorphism or suppresses the transcription and/or translation from the gene. Such gene therapy vectors have the advantage of being
35 highly specific. Therapeutic methods of treatment of patients suffering from SCH and/or BPD of the invention also include methods of treatment comprising a step of

modulating the activity of products of the genes comprising a polymorphism associated with SCH and/or BPD.

Definitions

5 Gene / gene sequence:

A compilation of

- the genomic sequences which are transcribed into a transcriptional entity
 - the genomic sequences in between
 - the genomic sequences involved in regulation of expression and splicing of the
- 10 gene comprising at least 2000 bp upstream and downstream from the transcribed entity.

The present invention relates to the genes identified in the NCBI database (<http://www.ncbi.nlm.nih.gov>) as

GeneID: 4809 (NHP2L1)

15 GeneID: 11252 (PACSIN2)

GeneID: 94009 (SERHL)

GeneID: 27254 (PIPPIN)

GeneID: 23774 (BRD1)

GeneID: 2033 (EP300)

20 GeneID: 2847 (GPR24)

GeneID: 25817 (FAM19A5)

Genomic sequences of the above genes (<http://genome.ucsc.edu/>) are identified in the present invention as

NHP2L1 gene SEQ ID NO: 1

25 PACSIN2 gene SEQ ID NO: 2

SERHL gene SEQ ID NO: 3

PIPPIN gene SEQ ID NO: 4

BRD1 gene SEQ ID NO: 5

EP300 gene SEQ ID NO: 6

30 GPR24 gene SEQ ID NO: 7

FAM19A5 SEQ ID NO: 94

The term "chromosome region containing a gene" means herein a part of a human chromosome containing a gene of the invention and nucleotide sequences of 2 to

35 2000 base pairs adjacent to both ends of the defined gene sequence (SEQ ID NO: 1-

7 and 94), wherein one end of the gene corresponds to the first nucleotide of the gene sequence, and another end corresponds to the last nucleotide of the gene sequence.

- 5 The term "adjacent" is used in connection with
- (i) a gene sequence to indicate a nucleotide sequence/chromosome region that is sufficiently closely located to said gene sequence in a chromosome, such as for instance less than 1 000 000 for example within 900 000-200 000 nucleotide positions, such as 100 000, e. g. less than 90 000, such as less than 80 000, e. g. less than 70 000, such as less than 60 000, e. g. from 10 000 to 50 000, e. g. 20 000 or 10 000 nucleotide positions, or from 1 to 10 000 nucleotide positions. It is preferred that the adjacent region is in linkage disequilibrium with said gene sequence;
 - (ii) an oligonucleotide sequence to indicate that said oligonucleotide sequence recognises a sequence that is sufficiently closely located to a specific nucleotide of interest for the oligonucleotide sequence to be suitable for the desired detection technique, such as for instance as a primer for amplification of a target nucleotide sequence. Preferably, adjacent means less than 500, such as less than 400, e.g. less than 300, such as less than 200, e.g. less than 100, such as less than 50 nucleotide positions away from the nucleotide or nucleotide sequence of interest.

As used herein, the term "coding sequence" refers to that portion of a gene that encodes the amino acid sequence of a protein. Exons contain the coding sequence of the gene.

Coding sequences of the above genes (cDNA) are identified herein as

NHP2L1 cDNA	SEQ ID NO: 8
PACSIN2 cDNA	SEQ ID NO: 9
30 SERHL cDNA	SEQ ID NO: 10
PIPPIN cDNA	SEQ ID NO: 11
BRD1 cDNA	SEQ ID NO: 12
EP300 cDNA	SEQ ID NO: 13
GPR24 cDNA	SEQ ID NO: 14..
35 FAM19A5	SEQ ID NO: 95

The promoter, UTR and intron regions referred herein as the "non-coding region(s)/sequence(s)" of the given genes. As used herein, "intron" refers to a DNA sequence present in a given gene that is spliced out during mRNA maturation. The term "promoter region" refers to the portion of DNA of a gene that controls transcription of the DNA to which it is operatively linked. The promoter region includes specific sequences of DNA that are sufficient for RNA polymerase recognition, binding and transcription initiation. This portion of the promoter region is referred to as the promoter. In addition, the promoter region includes sequences that modulate the recognition, binding and transcription initiation activity of the RNA polymerase. The UTR regions of the gene corresponds to the 5' and 3' sequences that are transcribed into (mature) mRNA but are not translated into protein.

The term "fragment" when used in connection with nucleotide sequences means any fragment of the nucleotide sequence consisting of at least 20 consecutive nucleotides of that sequence.

As used herein, the term "polymorphism" refers to the coexistence of more than one form of a gene or portion thereof. A portion of a gene of which there are at least two different forms, i. e., two different nucleotide sequences, is referred to as a "polymorphic region of a gene". A polymorphic region can be a single nucleotide, the identity of which differs in different alleles. Such polymorphism is referred herein as "single nucleotide polymorphism" or SNP. A polymorphic region also can be several nucleotides in length. A gene having at least one polymorphic region is referred to as a "polymorphic gene".

SNPs, which are known in the art, are identified herein with the numbers corresponding to the refSNP ID NOs (rs numbers) of the NCBI SNP database (<http://www.ncbi.nlm.nih.gov/SNP/>) and UCSC Genome SNP database (<http://www.genome.ucsc.edu/>), for example such as

rs11561, rs5758405 rs8779, rs132806,
rs2068943, rs2267487
rs881542, rs926333,rs1060387
rs1006407
rs4468, rs138855, rs2239848, rs138880, rs138881
rs20551, rs2294976, rs2076578, rs1046088
rs133068, rs133069

rs133070, rs1330739.

By the term "SNP type" is meant the promoter, UTR, intron, synonymous or non-synonymous SNP.

5

By "promoter SNP" is meant an SNP located in the promoter region of a gene. This type of SNP may affect expression of the gene.

10 By the term "UTR SNP" is meant a SNP located in part of the genome that is transcribed into mRNA, but this part of the mRNA is not translated into protein. SNPs in this part of the genome may for example affect splicing, regulation of transcription, the fate of the mRNA in the cell changing the stability or the location where the mRNA is transported.

15 By the term "Intron SNP" is meant a SNP located in an intronic region of the gene. Introns are spliced out of the mature mRNA before it leaves the cell nucleus. This type SNP may affect splicing of the primary transcript into the mature mRNA and thereby influence what is translated into protein (e.g. exon-skipping leading to defect protein, or change of reading-frame leading to truncated protein, or nonsense-mediated-decay). Intron SNPs may also have a regulatory potential.

20 By the term "synonymous SNP" (syn) is meant a SNP which is associated with the change of a nucleotide in the (coding) exon sequence that doesn't lead to the change of an amino acid in the protein encoded by the gene. A syn-SNPs may have an impact on the regulation of the gene and on splicing e.g. by introduction of cryptic splice sites

30 The term "non-synonymous SNP" (non-syn) designates a SNP which is associated with a nucleotide change in the coding DNA sequence that leads to the change of an amino acid in the sequence of the encoded protein. This type of SNP may lead to a change in the function and/or structure of the protein.

35 As used herein, "allele", which is used interchangeably herein with "allelic variant" refers to alternative forms of a gene or portions thereof. Alleles occupy the same locus or position on homologous chromosomes. When an individual has two identical alleles of a gene, the individual is said to be homozygous for the gene or allele.

When an individual has two different alleles of a gene, the individual is said to be heterozygous for the gene or alleles. Alleles of a specific gene can differ from each other in a single nucleotide, or several nucleotides, and can include substitutions, deletions, and insertions of nucleotides. An allele of a gene also can be a form of a gene containing a mutation.

By the term "susceptibility/risky allele" is in the present content meant the allele of a polymorphism, e. g. SNP, which is associated with a predisposition of an individual carrying this allele for development of a disease, e. g. SCH and/or BPD. By the term "protective allele" is meant the allele of a polymorphism, e. g. SNP, said polymorphism being associated with a predisposition of an individual for development of a disease, the presence of which in the genetic material from an individual, e. g. a DNA sample, indicates no predisposition of an individual carrying this allele for development of a disease, e. g. SCH and/or BPD.

As used herein, "predisposition" means that an individual having a particular genotype and/or haplotype has a higher likelihood than one not having such a genotype and/or haplotype for a particular condition/disease as one of the described herein.

As used herein, the term "haplotype" refers to a set of closely positioned alleles present on one chromosome which tend to be inherited together (not easily separable by recombination).

As used herein, the term "genetic marker" refers to an identifiable physical location on a chromosome (e.g., single nucleotide polymorphism (SNP), restriction enzyme cutting site) whose inheritance can be monitored. Markers can be expressed regions of DNA (genes) or some segment of DNA with no known coding function but whose pattern of inheritance can be determined.

As used herein, the term "linkage" refers to an association in inheritance between genetic markers such that the parental genetic marker combinations appear among the progeny more often than the non-parental.

As used herein, the term "linkage disequilibrium" (LD) means that the observed frequencies of haplotypes in a population does not agree with haplotype frequencies

predicted by multiplying the frequencies of individual genetic marker alleles in each haplotype; LD means that there exist correlations among neighbouring alleles, reflecting 'haplotypes' descended from single, ancestral chromosomes.

- 5 Target nucleic acid: a nucleic acid isolated from an individual and comprising at least one polymorphism identified in the present invention as well as further nucleotides upstream or downstream. The target nucleic acid can be used for hybridisation, for sequencing or other analytical purposes.
- 10 By the term "gene products" is meant herein products of gene transcription, such as an mRNA transcript and said mRNA transcript splicing products, and products of gene translation, such as polypeptide(s) translated from any of the gene mRNA transcripts and various products of post-translational processing of said polypeptides, such as the products of post-translational proteolytic processing of the polypeptide(s)
- 15 or products of various post-translational modifications of said polypeptide(s).

Alignment. When reference is made to alignment of protein sequences alignment is carried out using the MultAlin algorithm with default settings ("Multiple sequence alignment with hierarchical clustering", F Corpet, 1988, Nucl. Acids Res., 16 (22), 10881-10890), which is available at the internet address:

20 <http://prodes.toulouse.inra.fr/multalin/multalin.html>.

Conservative amino acid substitutions:

Substitutions within the groups of amino acids are considered conservative amino acid substitutions. Substitutions among the groups of amino acids are considered

25 non-conservative amino acid substitutions.

P, A, G, S, T (neutral, weakly hydrophobic)

Q, N, E, D, B, Z (hydrophilic, acid amine)

H, K, R (hydrophilic, basic)

30 F, Y, W (hydrophobic, aromatic)

L, I, V, M (hydrophobic)

C (cross-link forming)

SCHIZOPHRENIA (SCH) A mental disorder or heterogeneous group of mental disorders (the schizophrenias or schizophrenic disorders) comprising most major

35 psychotic disorders and characterized by disturbances in form and content of thought

(loosening of associations, delusions and hallucinations) mood (blunted, flattened or inappropriate affect), sense of self and relationship to the external world (loss of ego boundaries, dereistic thinking and autistic withdrawal) and behavior (bizarre, apparently purposeless and stereotyped activity or inactivity).

5

BIPOLAR AFFECTIVE DISORDER (BPD) is one of the most common, severe, and persistent mental diseases. It is characterized by periods of deep, prolonged, and profound depression that alternate with periods of excessively elevated and/or irritable mood known as mania. The symptoms of mania include a decreased need
10 for sleep, pressured speech, increased libido, reckless behaviour without regard for consequences, grandiosity, and severe thought disturbances, which may or may not include psychosis.

Description of drawings

15

Figure 1 presents the results of statistical evaluation of the coincidence of the presence of specific haplotypes comprising the SNPs of the BRD1 gene and occurrence of the SCH, BPD or combined SCH and BPD phenotype.

20

Figure 2 presents the results of statistical evaluation of the coincidence of the presence of specific haplotypes comprising the SNPs of the GPR24 gene and selected SNPs positioned on chromosome 22q and occurrence of the SCH, BPD or combined SCH and BPD phenotype.

25

Figure 3 illustrates one way to connect the cases from the Faroe Islands to a common ancestor. The sex of the individuals in the pedigree is not distinguished.

30

Figure 4 shows the statistical analysis of the observed distribution of the overall 1-5 marker haplotypes in the sample from the Faeroe Islands calculated using the program CLUMP. Significant associations with p-values less than 0.05 and less than 0.01 are shaded in light gray and dark grey, respectively.

35

Figure 5 presents the degree of pair-wise Linkage Disequilibrium (LD) between the SNPs genotyped in the Scottish case-control sample calculated using LDmax from the GOLD software package..

Figure 6 demonstrates the results of association analysis of haplotypes comprising 1-5 SNPs (BPD and SCH vs. controls)

Figure 7 demonstrates the results of association analysis of haplotypes comprising 1-5 SNPs (SCH vs. controls)

Figure 8 demonstrates the results of association analysis of haplotypes comprising 1-5 SNPs (BPD vs. controls)

Fig. 9: BRD1 immunostaining of adult rat, rabbit and human cerebral cortex. The 3 species exhibit a similar BRD1 staining pattern. (A) Rat cerebral cortex layer I-VI, (B) Rabbit cerebral cortex layer II-VI, (C) Human cerebral cortex layer II-V, (D) Rat layer II neurons, (E) Rabbit layer VI neurons, (F) Human layer V neurons.

Fig. 10: Confocal microscopic images of BRD1 immunofluorescence stained rat cerebral cortex layer II-III neurons (A), TO-PRO-3 stained cell nuclei in the same area (B), BRD1 and TO-PRO-3 double-stained section (C). The BRD1-immunoreactivity is located both in the cytosol and the nucleus of the neurons whereas surrounding glial cells with small TO-PRO-3 positive nuclei seem BRD1-negative.

Fig. 11: BRD1-positive neurons in different adult rat and rabbit CNS areas. (A) BRD1-positive pyramidal cells in CA1 of the rabbit hippocampus. Note the prominent staining of the proximal pyramidal dendrites in stratum radiatum. (B) Rabbit striatum. (C) Rat ventromedial hypothalamic nucleus. (D) Rat cerebellar cortex, note BRD1-staining of granule cells, purkinje cells, and molecular layer interneurons. (E) Rat brainstem trigeminal motor nucleus. (F) Caudal part of the trigeminal spinal nucleus in the rat cervical spinal cord.

Fig. 12: BRD1 immunostaining of subventricular myelencephalon from fetal pig of embryonic day 60 (A-C) and day of birth/embryonic day 115 (D-F). (A) Next to the lumen of the 4th ventricle, a cell rich neuroepithelial layer is noted wherefrom maturing neuroblasts migrate into the surrounding part of the myelencephalon. (B) Close view of the neuroepithelial layer seen in A. The neuroepithelial cells display prominent nuclear BRD1 immunopositivity. (C) Close view of maturing neuroblasts seen in A. Prominent BRD1-staining is seen both in the nucleus and the surrounding

cytosol. (D) Subventricular medulla at the day of birth corresponding to the area depicted in A, displays less prominent BRD1 immunoreactivity. (E) Close view of the subventricular layer seen in D. The subventricular cells display a weaker nuclear BRD1-immunoreactivity than seen in B. (F) Close view of the medullar neurons seen in D. Note the pronounced decrease in nuclear BRD1 immunoreactivity.

Fig. 13: Relative quantification of *BRD1* mRNA expression in terms of fold changes ($2^{-\Delta\Delta C_T}$ and $2^{-\Delta C_T}$) found by analysis of variance (left panel) contrasting to embryonic day 115 the mean threshold difference from each of the other days. The significance of the contrasts is indicated (*: $p < 0.05$, **: $p < 0.01$ and ***: $p < 0.001$) with p-values adjusted for the number of contrasts tested using the Sidák method. A polynomial regression model up to third order (right panel) was determined using a forward inclusion stepwise procedure. The significance of the highest order term is indicated.

15 Detailed description of the invention

Gene Polymorphism

In the first aspect the invention relates to a method for determining a predisposition to a mental disease in a subject comprising determining in a biological sample isolated from said subject one or polymorphisms in the chromosome regions containing the NHP2L1, PACSIN2, SERHL, PIPPIN, BRD1, EP300, FAM19A5 and/or GPR24 genes, or in a translational or transcriptional product from said regions, said polymorphism being indicative of said predisposition.

25 The invention relates to a mental disease being a disease characterised by at least two symptoms, preferably three or more symptoms, which are related to schizophrenia (SCH) and/or bipolar disorder (BPD). SCH and BPD are preferred mental diseases of the invention. In some embodiments a preferred mental disease is characterised by symptoms of both SCH and BPD.

30

Position of polymorphisms

In one embodiment the present invention relates to polymorphisms of the above identified genes, wherein the polymorphisms are located in the non-coding regions of the genes, such as the regions controlling expression of the genes, for example promoter and UTR regions, and the regions controlling the splicing of the gene transcript, such as introns or exon/intron boundaries. Such polymorphisms according

to the invention may influence expression of the gene or affect the splicing or maturation of the gene transcript, mRNA.

In another embodiment the invention relates to polymorphisms located in the coding regions of the gene, such as exons. Such polymorphisms according to the invention may lead to the production of variant proteins. Variant proteins are the proteins, amino acid sequence of which contains an amino acid change, such as substitution(s), insertion(s) and/or deletion(s), corresponding to the polymorphism(s) of the gene. A variant protein may have functional activity altered due to the latter gene polymorphism.

In one aspect the present invention relates to a method comprising the determination of one or more polymorphisms in the chromosome regions, said regions containing the NHP2L1, PACSIN2, SERHL, PIPPIN, BRD1, EP300, FAM19A5 and/or GPR24, and relating said one or more polymorphisms to the predisposition to SCH and BPD. The polymorphisms may be located either/both in the coding region and/or non-coding region of said genes. Another aspect of the invention relates to a method for determining one or more polymorphisms in the chromosome regions, which are in linkage disequilibrium with the polymorphisms of the above genes, and relating said one or more polymorphisms to a predisposition to SCH and/or BPD. The polymorphisms may be located in one individual gene as well as in two or more different individual genes. The polymorphism(s) may be located either or both in the coding and/or non-coding regions of these genes.

In further embodiments a predisposition to a mental disease comprises determining more than one polymorphism in any identified herein genes, or may be determined specifically for a selected disease by determining two or more selected polymorphisms of the invention, wherein said polymorphisms have stronger correlation with said selected disease than other polymorphisms identified in the present application. Thus, the invention relates to determining specific haplotypes comprising two or more identified herein polymorphisms which are associated with a specific mental disease, in particular SCH or BPD. The specific haplotypes of the invention may further comprise one or more polymorphisms of the DNA sequences adjusted to the genes of the invention.

Preferably, the invention relates to polymorphism of DNA being an SNP.

Preferred SNPs according to the invention are the SNPs having refSNP Nos.

rs11561, rs5758405 rs8779, rs132806,

rs2068943, rs2267487

5 rs881542, rs926333,rs1060387

rs1006407, rs6002408

rs4468, rs138855, rs2239848, rs138880, rs138881

rs20551, rs2294976, rs2076578, rs1046088

rs133068, rs133069, rs133070, rs133073

10 The above identified group of SNPs consists of the SNPs identified in the genomic sequences of the NHP2L1, PACSIN2, SERHL, PIPPIN, BRD1, EP300, FAM19A5 and/or GPR24 (SEQ ID NO: 1-7 and 94).

Positions of the SNPs within the genomic sequences of the genes (SEQ ID NOS: 1-7 and 94) are identified in Table 1.

Table 1

Gene	SEQ ID NO	Chromosome ID	SNP ID	Nucleotide No (position of SNP*)	SNP	Type of SNP
NHP2L1	1	22q13.2	rs8779	40313315	C/T	UTR
NHP2L1	1	22q13.2	rs132806	40313832	C/T	UTR
NHP2L1	1	22q13.2	rs11561	40314236	T/C	Non-syn
NHP2L1	1	22q13.2	rs5758405	40319420	C/A	Intron
PACSIN2	2	22q13.2	rs2068943	41587252	C/T	Promoter
PACSIN2	2	22q13.2	rs2267487	41642330	C/T	Promoter
SERHL	3	22q13.2	rs881542	41126941	C/G	Prom
SERHL	3	22q13.2	rs926333	41128686	A/G	Non-syn
SERHL	3	22q13.2	rs1060387	41136001	C/T	Non-syn
PIPPIN	4	22q13.2	rs6002408	40210909	C/T	UTR
PIPPIN	4	22q13.2	rs1006407	40215071	C/T	UTR
BRD1	5	22q13.3	rs4468	48386988	T/C	UTR
BRD1	5	22q13.3	rs138855	48417818	G/C	Intron
BRD1	5	22q13.3	rs2239848	48436090	G/A	Syn
BRD1	5	22q13.3	rs138880	48437947	A/C	Promoter
BRD1	5	22q13.3	rs138881	48439818	G/A	Promoter
EP300	6	22q13.2	rs20551	39781047	A/G	Non-syn
EP300	6	22q13.2	rs2294976	39807747	C/A	Intron
EP300	6	22q13.2	rs2076578	39812648	C/T	Intron
EP300	6	22q13.2	rs1046088	39817422	A/C	Non-syn
GPR24	7	22q13.2	rs133068	39317446	C/G	Promoter
GPR24	7	22q13.2	rs133069	39317501	A/C	Promoter
GPR24	7	22q13.2	rs133070	39317812	A/G	Promoter
GPR24	7	22q13.2	rs133073	39318734	C/T	Syn
FAM19A	84	22q13.3	Rs132234	47424523	C/T	Intron

5							
FAM19A	84	22q13.3	Rs3752466	47466709	C/T	3'UTR	
5							

According to the invention the above SNPs are genetic markers of a mental disease of the invention. * Nucleotide position in a genomic sequence is given according to the sequences of the UCSC database, July 2003 assembly
 5 (<http://www.genome.ucsc.edu/>).

The invention describes several haplotypes of the above SNPs in association with SCH and/or BPD.

The following is non-limited examples of specific haplotypes of the invention:

- 10 A three SNP haplotype consisting of the SNPs rs133069, rs133070, rs133073 of the GRP24 gene is according to invention indicative of BPD,
 A four SNP haplotype consisting of rs133069, rs133070, rs133073 rs133068 of the GRP24 gene is according to invention indicative of SCH.

More examples of the SNPs and haplotypes of the invention are shown in Figures 1
 15 and 2 of the present application and in Table 2 below.

Table 2.

Gene	SNP ID NO	SNP	
BRD1	rs4468	T/C	C over-represented in SCH/Part of haplotype associated with SCH and combined BPD&SCH
BRD1	rs138855	G/C	Part of haplotype associated with SCH and BPD
BRD1	rs2239848	G/A	Part of haplotype associated with SCH, BPD and combined BPD&SCH
BRD1	rs138880	A/C	C over-represented in BPD and SCH /Part of haplotype associated with SCH, BPD and combined BPD&SCH
BRD1	rs138881	G/A	Part of haplotype associated with SCH, BPD and combined BPD&SCH

EP300	rs20551	A/G	A over-represented in BPD/part of haplotype associated with BPD
EP300	rs2294976	C/A	part of haplotype associated with BPD
EP300	rs2076578	C/T	part of haplotype associated with BPD
EP300	rs1046088	A/C	A over-represented in BPD/part of haplotype associated with BPD

GPR24	rs133068	C/G	Part of haplotype associated with SCH, BPD and Combined BPD&SCH
GPR24	rs133069	A/C	Part of haplotype associated with SCH, BPD and Combined BPD&SCH
GPR24	rs133070	A/G	Part of haplotype associated with SCH, BPD and Combined BPD&SCH
GPR24	rs133073	C/T	Part of haplotype associated with SCH, BPD and Combined BPD&SCH

NHP2L1	rs8779	C/T	Part of haplotype associated with SCH, BPD and combined BPD&SCH
NHP2L1	rs132806	C/T	Part of haplotype associated with SCH
NHP2L1	rs11561	T/C	C over-represented in SCH /Part of haplotype associated with SCH
NHP2L1	rs5758405	C/A	Part of haplotype associated with SCH

PACSIN2	rs2068943	C/T	C over-represented in SCH /part of haplotype associated with SCH
PACSIN2	rs2267487	C/T	part of haplotype associated with SCH

SERHL	rs881542	C/G	C over-represented in SCH
SERHL	rs926333	A/G	A over-represented in SCH/Part of haplotype associated with SCH
SERHL	rs1060387	C/T	Part of haplotype associated with SCH

PIPPIN	rs6002408	C/T	Part of haplotype associated with SCH, BPD and Combined BPD&SCH
PIPPIN	rs1006407	C/T	Part of haplotype associated with SCH, BPD and Combined BPD&SCH

5

The invention features haplotypes of the above SNPs that are present on one chromosome, e. g. within the sequence of one particular gene, or located within the sequences of two different juxtaposed genes, which tend to be inherited together. The latter group of SNPs may for example consist of the SNPs having refSNP ID NO:

10 rs909660, rs710193, rs1573745, rs132234, rs3752466, rs6010260, rs137931, rs137932, rs3810971, rs2272843, rs1063900, rs715519, rs916005. This group of SNPs includes the SNPs which are identified in the present invention as parts of different haplotypes, said haplotypes being associated with SCH and/or BPD (as for example shown in Figure 1, 2 and 4 of the present application.

15

In another aspect the invention relates to polymorphisms located in the chromosome regions, which do not contain the above identified genes, said polymorphisms being in linkage disequilibrium with at least one of the above identified SNPs. In particular, the invention relates to polymorphisms in the human chromosome 22q being in

20 linkage disequilibrium with one or more polymorphisms in the NHP2L1, PACSIN2, SERHL, PIPPIN, BRD1, EP300, FAM19A5 and/or GPR24 genes, such as an SNP selected from rs11561, rs5758405 rs8779, rs132806, rs2068943, rs2267487, rs881542, rs926333, rs1060387, rs1006407, rs4468, rs138855, rs2239848, rs138880, rs138881, rs20551, rs2294976, rs2076578, rs1046088, rs133068,

25 rs133069, rs133070, rs133073, rs6002408 as well as the microsatellites D22S922 and D22S1169

Likewise, any polymorphism of the genes being adjacent to the genes of the invention, in particular a polymorphism(s) being located within the sequence of 1 to 1 000 000 nucleotides adjacent to said genes, and is in linkage disequilibrium with any of the SNPs identified above, is in the scope of the invention. The SNPs located
 5 within 50-100 000 nucleotides sequence adjusted to the nucleotide corresponding to the first or the last nucleotide of a sequence selected from the sequences identified as SEQ ID NO: 1-7 and 94 are preferred.

Gene Products

10 The invention relates to a method for determining a predisposition to a mental disease comprising determining at least one polymorphism in any of the above genes or in transcriptional or translational products of the genes.

As used herein, the term "transcriptional product of the gene" refers to a pre-messenger RNA molecule, pre-mRNA, that contains the same sequence information
 15 (albeit that U nucleotides replace T nucleotides) as the gene, or mature messenger RNA molecule, mRNA, which was produced due to splicing of the pre-mRNA, and is a template for translation of genetic information of the gene into a protein.

20 As used herein, the term "translational product of the gene" refers to a protein, which is encoded by the gene.

Thus, the invention includes in the scope of protection nucleic acids comprising the coding nucleotide sequences of the above genes comprising a polymorphism and
 25 proteins comprising a polymorphism corresponding to the polymorphism of the encoding nucleic acid sequence.

In particular, the invention relates to transcriptional products of the above genes being

- 30 (i) nucleic acid sequences identified in the invention as SEQ ID NO: 8-14, or fragments thereof,
- (ii) nucleic acid sequences having at least 90% identity with SEQ ID NO: 8-14, or fragments thereof,
- (iii) nucleic acid sequences being complementary to any of the sequences of
 35 (i) or (ii),

said nucleic acid sequences comprising the polymorphisms of the genomic sequences described above.

Translational products of the genes of the invention are identified herein as

- 5 (i) variant proteins corresponding to the proteins identified in the NCBI database under Ass. Nos.: NP_001003796/004999 (NHP2L1) (SEQ ID NO: 87), NP_009160 (PACSIN2) (SEQ ID NO: 88) , NP_733795 (SERHL) (SEQ ID NO: 89), NP_055275 (PIPPIN) (SEQ ID NO: 90), NP_055392 (BRD1) (SEQ ID NO: 91), NP_001420 (EP300) (SEQ ID NO: 92),
10 NP_005288 (GPR24) (SEQ ID NO: 93), or variants, or fragments thereof,
(ii) polypeptide sequences having at least 90% identity with the variant proteins, or fragments thereof, of (i),

15 said variant proteins, fragments thereof and said polypeptide sequences are comprising polymorphism corresponding to the polymorphism of the corresponding genomic sequences or transcriptional products of said genomic sequences.

20 Selected, but non-limited examples of protein polymorphism of the invention are given in Table 3 below:

Table 3

Gene	SNP ID.	Alleles	Protein polymorphism
EP300	rs1046088	A/C	Gln2223Pro
EP300	rs20551	A/G	Ile997Val
SERHL	rs926333	A/G	Ser46Ala
SERHL	rs1060387	C/T	Ala2Val
NHP2L1	rs11561	T/C	Thr43Ala

25 Thus, it is an embodiment of the invention to use the above identified variant proteins for the purpose of

- i) diagnosis of SCH and/or BPD in a neuropsychiatric patient, and/or
ii) prognosis of likelihood of development of SCH and/or BPD by an individual, and/or
30 iii) development new drug candidates for the treatment of SCH and/or BPD, and/or

- iv) therapeutic treatment of SCH and/or BPD.

Methods of determining polymorphisms

1 SNP

- 5 Many methods (see Table 4 below) are known in the prior art for determining the presence of particular nucleotide sequences or for determining particular proteins having particular amino acid sequences. All of these methods may be adapted for determining the polymorphisms according to the present invention.

Table 4.

Method	Result
Restriction fragment length polymorphism	Cleavage or non-cleavage based on SNP results in difference in length
Amplified fragment length polymorphism	Cleavage or non-cleavage based on SNP results in difference in length
Mass spectrometry	Difference in molecular weight of hybrids between a probe and the different alleles
Single strand conformation polymorphism (SSCP). SSCP heteroduplex.	Different separation in gel based on different conformation caused by single nucleotide polymorphism.
Single nucleotide extension	Difference in signal through incorporation of differently labelled nucleotide or labelled/non-labelled nucleotide
Sequencing	Difference in sequence
Hybridisation	Hybridisation or non-hybridisation at high stringency. Often detected by using differently labelled probes.
Determination of T_m profile	Difference in T_m profile between target and homologous vs. non-homologous probe.
Cleavage of single-stranded DNA	
Denaturing HPLC	DHPLC is based on resolving heteroduplex from homoduplex DNA fragments produced by PCR amplification using temperature-modulated heteroduplex analysis.

TAQMAN	PCR based technique.
--------	----------------------

One common method for detecting SNPs comprises the use of a probe bound to a detectable label. By carrying out hybridisation under conditions of high stringency it is ensured that the probe only hybridises to a sequence which is 100% complementary to the probe. According to the present invention this method comprises hybridising a probe to a target nucleic acid sequence comprising at least one of the SNPs at the positions identified in Table 1 (see above). For other polymorphisms or mutations within the defined region, similar probes can be designed by the skilled practitioner and used for hybridisation to a target nucleic acid sequence. The design and optimisation of probes and hybridisation conditions lies within the capabilities of the skilled practitioner.

In the scope of the present invention the term "hybridisation" signifies hybridisation under conventional hybridising conditions, preferably under stringent conditions, as described for example in Sambrook et al., Molecular Cloning, A Laboratory Manual, 2nd Edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.). The term "stringent" when used in conjunction with hybridisation conditions is as defined in the art, i.e. 15-20°C under the melting point T_m , cf. Sambrook et al, 1989, pages 11.45-11.49. Preferably, the conditions are "highly stringent", i.e. 5-10°C under the melting point T_m . Under highly stringent conditions hybridisation only occurs if the identity between the oligonucleotide sequence and the locus of interest is 100 %, while no hybridisation occurs if there is just one mismatch between oligonucleotide and DNA locus. Such optimised hybridisation results are reached by adjusting the temperature and/or the ionic strength of the hybridisation buffer as described in the art. However equally high specificity may be obtained using high-affinity DNA analogues. One such high-affinity DNA analogues has been termed "locked nucleic acid" (LNA). LNA is a novel class of bicyclic nucleic acid analogues in which the furanose ring conformation is restricted in by a methylene linker that connects the 2'-O position to the 4'-C position. Common to all of these LNA variants is an affinity toward complementary nucleic acids, which is by far the highest reported for a DNA analogue (Ørum et al. (1999) Clinical Chemistry 45, 1898-1905; WO 99/14226 EXIQON). LNA probes are commercially available from Proligo LLC, Boulder, Colorado, USA. Another high-affinity DNA analogue is the so-called protein nucleic acid (PNA). In PNA compounds, the sugar backbone of an oligonucleotide is replaced with an amide containing backbone, in particular an aminoethylglycine

backbone. The nucleobases are retained and are bound directly or indirectly to aza nitrogen atoms of the amide portion of the backbone (*Science* (1991) 254: 1497-1500).

- 5 Various different labels can be coupled to the probe. Among these fluorescent reporter groups are preferred because they result in a high signal/noise ratio.

Suitable examples of the fluorescent group include fluorescein, Cy2, Cy3, Cy3.5, Cy5, Cy5.5, Cy7, acridin, Hoechst 33258, Rhodamine, Rhodamine Green,
10 Tetramethylrhodamine, Texas Red, Cascade Blue, Oregon Green, Alexa Fluor, europium and samarium.

Another type of labels are enzyme tags. After hybridisation to the target nucleic acid sequence a substrate for the enzyme is added and the formation of a coloured
15 product is measured. Examples of enzyme tags include a beta-Galactosidase, a peroxidase, horseradish peroxidase, a urease, a glycosidase, alkaline phosphatase, chloramphenicol acetyltransferase and a luciferase.

A further group of labels include chemiluminescent group, such as hydrazides such
20 as luminol and oxalate esters.

A still further possibility is to use a radioisotope and detect the hybrid using scintillation counting. The radioisotope may be selected from the group consisting of ^{32}P , ^{33}P , ^{35}S , ^{125}I , ^{45}Ca , ^{14}C and ^3H .
25

One particularly preferred embodiment of the probe based detection comprises the use of a capture probe for capturing a target nucleic acid sequence. The capture probe is bound to a solid surface such as a bead, a well or a stick. The captured target nucleic acid sequence can then be contacted with the detection probe under
30 conditions of high stringency and the allele be detected.

One embodiment of the probe based technique based on TAQMAN technique. This is a method for measuring PCR product accumulation using a dual-labeled fluorescent oligonucleotide probe called a TAQMAN® probe. This probe is composed
35 of a short (ca. 20-25 bases) oligodeoxynucleotide that is labeled with two different fluorescent dyes. On the 5' terminus is a reporter dye and on the 3' terminus is a

quenching dye. This oligonucleotide probe sequence is homologous to an internal target sequence present in the PCR amplicon. When the probe is intact, energy transfer occurs between the two flourophors and emission from the reporter is quenched by the quencher. During the extension phase of PCR, the probe is cleaved
 5 by 5' nuclease activity of Taq polymerase thereby releasing the reporter from the oligonucleotide-quencher and producing an increase in reporter emission intensity.

Other suitable methods include using mass spectrometry, single base extension, determining the T_m profile of a hybrid between a probe and a target nucleic acid
 10 sequence, using single strand conformation polymorphism, using single strand conformation polymorphism heteroduplex, using RFLP or RAPD, using HPLC, using sequencing of a target nucleic acid sequence from said biological sample.

Denaturing high-performance liquid chromatography (DHPLC) has been proven
 15 useful in human and animal genetic studies for detecting single nucleotide polymorphisms (SNPs). In contrary to most SNP detection methods that are currently in used, SNP detection by DHPLC is not based on a re-sequencing strategy that is expensive to implement, nor does it require gel-based genotyping procedures. Instead, SNP detection by DHPLC is based on resolving heteroduplex from
 20 homoduplex DNA fragments produced by PCR amplification using temperature-modulated heteroduplex analysis.

In connection with several of these methods there is a need for amplifying the amount of target nucleic acid in the biological sample isolated from the subject.
 25 Amplification may be performed by any known method including methods selected from the group consisting of polymerase chain reaction (PCR), Ligase Chain Reaction (LCR), Nucleic Acid Sequence-Based Amplification (NASBA), strand displacement amplification, rolling circle amplification, and T7-polymerase amplification.

30 More particularly, PCR-based amplification can be carried out using for example a primer pair comprising appropriate sequences selected from the sequences identified in Table 5 below.

Table 5

Gene	SNP Rs ID No.	Primer	SEQ ID NO
NHP2L1	rs8779 F	AGC AGC ATA CCA AAG AGA TG	15

		R	ACA CAT GAA CTA GCA CTC TC	16
		snp	AAT GCT GCA GGA CTC CGG AGG C	17
	rs132806	F	TTC AAC GGT GCC ACC ACC	18
		R	ATC CAT TCA GCA GTC CAT TG	19
		snp	AAC TGA CTA AAC TAG GTG CCA CGT CGT	20
			GAA ACC TGT AAA ACA GAA CAA AAA	
	rs11561	F	TGT AGG GCT TGA GAA AAT GC	21
		R	CAG GTG CAG AAT GAT CTC C	22
		snp	AAC TGA CTA AAC TAG GTG CCA CGC CCC	23
			TGT TGA GGG TTT TGG	
	rs5758405	F	TTA CAT GAC TGC TGA ACG AG	24
		R	GTC CAT TGT TAA CTA CCT GG	25
		snp	CTC AGT CTA TGG GGG GGA	26
PIPPIN	rs6002408	F	CGA ACC ATA TTC TAG ACC ATC	27
		R	TCG ACT CTG AAG TCA TGG TG	28
		snp	GAG CCC GTG GCT GGT GAG GC	29
	rs1006407	F	CCA AGA TGT GTG TGC ACC C	30
		R	CGG CCC CAC TTC TCC CC	31
		snp	CCA GGA AGG AGG GCC CTG TC	32
EP300	Rs20551	F	ATC CTC CAT CTA CTA GTA GC	33
		R	GAA GTA CTT GGC TGG TCT TC	34
		snp	AAC TGA CTA AAC TAG GTG CCA CGT	35
			CGT GAA CAG ATA CGC AGC CGG AGG AT	
	Rs2294976	F	ATG GTT ACT ATT GGT GAT TCC	36
		R	TCG GTA TGG AAA GGA TTC TG	37
		snp	AAC TGA CTA AAC TAG GTG CCA CGT	38
			CGT GAA AGT CTG ACA AGT TAT GTT GTG	
	Rs2076578	F	GGA GAT ATT CTG TGC TAT TCC	39
		R	TGC TGT CTC CCT TGG TCA C	40
		snp	ACT GAC TAA ACT AGG TGC AAA CCT TAT	41
			TTT CTT GTC TC	
	Rs1046088	F	CAA CCA TAA CCA GTT CCA GC	42
		R	CAT ATT TCC TTG TTG CAT CTG	43
		snp	AAC TGA CTA AAC TAG AGT TGG CTA CCC	44
			ACC ACA GC	
PACSIN2	rs2068943	F	CAG CAG TCA GTG GGA CAG	45
		R	TCT TAG GAG ACA GTC GTG C	46
		snp	AAC TGA CTA AAC TAG GTG CCA CGT	47
			CGT GAA AGT CTG ACT TGC ACA GGA	
	rs2267487	F	TGA GTA GAG AGG CGG CTC	48
		R	GAT TAC CTT CCC TGA ACT TC	49
		snp	AAC TGA CTA AAC TAG GTG GCG AGC	50
BRD1	rs13881	F	ATC CGC CGA TTC CAC TAA C	51
		R	ACG AGT GAC ACG ATT GAG G	52
		snp	AAG GAG TTC TGA GAA TCC GTA G	53
	rs138880	F	TAC CTG GTC CTA TCG GAT G	54
		R	TTT CCA TGG TGG TCA CTG C	55
		snp	AAC TGA CAC AGG CCC TCC ACA CGC	56
			ACA	
	rs4468	F	TCT TTC CAC AAA GTA CCA TTC	57
		R	TCC AAG TTT TCT AGG AAC CC	58
		snp	AAC TGA CTA AAC TAG GTG CGT GAC	59
			GCC AGC AGC ACG GCC	
	rs138855	F	GAG GTA CAT GCG TGT AGT C	60
		R	GCT GAC ACA GAT GTG GTT G	61
		snp	AAC TGA CTA AAC TAG GGC ACT CCA	62
			GCC TGG GTG ACA	
	rs2239848	F	ACA CGG TCG GCA GGA CC	63

		R	AAA GAC CGC TTA CTG TGA TG	64
		snp	AAC TGA CAG ACG CCA TTT TTC ATT TC	65
SERHL	rs1060387	F	AAC AGG ACA TGG GCC TGC	66
		R	CCT TGT GAC GAG AAT TCA CC	67
		snp	AAC TGA CTA AAC TAG GTG CCA CGT	68
			CGT GAA AAC AGC ATT GAC TTC GTC AG	
	rs881542	F	TTT AGC CAT TTT GAC CAG GC	69
		R	CTG GTC TCG AAC TCT CAA C	70
		snp	AAC TGA CTA AAC TAG GTG CCA CGT	71
			CGT GAA AGT CTG ACA ATC CCA GCA	
	rs881542	F	TTT TGG GAG G	
		R	TGT GTC TTT CAG GTC TGA TC	72
		snp	GAA GAG GGA TGA GTC TGT C	73
			AAC TGA CTA AAC TAG GTG CCA CGT	74
GPR24	rs133068	F	CGT CGG CTG GCT GGA CAA TGC CA	
		R	AAG CAT GGT TCC TCA CTT CC	75
		snp	AGA TCC CTA CAT GTG TCC TC	76
			AAC TGA CTA AAG AGG GGT TTT CTG	77
	rs133069	F	TCT CCC C	
		R	AAG CAT GGT TCC TCA CTT CC	78
		snp	AGA TCC CTA CAT GTG TCC TC	79
			ATC CCC CAC CCC ACC CCC	80
	rs133070	F	CCT GAG TCT GGC AGT GGG	81
		R	CAC CCT CGG CCC CTT CC	82
		snp	AAC TGA CTA AAC TAG TCA GAT ATT	83
			TGT CTC AAA GA	
	rs133073	F	GGA GCT CAG CTC GGT TGT G	84
		R	GCT CTG CTC CTG CCC GTC	85
		snp	AAC TGA CTA AAC TAG GTG CCA CGT	86
			CGT CTG CTG CCC ACT GGT CCC AA	

F - forward PCR primer

R - reverse PCR primer

snp – primers for the single base extension detection method

5

One of the primers may comprise a moiety for subsequent immobilisation of the amplified fragments.

10

It is understood that the primers identified above may also be used as probes for determining the polymorphisms of the invention in a nucleic acid sequence using any of the methods known in the art and featured above.

15

To the extent that the polymorphisms as defined in the present invention are present in DNA sequences transcribed as mRNA transcripts these transcripts constitute a suitable target sequence for detection of the polymorphisms. Commercial protocols are available for isolation of total mRNA. Through the use of suitable primers the target mRNA can be amplified and the presence or absence of polymorphisms be

detected with any of the techniques described above for detection of polymorphisms in a DNA sequence.

2. Proteins

- 5 As discussed above, genetic polymorphism of the invention can also be detected as a polymorphism of a protein product of the gene, or a change in a biological response mediated by the protein.

10 The polymorphism located for example in the NHP2L1, SERHL or EP300 genes may also be detected by isolating the protein from a biological sample and determining the presence or absence of the mutated residue (according to Table 3 above) by sequencing said protein, or determining the presence or absence of another polymorphic amino acid located in a mutant gene. The polymorphism of any of the variant proteins of the invention may be detected likewise.

15

Isolated/identified variant proteins expressed by any of the other polymorphic genes of the invention are used as alternative diagnostic markers of the genetic polymorphism associated with a predisposition to a mental disease of the invention.

- 20 The invention also concerns using variant proteins described herein, or fragments or variants thereof, said fragments and variants comprising a polymorphism corresponding to the polymorphism of the variant protein, for the manufacture of an antibody, which may be used both in methods for diagnosis and methods of treatment of SCH and/or BPD.

25 ***Isolated oligonucleotides***

- In one aspect the invention relates to an isolated oligonucleotide comprising at least 10 contiguous nucleotides being 100% identical to a subsequence of the genes of the invention comprising or adjacent to a polymorphism or mutation being correlated to an mental disease, or being 100% identical to a subsequence of the human
30 genome which is in linkage disequilibrium with any of the genes of the invention comprising or adjacent to a polymorphism or mutation being correlated to a mental disease. As explained in the summary, such probes may be used for detecting the presence of a polymorphism of interest and/or they may constitute part of a primer pair and/or they may form part of a gene therapy vector used for treating the mental
35 diseases.

Preferably the isolated oligonucleotide comprises at least 10 contiguous bases of a sequence identified as SEQ ID NOs: 8-14 or the corresponding complementary strand, or a strand sharing at least 90% sequence identity more preferably at least 95% sequence identity with SEQ ID NOs: 8-14 or a complementary strand thereof, said isolated oligonucleotide comprising a polymorphism of the invention.

Further preferred isolated oligonucleotide may comprise at least 10 contiguous bases of any of the sequences identified as SEQ ID NOS: 1-7 and 94 or the corresponding complementary strand thereof, or a strand sharing at least 90% sequence identity more preferably at least 95% sequence identity with any of the SEQ ID NOS: 1-7 and 94 or a complementary strand thereof, said isolated oligonucleotide comprising a polymorphism of the invention.

These particular oligonucleotides may be used as probes for assessing the polymorphisms in the human BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 genes which are strongly correlated with mental diseases of the invention, such as SCH and/or BPD.

The length of the isolated oligonucleotide depends on the purpose. When being used for amplification from a sample of genomic DNA, the length of the primers should be at least 15 and more preferably even longer to ensure specific amplification of the desired target nucleotide sequence. When being used for amplification from mRNA the length of the primers can be shorter while still ensuring specific amplification. In one particular embodiment one of the pair of primers may be an allele specific primer in which case amplification only occurs if the specific allele is present in the sample. When the isolated oligonucleotides are used as hybridisation probes for detection, the length is preferably in the range of 10-15 nucleotides. This is enough to ensure specific hybridisation in a sample with an amplified target nucleic acid sequence. When using nucleotides which bind stronger than DNA (e.g. LNA and/or PNA), the length of the probe can be somewhat shorter, e.g. down to 7-8 bases.

The length may be at least 15 contiguous nucleotides, such as at least 20 nucleotides. An upper limit preferably determines the maximum length of the isolated oligonucleotide. Accordingly, the isolated oligonucleotide may be less than 1000 nucleotides, more preferably less than 500 nucleotides, more preferably less than

100 nucleotides, such as less than 75 nucleotides, for example less than 50 nucleotides, such as less than 40 nucleotides, for example less than 30 nucleotides, such as less than 20 nucleotides.

- 5 The isolated oligonucleotide may comprise from 10 to 50 nucleotides, such as from 10 to 15, from 15 to 20, from 20 to 25, or comprising from 20 to 30 nucleotides, or from 15 to 25 nucleotides.

- 10 Depending on the use the polymorphism may be located in the centre of the nucleic acid sequence, in the 5' end of the nucleic acid sequence, or in the 3' end of the nucleic acid sequence.

For detection based on single base extension the sequence of the oligonucleotide is adjacent to the mutation/polymorphism, either in the 3' or 5' direction.

15

The isolated oligonucleotide sequence may be complementary to a sub-sequence of the coding strand of a target nucleotide sequence or to a sub-sequence to the non-coding strand of a target nucleotide sequence as the polymorphism may be assessed with similar efficiency in the coding and the non-coding strand.

20

The isolated oligonucleotide sequence may be made from RNA, DNA, LNA, PNA monomers or from chemically modified nucleotides capable of hybridising to a target nucleic acid sequence. The oligonucleotides may also be made from mixtures of said monomers.

25

Antibody

- Antibodies directed against unimpaired (wild type) or mutant polypeptide products of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 genes or conserved variants or peptide fragments thereof, which are discussed, above, may also be used as compounds in the methods for diagnosis/prognosis of a BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene associated disorder or a neuropsychiatric disorder, in particular such as SCH and BPD.

- 35 Such methods may be used to detect abnormalities in the level of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene

encoded polypeptide synthesis or expression, or abnormalities in the structure, temporal expression, and/or physical location of BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 proteins. The antibodies and immunoassay methods described below have, for example, important in vitro applications in assessing the efficacy of treatments for the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 gene associated disorders, such as neuropsychiatric disorders, for example SCH and/or BDP. Antibodies, or fragments of antibodies, such as those described below, may be used to screen potentially therapeutic compounds in vitro to determine their effects on BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene expression and BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 protein production. The compounds that have beneficial effects on a BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene disorder, such as SCH and BPD, can be identified, and a therapeutically effective dose determined.

In vitro immunoassays may also be used, for example, to assess the efficacy of cell-based gene therapy for a BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene associated disorder, such as SCH and/or BPD. Antibodies directed against BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 polypeptides may be used in vitro to determine, for example, the level of BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene expression achieved in cells genetically engineered to produce BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 polypeptides. In the case of intracellular BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene products, including both transcriptional and translational products, such an assessment is done, preferably, using cell lysates or extracts. Such analysis will allow for a determination of the number of transformed cells necessary to achieve therapeutic efficacy in vivo, as well as optimization of the gene replacement protocol.

For the purpose of below discussion all molecules that produced due to activity of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 genes, such as transcriptional and translational products of the genes, are termed herein "gene products", if not specified otherwise.

The tissue or cell type to be analyzed will generally include those that are known, or suspected, to express the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 genes. The protein isolation methods employed herein may, for example, be such as those described in Harlow and Lane (1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York). The isolated cells can be derived from cell culture or from a patient. The analysis of cells taken from culture may be a necessary step in the assessment of cells to be used as part of a cell-based gene therapy technique or, alternatively, to test the effect of compounds on expression of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 genes.

Preferred diagnostic methods for the detection of BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene products, such as BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 proteins or conserved variants or peptide fragments thereof, may involve, for example, immunoassays wherein these products are detected by their interaction with an anti-BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24-specific antibody.

For example, antibodies, or fragments of antibodies, useful in the present invention may be used to quantitatively or qualitatively detect the presence of BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 proteins or conserved variants or peptide fragments thereof. This can be accomplished, for example, by immunofluorescence techniques employing a fluorescently labeled antibody coupled with light microscopic, flow cytometric, or fluorimetric detection. Such techniques are especially preferred for the proteins that are expressed on the cell surface.

The antibodies, or fragments thereof, useful in the present invention may, additionally, be employed histologically, as in immunofluorescence or immunoelectron microscopy, for in situ detection of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene derived polypeptide products or conserved variants or peptide fragments thereof. In situ detection may be accomplished by removing a histological specimen from a patient, and applying thereto a labelled antibody of the present invention. The antibody (or fragment) is preferably applied by overlaying the labeled antibody (or fragment) onto a biological sample. Through the use of such a procedure, it is possible to determine not only the

presence of BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 protein or conserved variants or peptide fragments, but also its distribution in the examined tissue. Using the present invention, those of ordinary skill will readily perceive that any of a wide variety of histological methods (such as staining
5 procedures) can be modified in order to achieve such in situ detection.

Immunoassays for BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 proteins or conserved variants or peptide fragments thereof will typically comprise incubating a sample, such as a biological fluid, a tissue extract,
10 freshly harvested cells, or lysates of cells, that have been incubated in cell culture, in the presence of a detectably labeled antibody capable of identifying BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 proteins or conserved variants or peptide fragments thereof, and detecting the bound antibody by any of a number of techniques well-known in the art.

15

The biological sample may be brought in contact with and immobilized onto a solid phase support or carrier such as nitrocellulose, or other solid support that is capable of immobilizing cells, cell particles or soluble proteins. The support may then be washed with suitable buffers followed by treatment with the detectably labeled BRD1,
20 NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 protein specific antibody. The solid phase support may then be washed with the buffer a second time to remove unbound antibody. The amount of bound label on solid support may then be detected by conventional means.

25 By "solid phase support or carrier" is intended any support capable of binding an antigen or an antibody. Well-known supports or carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, gabbros, and magnetite. The nature of the carrier can be either soluble to some extent or insoluble for the purposes of the present invention.

30 The support material may have virtually any possible structural configuration so long as the coupled molecule is capable of binding to an antigen or antibody. Thus, the support configuration may be spherical, as in a bead, or cylindrical, as in the inside surface of a test tube, or the external surface of a rod. Alternatively, the surface may be flat such as a sheet, test strip, etc. Preferred supports include polystyrene beads.

35 Those skilled in the art will know many other suitable carriers for binding antibody or antigen, or will be able to ascertain the same by use of routine experimentation.

The binding activity of an anti- BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 antibody may be determined according to well known methods. Those skilled in the art will be able to determine operative and optimal
5 assay conditions for each determination by employing routine experimentation.

One of the ways in which an BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24-specific antibody can be detectably labeled is by linking the same to an enzyme and use in an enzyme immunoassay (EIA) (Voller, A., "The
10 Enzyme Linked Immunosorbent Assay (ELISA)", 1978, Diagnostic Horizons 2, 1-7 and 94, Microbiological Associates Quarterly Publication, Walkersville, Md.); Voller, A. et al., 1978, J. Clin. Pathol. 31, 507-520; Butler, J. E., 1981, Meth. Enzymol. 73, 482-523; Maggio, E. (ed.), 1980, Enzyme Immunoassay, CRC Press, Boca Raton, Fla.,; Ishikawa, E. et al., (eds.), 1981, Enzyme Immunoassay, Kaku Shoin, Tokyo).
15 The enzyme which is bound to the antibody will react with an appropriate substrate, preferably a chromogenic substrate, in such a manner as to produce a chemical moiety that can be detected, for example, by spectrophotometric, fluorimetric or by visual means. Enzymes that can be used to detectably label the antibody include, but are not limited to, malate dehydrogenase, staphylococcal nuclease, delta-5-steroid
20 isomerase, yeast alcohol dehydrogenase, .alpha.-glycerophosphate, dehydrogenase, triose phosphate isomerase, horseradish peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, .beta.-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase and acetylcholinesterase. The detection can be accomplished by colorimetric methods that employ a chromogenic
25 substrate for the enzyme. Detection may also be accomplished by visual comparison of the extent of enzymatic reaction of a substrate in comparison with similarly prepared standards.

Detection may also be accomplished using any of a variety of other immunoassays.
30 For example, by radioactively labelling the antibodies or antibody fragments, it is possible to detect the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 polypeptides through the use of a radioimmunoassay (RIA) (see, for example, Weintraub, B., Principles of Radioimmunoassays, Seventh Training Course on Radioligand Assay Techniques, The Endocrine Society, March,
35 1986). The radioactive isotope can be detected by such means as the use of a gamma counter or a scintillation counter or by autoradiography.

It is also possible to label the antibody with a fluorescent compound. When the fluorescently labeled antibody is exposed to light of the proper wave length, its presence can then be detected due to fluorescence. Among the most commonly
5 used fluorescent labeling compounds are fluorescein isothiocyanate, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, o-phthaldehyde and fluorescamine.

The antibody can also be detectably labeled using fluorescence emitting metals such as ¹⁵²Eu, or others of the lanthanide series. These metals can be attached to
10 the antibody using such metal chelating groups as diethylenetriaminepentacetic acid (DTPA) or ethylenediaminetetraacetic acid (EDTA).

The antibody also can be detectably labeled by coupling it to a chemiluminescent compound. The presence of the chemiluminescent-tagged antibody is then
15 determined by detecting the presence of luminescence that arises during the course of a chemical reaction. Examples of particularly useful chemiluminescent labeling compounds are luminol, isoluminol, thionin, acridinium ester, imidazole, acridinium salt and oxalate ester.

20 Likewise, a bioluminescent compound may be used to label the antibody of the present invention. Bioluminescence is a type of chemiluminescence found in biological systems in which a catalytic protein increases the efficiency of the chemiluminescent reaction. The presence of a bioluminescent protein is determined by detecting the presence of luminescence. Important bioluminescent compounds for
25 purposes of labeling are luciferin, luciferase, green fluorescent protein and aequorin.

Biological sample

The biological sample used in the present invention may be any suitable biological sample comprising genetic material and/or proteins. In a preferred embodiment the
30 sample is a blood sample, a tissue sample, a secretion sample, semen, ovum, hairs, nails, tears, and urine. The most convenient sample type is a blood sample.

Kits

In one aspect there is provided a kit for predicting the risk of a subject for developing
35 mental diseases, such as SCH and/or BPD or for other diagnostic and classification

purposes of mental diseases such as SCH and/or BPD comprising at least one probe comprising a nucleic acid sequence as defined in the previous section.

In one embodiment the probe is linked to a detectable label.

5

In another embodiment based on single nucleotide extension the kit further comprises at least one nucleotide monomer labelled with a detectable label, a polymerase and suitable buffers and reagents.

- 10 The kit preferably also comprises set of primers for amplifying a region comprising at least one of the identified above polymorphisms in any of the genes selected from the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 genes or transcriptional products of said genes, or the corresponding complementary strands. The primers preferably are at least 15 bases long and may be coupled to an
- 15 entity suitable for subsequent immobilisation.

In another embodiment a diagnostic kit of the invention may comprise an antibody as described above.

20 Diseases

According to the invention the association of an SNP with a particular disease means the association a particular allele at the position of said SNP with a predisposition to said disease or with a protection against said disease. Non-limited examples of the protective/susceptibility alleles of the SNP having the positions as of Table 1 are

25 given in Table 6 below.

Table 6

Gene	SEQ ID NO	SNP	SNP ID	Allele	
				protective	susceptibility
NHP2L1	1	A/G	rs11561	A	G in SCH
PACSIN2	2	C/T	rs2068943	T	C in SCH
SERHL	3	C/G	rs881542	G	C in SCH
	3	A/G	rs926333	G	A in SCH
BRD1	5	T/C	rs4468	T	C in SCH
	5	A/C	rs138880	A	C in SCH
EP300	6	A/G	rs20551	G	A in BPD
	6	A/C	rs1046088	C	A in BPD

According to the invention individuals carrying the protective alleles at positions identified herein are less likely to develop a mental disease of the invention. In

contrary, the presence of the susceptibility allele is indicative of a predisposition to a mental disease of the invention.

Thus, in one embodiment the invention relates to a method for determining a
5 predisposition of an individual to a mental disease, in particular SCH and/BDP,, said
method comprising determining the presence in a biological sample from said
individual the susceptibility allele of at least one SNP selected from the SNPs
identified above as rs11561, rs5758405 rs8779, rs132806, rs2068943, rs2267487,
rs881542, rs926333, rs1060387, rs1006407, rs4468, rs138855, rs2239848,
10 rs138880, rs138881, rs20551, rs2294976, rs2076578, rs1046088, rs133068,
rs133069, rs133070, rs133073, rs6002408.

In another embodiment the invention relates to a method for determining a
predisposition of an individual to a mental disease, in particular SCH and/BDP, said
15 method comprising determining the presence in a biological sample from said
individual the susceptibility allele of two or more of the SNPs identified above as
rs11561, rs5758405 rs8779, rs132806, rs2068943, rs2267487, rs881542, rs926333,
rs1060387, rs1006407, rs4468, rs138855, rs2239848, rs138880, rs138881, rs20551,
rs2294976, rs2076578, rs1046088, rs133068, rs133069, rs133070, rs133073,
20 rs6002408.

In still another embodiment, the invention relates to a method for determining a
predisposition of an individual to a mental disease, in particular SCH and/BDP, said
method comprising determining the presence in a biological sample from said
25 individual the susceptibility allele of one or more SNPs of the identified above as
rs11561, rs5758405 rs8779, rs132806, rs2068943, rs2267487, rs881542, rs926333,
rs1060387, rs1006407, rs4468, rs138855, rs2239848, rs138880, rs138881, rs20551,
rs2294976, rs2076578, rs1046088, rs133068, rs133069, rs133070, rs133073,
rs6002408 and the presence at least one of the SNPs of the identified above as
30 rs909660, rs710193, rs1573745, rs132234, rs3752466, rs6010260, rs137931,
rs137932, rs3810971, rs2272843, rs1063900, rs715519, rs916005.

In yet another embodiment, the invention relates to a method for determining a
predisposition of an individual to a mental disease, in particular SCH and/BDP, said
35 method comprising determining the presence in a biological sample from said
individual one or more SNPs selected from the SNPs identified above rs909660,

rs710193, rs1573745, rs132234, rs3752466, rs6010260, rs137931, rs137932, rs3810971, rs2272843, rs1063900, rs715519, rs916005.

5 In yet another embodiment, the invention relates to a method for determining a predisposition of an individual to not having a mental disease, in particular SCH and/BDP, said method comprising determining the presence in a biological sample from said individual the protective allele of one or more SNPs selected from the SNPs identified above as rs11561, rs5758405 rs8779, rs132806, rs2068943, rs2267487, rs881542, rs926333, rs1060387, rs1006407, rs4468, rs138855, 10 rs2239848, rs138880, rs138881, rs20551, rs2294976, rs2076578, rs1046088, rs133068, rs133069, rs133070, rs133073, rs6002408.

15 In some embodiments the method for determining a predisposition/no predisposition to a mental disease of the invention, in particular SCH and/or BDP, may concern determining two or more of the above described SNPs

Screening for new candidate drugs for therapeutic treatment SCH and/or BPD

20 Screening methods for compounds with are capable of modulating the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 protein-protein interactions are within the scope of the invention.

25 For the purpose of below discussion molecules that produced in the cells due to activity of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 genes, such as transcriptional and translational products of the genes, are termed herein "gene products", if not specified otherwise.

30 Any method suitable for detecting protein-protein interactions may be employed for identifying the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 protein-protein interactions.

Among the traditional methods that may be employed are co-immunoprecipitation, cross-linking and co-purification through gradients or chromatographic columns. Utilizing procedures such as these allows for the identification of proteins, including intracellular proteins, which interact with BRD1, NHP2L1, PACSIN2, SERHL, 35 PIPPIN, EP300, FAM19A5 and/or GPR24 proteins. Once isolated, such a protein can be identified and can be used in conjunction with standard techniques, to identify

proteins it interacts with. For example, at least a portion of the amino acid sequence of a protein that interacts with BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 protein can be ascertained using techniques well known to those of skill in the art, such as via the Edman degradation technique (see, e.g.,
5 Creighton, 1983, "Proteins: Structures and Molecular Principles," W.H. Freeman & Co., N.Y., pp.34-49). The amino acid sequence obtained may be used as a guide for the generation of oligonucleotide mixtures that can be used to screen for gene sequences encoding such proteins. Screening made be accomplished, for example, by standard hybridization or PCR techniques. Techniques for the generation of
10 oligonucleotide mixtures and the screening are well-known. (See, e.g., Ausubel, supra, and 1990, "PCR Protocols: A Guide to Methods and Applications," Innis, et al., eds. Academic Press, Inc., New York).

Additionally, methods may be employed that result in the simultaneous identification
15 of genes that encode a protein which interacts with BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 protein. These methods include, for example, probing expression libraries with labelled BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 polypeptides, using BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 proteins in a
20 manner similar to the well known technique of antibody probing of lambda.gt11 and lambda.gt10 libraries.

One method that detects protein interactions in vivo, the two-hybrid system, is described in detail for illustration only and not by way of limitation. One version of this
25 system has been described (Chien, et al., 1991, Proc. Natl. Acad. Sci. USA, 88, 9578-9582) and is commercially available from Clontech (Palo Alto, Calif.).

Briefly, utilizing such a system, plasmids are constructed that encode two hybrid proteins: one consists of the DNA-binding domain of a transcription activator protein
30 fused to the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene peptide product and the other consists of the transcription activator protein's activation domain fused to an unknown protein that is encoded by a cDNA that has been recombined into this plasmid as part of a cDNA library. The DNA-binding domain fusion plasmid and the cDNA library are transformed into a strain of
35 the yeast *Saccharomyces cerevisiae* that contains a reporter gene (e.g., HBS or lacZ) whose regulatory region contains the transcription activator's binding site.

Either hybrid protein alone cannot activate transcription of the reporter gene: the DNA-binding domain hybrid cannot because it does not provide activation function and the activation domain hybrid cannot because it cannot localize to the activator's binding sites. Interaction of the two hybrid proteins reconstitutes the functional
5 activator protein and results in expression of the reporter gene, which is detected by an assay for the reporter gene product.

The two-hybrid system or related methodology may be used to screen activation domain libraries for proteins that interact with the "bait" gene product. By way of
10 example, and not by way of limitation, BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene derived peptide products may be used as the bait gene product. Total genomic or cDNA sequences are fused to the DNA encoding an activation domain. This library and a plasmid encoding a hybrid of a bait BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 protein, or a
15 fragment thereof, fused to the DNA-binding domain are co-transformed into a yeast reporter strain, and the resulting transformants are screened for those that express the reporter gene. For example, and not by way of limitation, a bait BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene sequence, such as the open reading frame of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN,
20 EP300, FAM19A5 and/or GPR24 gene, can be cloned into a vector such that it is translationally fused to the DNA encoding the DNA-binding domain of the GAL4 protein. These colonies are purified and the library plasmids responsible for reporter gene expression are isolated. DNA sequencing is then used to identify the proteins encoded by the library plasmids.

25 A cDNA library of the cell line from which proteins that interact with bait BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/ GPR24 gene product are to be detected can be made using methods routinely practiced in the art. According to the particular system described herein, for example, the cDNA
30 fragments can be inserted into a vector such that they are translationally fused to the transcriptional activation domain of GAL4. This library can be co-transformed along with the bait BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene sequence-GAL4 fusion plasmid into a yeast strain that contains a lacZ gene driven by a promoter that contains GAL4 activation sequence. A cDNA
35 encoded protein, fused to GAL4 transcriptional activation domain, that interacts with bait BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24

gene product will reconstitute an active GAL4 protein and thereby drive expression of the HIS3 gene. Colonies that express HIS3 can be detected by their growth on petri dishes containing semi-solid agar based media lacking histidine. The cDNA can then be purified from these strains, and used to produce and isolate the bait BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 protein-interacting protein using techniques routinely practiced in the art.

The invention also related to screening assays for compounds that interfere with the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene products macromolecule interaction.

The BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene products of the invention may, in vivo, interact with one or more macromolecules, including intracellular macromolecules, such as proteins. Such macromolecules may include, but are not limited to, nucleic acid molecules and those proteins identified via methods such as those described above. For purposes of this discussion, the macromolecules are referred to herein as "binding partners". Compounds that are able to disrupt the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 gene products binding in this way may be useful in regulating the activity of products of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 genes, especially variant BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 proteins and thereof derived peptide products. Such compounds may include, but are not limited to molecules such as peptides, and the like, which would be capable of gaining access to a BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene product.

The basic principle of the assay systems used to identify compounds that interfere with the interaction between BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 gene products and their binding partner or partners involves preparing a reaction mixture containing the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene product, and the binding partner under conditions and for a time sufficient to allow the two to interact and bind, thus forming a complex. In order to test a compound for inhibitory activity, the reaction mixture is prepared in the presence and absence of the test compound. The test compound may be initially included in the reaction mixture, or may be added at a

time subsequent to the addition of BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene product and its binding partner. Control reaction mixtures are incubated without the test compound or with a placebo. The formation of any complexes between the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene product and the binding partner is then detected. The formation of a complex in the control reaction, but not in the reaction mixture containing the test compound, indicates that the compound interferes with the interaction of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene product and the interactive binding partner. Additionally, complex formation within reaction mixtures containing the test compound and for example normal (wild type) BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 protein may also be compared to complex formation within reaction mixtures containing the test compound and a variant BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 protein. This comparison may be important in those cases wherein it is desirable to identify compounds that disrupt interactions of mutant but not wild type BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 protein.

The assay for compounds that interfere with the interaction of BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/ GPR24 gene products and their binding partners can be conducted in a heterogeneous or homogeneous format. Heterogeneous assays involve anchoring either the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene product or the binding partner onto a solid phase and detecting complexes anchored on the solid phase at the end of the reaction. In homogeneous assays, the entire reaction is carried out in a liquid phase. In either approach, the order of addition of reactants can be varied to obtain different information about the compounds being tested. For example, test compounds that interfere with the interaction between the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene products and the binding partners, e.g., by competition, can be identified by conducting the reaction in the presence of the test substance; i.e., by adding the test substance to the reaction mixture prior to or simultaneously with the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene protein and interactive intracellular binding partner. Alternatively, test compounds that disrupt preformed complexes, e.g., compounds with higher binding constants that displace one of the components

from the complex, can be tested by adding the test compound to the reaction mixture after complexes have been formed. The various formats are described briefly below.

In a heterogeneous assay system, either the BRD1, NHP2L1, PACSIN2, SERHL,
5 PIPPIN, EP300, FAM19A5 and/or GPR24 gene product or the interactive binding partner, is anchored onto a solid surface, while the non-anchored species is labeled, either directly or indirectly. In practice, microtiter plates are conveniently utilized. The anchored species may be immobilized by non-covalent or covalent attachments. Non-covalent attachment may be accomplished simply by coating the solid surface
10 with a solution of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene product or binding partner and drying. Alternatively, an immobilized antibody specific for the species to be anchored may be used to anchor the species to the solid surface. The surfaces may be prepared in advance and stored.

15 In order to conduct the assay, the partner of the immobilized species is exposed to the coated surface with or without the test compound. After the reaction is complete, unreacted components are removed (e.g., by washing) and any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored
20 on the solid surface can be accomplished in a number of ways. Where the non-immobilized species is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the non-immobilized species is not pre-labelled, an indirect label can be used to detect complexes anchored on the surface; e.g., using a labeled antibody specific for the initially non-immobilized
25 species (the antibody, in turn, may be directly labeled or indirectly labeled with a labeled anti-Ig antibody). Depending upon the order of addition of reaction components, test compounds that inhibit complex formation or that disrupt preformed complexes can be detected.

30 Alternatively, the reaction can be conducted in a liquid phase in the presence or absence of the test compound, the reaction products separated from unreacted components, and complexes detected; e.g., using an immobilized antibody specific for one of the binding components to anchor any complexes formed in solution, and a labeled antibody specific for the other partner to detect anchored complexes. Again,
35 depending upon the order of addition of reactants to the liquid phase, test

compounds that inhibit complex or that disrupt preformed complexes can be identified.

In an alternate embodiment of the invention, a homogeneous assay can be used. In this approach, a preformed complex of a BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene product and the interactive binding partner is prepared in which either the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene product or its binding partners is labeled, but the signal generated by the label is quenched due to complex formation (see, e.g., U.S. Pat. No.4,109,496 by Rubenstein which utilizes this approach for immunoassays). The addition of a test substance that competes with and displaces one of the species from the preformed complex will result in the generation of a signal above background. In this way, test substances that disrupt the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 gene product/binding partner interaction can be identified.

In another embodiment, the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene product can be prepared for immobilization using recombinant DNA techniques. For example, the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene coding region can be fused to the glutathione-S-transferase (GST) gene using a fusion vector, such as pGEX-5X-1, in such a manner that its binding activity is maintained in the resulting fusion protein. The interactive binding partner can be purified and used to raise an antibody, using methods routinely practiced in the art. The antibody can then be labeled with a radioactive isotope such as ¹²⁵I, for example, by methods routinely practiced in the art. In a heterogeneous assay, e.g., the GST- BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 fusion protein can be anchored to glutathione-agarose beads. The interactive binding partner can then be added in the presence or absence of the test compound in a manner that allows interaction and binding to occur. At the end of the reaction period, unbound material can be washed away, and the labeled monoclonal antibody can be added to the system and allowed to bind to the complexed components. The interaction between the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene product and the interactive binding partner can be detected by measuring the amount of radioactivity that remains associated with the glutathione-agarose beads. A successful inhibition

of the interaction by the test compound will result in a decrease in measured radioactivity.

Alternatively, the GST-BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300,
5 FAM19A5 and/or GPR24 fusion protein and the interactive binding partner can be mixed together in liquid in the absence of the solid glutathione-agarose beads. The test compound can be added either during or after the species are allowed to interact. This mixture can then be added to the glutathione-agarose beads and unbound material is washed away. Again the extent of inhibition of the BRD1,
10 NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene product/binding partner interaction can be detected by adding the labelled antibody and measuring the radioactivity associated with the beads.

In still another embodiment of the invention, these same techniques can be
15 employed using peptide fragments that correspond to the binding domains of BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 proteins and/or the interactive or binding partner (in cases where the binding partner is a protein), in place of one or both of the full length proteins. Any number of methods routinely practiced in the art can be used to identify and isolate the binding sites.
20 These methods include, but are not limited to, mutagenesis of the gene encoding one of the proteins and screening for disruption of binding in a co-immunoprecipitation assay. Compensating mutations in the gene encoding the second species in the complex can then be selected. Sequence analysis of the genes encoding the respective proteins will reveal the mutations that correspond to the region of the
25 protein involved in interactive binding. Alternatively, one protein can be anchored to a solid surface using methods described in this Section above, and allowed to interact with and bind to its labeled binding partner, which has been treated with a proteolytic enzyme, such as trypsin. After washing, a short, labelled peptide comprising the binding domain may remain associated with the solid material, which can be isolated
30 and identified by amino acid sequencing. Also, once the gene coding for the segments can be engineered to express peptide fragments of the protein, which can then be tested for binding activity and purified or synthesized.

For example, and not by way of limitation, a BRD1, NHP2L1, PACSIN2, SERHL,
35 PIPPIN, EP300, FAM19A5 and/or GPR24 gene product can be anchored to a solid material as described above by making a GST- BRD1, NHP2L1, PACSIN2, SERHL,

PIPPIN, EP300, FAM19A5 and/or GPR24 fusion protein and allowing it to bind to glutathione agarose beads. The interactive binding partner obtained can be labeled with a radioactive isotope, such as ³⁵S, and cleaved with a proteolytic enzyme such as trypsin. Cleavage products can then be added to the anchored GST-BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 fusion protein and allowed to bind. After washing away unbound peptides, labelled bound material, representing the binding partner binding domain, can be eluted, purified, and analyzed for amino acid sequence by well-known methods. Peptides so identified can be produced synthetically or fused to appropriate facilitative proteins using recombinant DNA technology.

The invention also provides assays for identification of compounds that ameliorate the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 gene associated disorders, such as SCH and/or BPD.

Compounds, including but not limited to binding compounds identified via assay techniques such as those described above can be tested for the ability to ameliorate symptoms of a BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene associated disorder including a disorder of thought and/or mood, neuropsychiatric disorders including bipolar (BPD), genetically related unipolar affective disorders, delusional disorders, paraphrenia, paranoid psychosis, SCH, schizotypal disorder, schizoaffective disorder, schizoaffective bipolar and genetically related unipolar affective disorders, psychogenic psychosis, catatonia, periodic bipolar and genetically related unipolar affective disorders, cycloid psychosis, schizoid personality disorder, paranoid personality disorder, bipolar and genetically related unipolar affective disorders related affective disorders and subtypes of unipolar affective disorder.

It should be noted that the assays described herein can identify compounds that affect the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene activity by either affecting BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 gene expression or by affecting the level of BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 gene product activity. For example, compounds may be identified that are involved in another step in the pathway in which the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene and/or the BRD1, NHP2L1, PACSIN2, SERHL,

PIPPIN, EP300, FAM19A5 and/or GPR24 gene product is involved and, by affecting this same pathway may modulate the effect of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene on the development of a neuropsychiatric disorder such as SCH and/or BPD. Such compounds can be used
5 as part of a therapeutic method for the treatment of the disorder.

Described below are cell-based and animal model-based assays for the identification of compounds exhibiting such an ability to ameliorate symptoms of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 gene activity
10 associated with a neuropsychiatric disorder, such as SCH and/or BPD.

First, cell-based systems can be used to identify compounds that may act to ameliorate symptoms of a BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene associated disorder, such as BPD and/or SCH. Such
15 cell systems can include, for example, recombinant or non-recombinant cell, such as cell lines, that express the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene.

In utilizing such cell systems, cells that express the BRD1, NHP2L1, PACSIN2,
20 SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene may be exposed to a compound suspected of exhibiting an ability to ameliorate symptoms of a BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene disorder, such as SCH and/or BPD, at a sufficient concentration and for a sufficient time to elicit such an amelioration of such symptoms in the exposed cells. After
25 exposure, the cells can be assayed to measure alterations in the expression of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene, e.g., by assaying cell lysates for the presence of BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene transcripts (e.g., by Northern analysis) or for the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5
30 and GPR24 gene translation products expressed by the cell. Compounds that modulate expression of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene are considered to be good candidates as therapeutics.

35 Alternatively, the cells are examined to determine whether one or more cellular phenotypes associated with a BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300,

FAM19A5 and/or GPR24 gene disorder, such as SCH and/or BPD, has been altered to resemble a more normal or unimpaired, unaffected phenotype, or a phenotype more likely to produce a lower incidence or severity of disorder symptoms.

- 5 In addition, animal-based systems or models for a BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene associated disorder, such as SCH and/or BPD, which may include, for example mice, may be used to identify compounds capable of ameliorating symptoms of the disorder. Such animal models may be used as test substrates for the identification of drugs, pharmaceuticals, therapies and interventions that may be effective in treating such disorders. For example, animal models may be exposed to a compound suspected of exhibiting an ability to ameliorate symptoms, at a sufficient concentration and for a sufficient time to elicit such an amelioration of symptoms of a BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene associated disorder, such as SCH and/or BPD, in the exposed animals. The response of the animals to the exposure may be monitored by assessing the reversal of such symptoms.

- With regard to intervention, any treatments that reverse any aspect of symptoms of a BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene associated disorder, such as SCH and/or BPD, should be considered as candidates for human therapeutic intervention in such a disorder.

Therapeutic treatment

- The present invention relates to a BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene associated disorder including a disorder of thought and/or mood, neuropsychiatric disorders including bipolar (BPD), genetically related unipolar affective disorders, delusional disorders, paraphrenia, paranoid psychosis, SCH, schizotypal disorder, schizoaffective disorder, schizoaffective bipolar and genetically related unipolar affective disorders, psychogenic psychosis, catatonia, periodic bipolar and genetically related unipolar affective disorders, cycloid psychosis, schizoid personality disorder, paranoid personality disorder, bipolar and genetically related unipolar affective disorders related affective disorders and subtypes of unipolar affective disorder.

Having identified a group of subjects having a polymorphism as described in the present invention, the invention also relates to the use of compounds directed to decreasing or modulating the effect of the polymorphism for the preparation of a medicament for the treatment of SCH and/or BPD in said subjects.

5

The compounds that bind to a BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene product, intracellular proteins or portions of proteins that interact with a BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene product, compounds that interfere with the interaction of a
10 BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene product with intracellular proteins and compounds that modulate the activity of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 genes (i.e., modulate the level of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 gene expression and/or modulate the level of the BRD1,
15 NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 gene product activity) are considered to be good candidates for the manufacture of a medicament for treatment of a BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene associated disorder. Assays may additionally be utilized that identify compounds that bind to the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN,
20 EP300, FAM19A5 and/or GPR24 gene regulatory sequences (e.g., promoter sequences; see e.g., Platt, 1994, J. Biol. Chem. 269, 28558-28562), and that may modulate the level of BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene expression. Compounds may include, but are not limited to, small organic molecules, such as ones that are able to cross the blood-brain barrier,
25 gain entry into an appropriate cell and affect expression of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene or some other gene involved in a BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene dependent regulatory pathway, or intracellular proteins. Such intracellular proteins may for example be involved in the control and/or regulation of
30 mood. Further, among these compounds are compounds that affect the level of BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene expression and/or the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene product activity and that can be used as medicaments in the therapeutic treatment of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN,
35 EP300, FAM19A5 and/or GPR24 gene associated disorders, for example neuropsychiatric disorders such as SCH and/or BPD.

Compounds may include, but are not limited to, peptides such as, for example, soluble peptides, including but not limited to, Ig-tailed fusion peptides, and members of random peptide libraries; (see, e.g., Lam, et al., 1991, Nature 354, 82-84; 5 Houghten, et al., 1991, Nature 354, 84-86), and combinatorial chemistry-derived molecular library made of D- and/or L- configuration amino acids, phosphopeptides (including, but not limited to members of random or partially degenerate, directed phosphopeptide libraries; see, e.g., Songyang, et al., 1993, Cell 72, 767-778), antibodies (including, but not limited to, polyclonal, monoclonal, humanized, anti- 10 idiotypic, chimeric or single chain antibodies, and FAb, F(ab').sub.2 and Fab expression library fragments, and epitope-binding fragments thereof), and small organic or inorganic molecules. Such compounds may further comprise compounds, in particular drugs or members of classes or families of drugs, known to ameliorate or exacerbate the symptoms of a neuropsychiatric disorder such as bipolar and 15 genetically related unipolar affective disorders with the use of lithium salts, atypical antipsychotics such as ziprasadone, risperidone, clozapine, quetiapine, olanzapine, butyrophenone derivatives such as haloperidol and droperidol, phenothiazine derivatives such as chlorpromazine, prochlorperazine, promazine, trifluopromazine, thioxanthine derivatives such as flupenthixol, chlorprothixene and dibenzodiazepines 20 and antipsychotic antiepileptic drugs such as carbamazepine, and valproic acid. Antidepressant drugs such as imipramine, amitriptyline, nortriptyline, prothiaden, doxapine, other tricyclic antidepressants, tetracyclic antidepressants, serotonin reuptake inhibitor antidepressants such as fluoxetine, paroxetine, ciproamil, venlafaxine, monoamine oxidase inhibitor antidepressants such as phenelzine, 25 tranlycypromine, isocaboxazid, selegiline, and moclobamide. In addition psychotogenic drugs such as bromocriptine, apomorphine, amphetamine, methylphenidate, methylamphetamine, ketamine. Many of these drugs can be or have been used in combination.

30 *Compounds identified via assays such as those described herein may be useful, for example, in elaborating the biological function of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 gene products, and for ameliorating the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 gene associated disorders or neuropsychiatric disorders, such as for example SCH and/or BPD.*

35

In one embodiment of treatment methods, the compounds administered do not comprise compounds, in particular drugs, reported to ameliorate or exacerbate the

symptoms of a neuropsychiatric disorder, such as bipolar and genetically related unipolar affective disorders. Such compounds include antidepressants such as lithium salts, flupenthixol, risperidone, clozapine, quetiapine, olanzapine, haloperidol, droperidol, chlorpromazine, prochlorperazine, phenothiazine derivatives, promazine, trifluopromazine, butyrophenone derivatives, thioxanthine derivatives such as chlorprothixene and dibenzodiazepines and antipsychotic antiepileptic drugs such as carbamazepine, and valproic acid, reserpine. Psychotogenic drugs such as LSD, bromocriptine, apomorphine, amphetamine, methylphenidate, methylamphetaime, ketamine, Many of these drugs are used in combination.

10

1. Inhibitory antisense, ribozyme and triple helix approaches

In another embodiment, symptoms of certain neuropsychiatric disorders, such as SCH and/or BPD, may be ameliorated by decreasing the level of BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene expression and/or the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene product activity by using the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene derived nucleotide sequences in conjunction with well-known antisense, gene "knock-out," ribozyme and/or triple helix methods to decrease the level of BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene expression. Among the compounds that may exhibit the ability to modulate the activity, expression of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene and/or synthesis the gene products, including the ability to ameliorate the symptoms of a BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene disorder, are antisense, ribozyme, and triple helix molecules. Such molecules may be designed to reduce or inhibit either unimpaired, or if appropriate, mutant target gene activity. Techniques for the production and use of such molecules are well known to those of skill in the art.

Antisense RNA and DNA molecules act to directly block the translation of mRNA by hybridizing to targetted mRNA and preventing protein translation. Antisense approaches involve the design of oligonucleotides that are complementary to a target gene mRNA. The antisense oligonucleotides will bind to the complementary target gene mRNA transcripts and prevent translation. Absolute complementarity, although preferred, is not required.

A sequence "complementary" to a portion of an RNA, as referred to herein, means a sequence having sufficient complementarity to be able to hybridize with the RNA, forming a stable duplex; in the case of double-stranded antisense nucleic acids, a single strand of the duplex DNA may thus be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of complementarity and the length of the antisense nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with an RNA it may contain and still form a stable duplex (or triplex, as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

In one embodiment, oligonucleotides complementary to non-coding regions of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene could be used in an antisense approach to inhibit translation of endogenous BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 mRNA. Antisense nucleic acids should be at least six nucleotides in length, and are preferably oligonucleotides ranging from 6 to about 50 nucleotides in length. In specific aspects the oligonucleotide is at least 10 nucleotides, at least 17 nucleotides, at least 25 nucleotides or at least 50 nucleotides.

Regardless of the choice of target sequence, it is preferred that in vitro studies are first performed to quantitate the ability of the antisense oligonucleotide to inhibit gene expression. It is preferred that these studies utilize controls that distinguish between antisense gene inhibition and nonspecific biological effects of oligonucleotides. It is also preferred that these studies compare levels of the target RNA or protein with that of an internal control RNA or protein. Additionally, it is envisioned that results obtained using the antisense oligonucleotide are compared with those obtained using a control oligonucleotide. It is preferred that the control oligonucleotide is of approximately the same length as the test oligonucleotide and that the nucleotide sequence of the oligonucleotide differs from the antisense sequence no more than is necessary to prevent specific hybridization to the target sequence.

The oligonucleotides can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization, etc. The oligonucleotide

may include other appended groups such as peptides (e.g., for targeting host cell receptors in vivo), or agents facilitating transport across the cell membrane (see, e.g., Letsinger, et al., 1989, Proc. Natl. Acad. Sci. U.S.A. 86,6553-6556; Lemaitre, et al., 1987, Proc. Natl. Acad. Sci. 84, 648-652; PCT Publication No. WO88/09810, published Dec. 15, 1988) or the blood-brain barrier (see, e.g., PCT Publication No. WO89/10134, published Apr. 25, 1988), hybridization-triggered cleavage agents (see, e.g., Krol et al., 1988, BioTechniques 6, 958-976) or intercalating agents (see, e.g., Zon, 1988, Pharm. Res. 5, 539-549). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

The antisense oligonucleotide may comprise at least one modified base moiety which is selected from the group including but not limited to 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

The antisense oligonucleotide may also comprise at least one modified sugar moiety selected from the group including but not limited to arabinose, 2-fluoroarabinose, xylulose, and hexose.

30

In yet another embodiment, the antisense oligonucleotide comprises at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

35

In yet another embodiment, the antisense oligonucleotide is an .alpha.-anomeric oligonucleotide. An alpha.-anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual .beta.-units, the strands run parallel to each other (Gautier, et al., 1987, Nucl. Acids Res. 15, 6625-6641). The oligonucleotide is a 2'-O-methylribonucleotide (Inoue, et al., 1987, Nucl. Acids Res. 15, 6131-6148), or a chimeric RNA-DNA analogue (Inoue, et al., 1987, FEBS Lett. 215, 327-330).

Oligonucleotides of the invention may be synthesized by standard methods known in the art, e.g. by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides may be synthesized by the method of Stein, et al. (1988, Nucl. Acids Res. 16, 3209), methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin, et al., 1988, Proc. Natl. Acad. Sci. U.S.A. 85, 7448-7451), etc.

While antisense nucleotides complementary to the target gene coding region sequence could be used, those complementary to the transcribed, untranslated region are most preferred. For example, antisense oligonucleotides having the following sequences can be utilized in accordance with the invention:

Antisense molecules should be delivered to cells that express the target gene in vivo. A number of methods have been developed for delivering antisense DNA or RNA to cells; e.g., antisense molecules can be injected directly into the tissue site, or modified antisense molecules, designed to target the desired cells (e.g., antisense linked to peptides or antibodies that specifically bind receptors or antigens expressed on the target cell surface) can be administered systemically.

However, it is often difficult to achieve intracellular concentrations of the antisense sufficient to suppress translation of endogenous mRNAs. Therefore a preferred approach utilizes a recombinant DNA construct in which the antisense oligonucleotide is placed under the control of a strong pol III or pol II promoter. The use of such a construct to transfect target cells in the patient will result in the transcription of sufficient amounts of single stranded RNAs that will form complementary base pairs with the endogenous target gene transcripts and thereby prevent translation of the target gene mRNA. For example, a vector can be

- introduced e.g., such that it is taken up by a cell and directs the transcription of an antisense RNA. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA. Such vectors can be constructed by recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in mammalian cells. Expression of the sequence encoding the antisense RNA can be by any promoter known in the art to act in mammalian, preferably human cells. Such promoters can be inducible or constitutive. Such promoters include but are not limited to: the SV40 early promoter region (Bernoist and Chambon, 1981, Nature 290, 304-310), the promoter contained in the 31 long terminal repeat of Rous sarcoma virus (Yamamoto, et al., 1980, Cell 22, 787-797), the herpes thymidine kinase promoter (Wagner, et al., 1981, Proc. Natl. Acad. Sci. U.S.A. 78, 1441-1445), the regulatory sequences of the metallothionein gene (Brinster, et al., 1982, Nature 296, 39-42), etc. Any type of plasmid, cosmid, YAC or viral vector can be used to prepare the recombinant DNA construct which can be introduced directly into the tissue site. Alternatively, viral vectors can be used that selectively infect the desired tissue, in which case administration may be accomplished by another route (e.g., systemically).
- 20 Ribozyme molecules designed to catalytically cleave target gene mRNA transcripts can also be used to prevent translation of target gene mRNA and, therefore, expression of target gene product. (See, e.g., PCT International Publication WO90/11364, published Oct. 4, 1990; Sarver, et al., 1990, Science 247, 1222-1225).
- 25 Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. (For a review, see Rossi, 1994, Current Biology 4, 469-471). The mechanism of ribozyme action involves sequence specific hybridization of the ribozyme molecule to complementary target RNA, followed by an endonucleolytic cleavage event. The composition of ribozyme molecules must include one or more sequences complementary to the target gene mRNA, and must include the well known catalytic sequence responsible for mRNA cleavage. For this sequence, see, e.g., U.S. Pat. No. 5,093,246, which is incorporated herein by reference in its entirety.
- 30 While ribozymes that cleave mRNA at site specific recognition sequences can be used to destroy target gene mRNAs, the use of hammerhead ribozymes is preferred.
- 35

Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well known in the art and is described more fully in Myers, 1995, Molecular Biology and Biotechnology: A Comprehensive Desk Reference, VCH Publishers, New York, (see especially Figure. 4, page 833) and in Haseloff and Gerlach, 1988, Nature, 334, 585-591, which is incorporated herein by reference in its entirety.

- 10 Preferably the ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the target gene mRNA, i.e., to increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts. For example, hammerhead ribozymes having the following sequences can be utilized. The ribozymes of the present invention also include RNA endoribonucleases (hereinafter
- 15 "Cech-type ribozymes") such as the one that occurs naturally in Tetrahymena thermophila (known as the IVS, or L-19 IVS RNA) and that has been extensively described by Thomas Cech and collaborators (Zaug, et al., 1984, Science, 224, 574-578; Zaug and Cech, 1986, Science, 231, 470-475; Zaug, et al., 1986, Nature, 324, 429-433; published International patent application No. WO 88/04300 by University
- 20 Patents Inc.; Been and Cech, 1986, Cell, 47, 207-216). The Cech-type ribozymes have an eight base pair active site which hybridizes to a target RNA sequence where after cleavage of the target RNA takes place.

- As in the antisense approach, the ribozymes can be composed of modified
- 25 oligonucleotides (e.g., for improved stability, targeting, etc.) and should be delivered to cells that express the target gene in vivo. A preferred method of delivery involves using a DNA construct "encoding" the ribozyme under the control of a strong constitutive pol III or pol II promoter, so that transfected cells will produce sufficient quantities of the ribozyme to destroy endogenous target gene messages and inhibit
- 30 translation. Because ribozymes unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficiency.

- Endogenous target gene expression can also be reduced by inactivating or "knocking out" the target gene or its promoter using targeted homologous recombination (e.g.,
- 35 see Smithies, et al., 1985, Nature 317, 230-234; Thomas and Capecchi, 1987, Cell 51, 503-512; Thompson, et al., 1989, Cell 5, 313-321; each of which is incorporated

by reference herein in its entirety). For example, a mutant, non-functional target gene (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous target gene (either the coding regions or regulatory regions of the target gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express the target gene in vivo. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the target gene. Such approaches are particularly suited in the agricultural field where modifications to ES (embryonic stem) cells can be used to generate animal offspring with an inactive target gene (e.g., see Thomas and Capecchi, 1987 and Thompson, 1989, supra). However this approach can be adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site in vivo using appropriate viral vectors.

Alternatively, endogenous target gene expression can be reduced by targeting deoxyribonucleotide sequences complementary to the regulatory region of the target gene (i.e., the target gene promoter and/or enhancers) to form triple helical structures that prevent transcription of the target gene in target cells in the body. (See generally, Helene, 1991, *Anticancer Drug Des.*, 6(6), 569-584; Helene, et al., 1992, *Ann. N.Y. Acad. Sci.*, 660, 27-36; and Maher, 1992, *Bioassays* 14(12), 807-815).

Nucleic acid molecules to be used in triplex helix formation for the inhibition of transcription should be single stranded and composed of deoxynucleotides. The base composition of these oligonucleotides must be designed to promote triple helix formation via Hoogsteen base pairing rules, which generally require sizeable stretches of either purines or pyrimidines to be present on one strand of a duplex. Nucleotide sequences may be pyrimidine-based, which will result in TAT and CGC.sup.+ triplets across the three associated strands of the resulting triple helix. The pyrimidine-rich molecules provide base complementarity to a purine-rich region of a single strand of the duplex in a parallel orientation to that strand. In addition, nucleic acid molecules may be chosen that are purine-rich, for example, contain a stretch of G residues. These molecules will form a triple helix with a DNA duplex that is rich in GC pairs, in which the majority of the purine residues are located on a single strand of the targeted duplex, resulting in GGC triplets across the three strands in the triplex.

Alternatively, the potential sequences that can be targeted for triple helix formation may be increased by creating a so called "switchback" nucleic acid molecule. Switchback molecules are synthesized in an alternating 5'-3', 3'-5' manner, such that they base pair with first one strand of a duplex and then the other, eliminating the
5 necessity for a sizeable stretch of either purines or pyrimidines to be present on one strand of a duplex.

In instances wherein the antisense, ribozyme, and/or triple helix molecules described herein are utilized to inhibit mutant gene expression, it is possible that the technique
10 may so efficiently reduce or inhibit the transcription (triple helix) and/or translation (antisense, ribozyme) of mRNA produced by normal target gene alleles that the possibility may arise wherein the concentration of normal target gene product present may be lower than is necessary for a normal phenotype. In such cases, to ensure that substantially normal levels of target gene activity are maintained, therefore,
15 nucleic acid molecules that encode and express target gene polypeptides exhibiting normal target gene activity may, be introduced into cells via gene therapy methods such as those described, below, in Section 5.9.2 that do not contain sequences susceptible to whatever antisense, ribozyme, or triple helix treatments are being utilized. Alternatively, in instances whereby the target gene encodes an extracellular
20 protein, it may be preferable to co-administer normal target gene protein in order to maintain the requisite level of target gene activity.

Anti-sense RNA and DNA, ribozyme, and triple helix molecules of the invention may be prepared by any method known in the art for the synthesis of DNA and RNA
25 molecules, as discussed above. These include techniques for chemically synthesizing oligodeoxyribonucleotides and oligoribonucleotides well known in the art such as for example solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be
30 incorporated into a wide variety of vectors that incorporate suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

2 Gene therapy

Having identified polymorphism(s) as the cause of a disease it is also rendered possible with the present invention to provide a genetic therapy for subjects being diagnosed as having a predisposition according to the invention, said therapy
5 comprising administering to said subject a therapeutically effective amount of a gene therapy vector.

Having discovered the NHP2L1, PACSIN2, SERHL, PIPPIN, BRD1, EP300, FAM19A5 and/or GPR24 genes as etiological factors in mental diseases, such as
10 SCH and/or BPD, the inventors also provide methods for gene therapy and gene therapy vectors for use in subjects irrespective of whether they carry any of the susceptibility or protective alleles/haplotypes described in the present invention.

There are various different methods of gene therapy for the subjects defined in the
15 present invention.

The first two are based on activation of the repair system of the cells by introducing into those cells a gene therapy vector which causes "correction" of the polymorphism by presenting the repair mechanism with a template for carrying out the correction.
20 One such type includes the RNA/DNA chimeraplast, said chimeraplast being capable of correcting the polymorphism in cells of said subject. Examples of the design of such chimeraplasts can be found in e.g. US 5,760,012; US 5,888,983; US 5,731,181; US 6,010,970; US 6,211,351.

25 The second method is based on application of single stranded oligonucleotides, wherein the terminal nucleotides is protected from degradation by using 3' and 5' phosphorothioate-linkage of the monomers. This gene therapy vector is also capable of "correcting" the polymorphism by replacing one nucleotide with another.

30 These first two types of gene therapy vectors comprise a small sequence (less than 50 bases) which overlaps with the polymorphism in question. Suitable sequences for this purpose are genomic sequences located around the polymorphism.

Other types of gene therapy include the use of retrovirus (RNA-virus). Retrovirus can
35 be used to target many cells and integrate stably into the genome. Adenovirus and

adeno-associated virus can also be used. A suitable retrovirus or adenovirus for this purpose comprises an expression construct with the wildtype gene under the control of the wildtype promoter or a constitutive promoter or a regulatable promoter such as a repressible and/or inducible promoter or a promoter comprising both repressible and inducible elements.

A further group of gene therapy vectors includes vectors comprising interfering RNA (RNAi) for catalytic breakdown of mRNA carrying the polymorphism. RNAi can be used for lowering the expression of a given gene for a relatively short period of time. In particular these RNAi oligos may be used for therapy for both subjects carrying a susceptibility allele as described in the present invention as well as for subjects which do not carry such an allele.

Interfering RNA ("RNAi") is double stranded RNA that results in catalytic degradation of specific mRNAs, and can also be used to lower gene expression.

The gene therapy vectors carry the protective allele of the genes. The protective allele means in the present content that presence of this allele in an individual indicates protection against a mental disease of the invention.

Described below are methods and compositions whereby a BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene disorder or a disorder of thought and/or mood, such as bipolar and genetically related unipolar affective disorders, may be treated.

With respect to an increase in the level of normal BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 AND GPR24 gene expression and/or BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 GENE product activity, the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 gene derived nucleotide sequences, for example, be utilized for the treatment of a BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene associated disorder such as SCH and/or BPD. Such treatment can be performed, for example, in the form of gene replacement therapy. Specifically, one or more copies of a normal BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene or a portion of said gene that directs the production of

a gene product exhibiting normal BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene function, may be inserted into the appropriate cells within a patient, using vectors that include, but are not limited to adenovirus, adeno-associated virus, and retrovirus vectors, in addition to other particles that
5 introduce DNA into cells, such as liposomes.

Because the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 genes are expressed in the brain, such gene replacement therapy techniques should be capable delivering the BRD1, NHP2L1, PACSIN2, SERHL,
10 PIPPIN, EP300, FAM19A5 and/or GPR24 gene sequences to these cell types within patients. Thus, in one embodiment, techniques that are well known to those of skill in the art (see, e.g., PCT Publication No. WO89/10134, published Apr. 25, 1988) can be used to enable the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene sequences to cross the blood-brain barrier readily and to deliver
15 the sequences to cells in the brain. With respect to delivery that is capable of crossing the blood-brain barrier, viral vectors such as, for example, those described above, are preferable. Also included are methods using liposomes either in vivo ex vivo or in vitro. Wherein the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene sense or antisense DNA is delivered to the cytoplasm
20 and nucleus of target cells. Liposomes can deliver the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and GPR24 gene sense or nonsense RNA to humans and the human brain or in mammals through intrathecal delivery either as part of a viral vector or as DNA conjugated with nuclear localizing proteins or other proteins that increase take up into the cell nucleus.

25

In another embodiment, techniques for delivery involve direct administration of such BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene sequences to the site of the cells in which the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene sequences are to be expressed.
30 Additional methods that may be utilized to increase the overall level of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene expression and/or the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene product activity include the introduction of appropriate
35 BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene-expressing cells, preferably autologous cells, into a patient at positions and in numbers that are sufficient to ameliorate the symptoms of a BRD1, NHP2L1,

PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene associated disorder, such as SCH and/or BPD. Such cells may be either recombinant or non-recombinant.

- 5 Among the cells that can be administered to increase the overall level of BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene expression in a patient are normal cells, preferably brain cells and also choroid plexus cells within the CNS which are accessible through intrathecal injections. Alternatively, cells, preferably autologous cells, can be engineered to express BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene sequences, and may then be introduced into a patient in positions appropriate for the amelioration of the symptoms of a BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene associated disorder. Alternately, cells that express an unimpaired BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene and that are from a MHC matched individual can be utilized, and may include, for example, brain cells. The expression of the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene derived sequences is controlled by the appropriate gene regulatory sequences to allow such expression in the necessary cell types. Such gene regulatory sequences are well known to the skilled artisan. Such cell-based gene therapy techniques are well known to those skilled in the art, see, e.g., Anderson, U.S. Pat. No. 5,399,349.

When the cells to be administered are non-autologous cells, they can be administered using well known techniques that prevent a host immune response against the introduced cells from developing. For example, the cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

- 30 Additionally, compounds, such as those identified via techniques such as those described above that are capable of modulating the BRD1, NHP2L1, PACSIN2, SERHL, PIPPIN, EP300, FAM19A5 and/or GPR24 gene product activity can be administered using standard techniques that are well known to those of skill in the art. In instances in which the compounds to be administered are to involve an interaction with brain cells, the administration techniques should include well known

ones that allow for a crossing of the blood-brain barrier such as intrathecal injection and conjugation with compounds that allow transfer across the blood brain barrier.

Preferred embodiments of the invention concern the gene therapy vectors comprising

- 5 (i) a DNA sequence selected from the sequences identified as SEQ ID NO 1-7 and 94, or a fragment thereof, wherein said sequence or a said fragment comprises the protective allele of an SNP selected from the SNPs having refSNP IDs: rs11561, rs5758405 rs8779, rs132806, rs2068943, rs2267487, rs881542, rs926333, 10 rs1060387, rs1006407, rs6002408, rs4468, rs138855, rs2239848, rs138880, rs138881, rs20551, rs2294976, rs2076578, rs1046088, rs133068, rs133069, rs133070, rs133073, rs6002408, or
- 15 (ii) a DNA sequence selected from the sequences identified as SEQ ID NOs: 8-14 and 95, or a fragment of said DNA sequence, wherein said DNA sequence or said fragment comprises the protective allele of an SNP selected from the SNPs having refSNP IDs: rs11561, rs5758405 rs8779, rs132806, rs2068943, rs2267487, rs881542, rs926333, rs1060387, rs1006407, rs6002408, rs4468, rs138855, rs2239848, rs138880, rs138881, 20 rs20551, rs2294976, rs2076578, rs1046088, rs133068, rs133069, rs133070, rs133073, rs600240.

Pharmaceutical compositions and methods of administration

The compounds that are determined to affect BRD1, NHP2L1, PACSIN2, SERHL, 25 PIPPIN, EP300, FAM19A5 and/or GPR24 gene expression or gene product activity can be administered to a patient at therapeutically effective doses to treat or ameliorate a gene associated disorder, such as SCH and/or BPD. A therapeutically effective dose refers to that amount of the compound sufficient to result in amelioration of symptoms of such a disorder.

30

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio 35 LD50 /ED50. Compounds that exhibit large therapeutic indices are preferred. While

compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

5

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED₅₀ with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC_{sub}.50 (i.e., the concentration of the test compound that achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

20

Pharmaceutical compositions for use in accordance with the present invention may be formulated in conventional manner using one or more physiologically acceptable carriers or excipients.

25

Thus, the compounds and their physiologically acceptable salts and solvates may be formulated for administration by inhalation or insufflation (either through the mouth or the nose) or intrathecal, oral, buccal, parenteral or rectal administration.

30

For oral administration, the pharmaceutical compositions may take the form of, for example, tablets or capsules prepared by conventional means with pharmaceutically acceptable excipients such as binding agents (e.g., pregelatinised maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (e.g., lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (e.g., magnesium stearate, talc or silica); disintegrants (e.g., potato starch or sodium starch glycolate); or wetting agents (e.g., sodium lauryl sulphate). The tablets may be coated by methods well known in the art. Liquid preparations for oral administration may take the form of, for example, solutions, syrups or suspensions, or they may be presented as a dry product for constitution with water or other suitable vehicle before

35

use. Such liquid preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (e.g., sorbitol syrup, cellulose derivatives or hydrogenated edible fats); emulsifying agents (e.g., lecithin or acacia); non-aqueous vehicles (e.g., almond oil, oily esters, ethyl alcohol or fractionated vegetable oils); and preservatives (e.g., methyl or propyl-p-hydroxybenzoates or sorbic acid). The preparations may also contain buffer salts, flavoring, coloring and sweetening agents as appropriate. Preparations for oral administration may be suitably formulated to give controlled release of the active compound.

10

For buccal administration the compositions may take the form of tablets or lozenges formulated in conventional manner. For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of e.g., gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

15

20

The compounds may be formulated for parenteral administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampoules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

25

30

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides.

35

In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be

administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

The compositions may, if desired, be presented in a pack or dispenser device that may contain one or more unit dosage forms containing the active ingredient. The pack may for example comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration.

Examples

In order to identify potential susceptibility variants in the NHP2L1, PACSIN2, SERHL, PIPPIN, BRD1, EP300, FAM19A5 and/or GPR24 genes, the genes were sequenced in a subset of patients with mental disorders. The genomic sequences containing upstream promoter sequences, intronic sequences close to the exon/intron boundaries and coding sequences were analysed. In addition, potential SNPs were identified by searching databases such as the dbSNP (<http://www.ncbi.nlm.nih.gov/projects/SNP/>). The identified variants were analysed in a case-control sample from Scotland described below.

Example 1. Analysis of unrelated patients and ethnically matched unrelated controls from Scotland.

Subjects

The case-control sample from Scotland consisted of 103 patients with SZ, 162 patients with BPD and 200 ethnically matched controls. Informed consent was obtained from all patients prior to inclusion, and the study was approved by the local research ethical committees where patients were recruited. Subjects were interviewed by an experienced psychiatrist and venous blood taken for subsequent DNA extraction using routine procedures. Diagnoses were made according to DSM-IV criteria after case-note review and personal interview using the Schedule for Affective Disorders and Schizophrenia – Lifetime version. Final diagnoses were reached by consensus between two experienced psychiatrists (DB and WM). Control subjects (prescreened to exclude those with serious chronic illness) were drawn from the same population in South East and South Central Scotland and recruited from

Scottish National Blood Transfusion Service donors. We have reported these sample sets previously (17, 18).

Genotyping

5 The genes selected were included on the basis of their location, expression profile and existing knowledge of their function. The selection criteria for the single nucleotide polymorphisms (SNPs) were also based on a functional approach evaluating the type of SNP (prioritizing non-synonymous SNPs) and location (preferably promoter, UTR, intron/exon boundaries, and conserved regions).
10 Genotyping was performed using 40 ng of DNA per multiplex PCR. Exonuclease1 and Shrimp Alkaline Phosphatase were used for purification steps and the SNPs genotyped by multiplex single base extension technology using the ABI SNaP-shot kit and an ABI 310 Genetic Analyzer or a 3100 Avant Genetic Analyzer (Applied Biosystems, Foster City, CA) according to the manufacturer's recommendations. The
15 data was analyzed using the GeneScan 3.1.2 program (Applied Biosystems, Foster City, CA). Standard PCR conditions were used. The microsatellite markers were analyzed using fluorescent primers, standard conditions for (duplex) PCR amplification and separation of allelic fragments on an ABI 310 Genetic Analyzer. All primer sequences are available on request. To minimize genotyping errors all
20 polymorphisms were scored independently by two experienced investigators. Any discordances lead to re-analysis of the sample. In addition a number of SNPs (including *BRD1* rs4468 and rs138880) were analyzed twice in at least 85 individuals. No divergent genotypes were observed, indicating a very low error rate (for allele calls less than 0.006).

25

Statistical analysis

Chi-square and Fisher's Exact test were used to assess allele and genotype distributions. Haplotype Trend Regression (HTR) was used to estimate the frequency and analyze the distribution of haplotypes (19). When comparing two groups, HTR
30 produces an overall p-value for the observed distribution of all the haplotypes at an interval defined by a set of neighboring markers and also a haplotype-specific p-value describing the likelihood of the observed distribution of each of the individual haplotypes. The p-values from HTR presented in this study are empirical values based on up to 100,000,000 permutations. P-values less than 0.05 are referred to as
35 significant. The p-values presented are not corrected for multiple testing. However, the highly significant haplotype associations remained significant even after a

Bonferroni correction, which is overly conservative as the tests performed are not independent.

Tests of linkage disequilibrium were performed using the program Ldmax from the GOLD software package (<http://www.sph.umich.edu/csg/abecasis/GOLD>) which uses
5 the Slatkin and Excoffier expectation-maximization based approach (20).

In silico analysis

The impact of a promoter SNP on potential binding sites for transcription factors was analyzed using the program MatInspector (www.genomatix.de) (21, 22). This
10 program utilizes a library of matrix descriptions for transcription factor binding sites to identify potential sites in a sequence analyzed and assign a quality rating of matches (core and matrix similarity) estimating the influence of a SNP on the binding of transcription factors.

The possible effect of an intragenic SNP on splicing was investigated using the
15 programs ESEfinder release 2.0 (<http://rulai.cshl.edu/tools/ESE>) (21-23), RESCUE-ESE Web Server (<http://genes.mit.edu/burgelab/rescue-ese>) (24), FAS-ESS web server (<http://genes.mit.edu/fas-ess/>) (25), ExonScan Web Server (<http://genes.mit.edu/exonscan/>) (24-26) and NNSPLICE (http://www.fruitfly.org/seq_tools/splice) (27).

20 The effects of 3' UTR SNPs on microRNA binding sites were analyzed using the miRBase Targets Pre-release Version 1.0 (<http://microrna.sanger.ac.uk/targets/v1/>) which is a web resource provided by the Wellcome Trust Sanger Institute containing computationally predicted targets for microRNAs across a number of species.

25 **Northern blotting**

Human Multiple Tissue Northern Blots II and V (Clontech Laboratories, CA, USA) were probed with a α -³²P-CTP-labelled probe against *BRD1*. Full length *BRD1* was cloned from human brain cDNA (Clontech, #639300), gel purified (Qiagen, #28704) and used as template for the probe synthesis. The probe was synthesized and
30 labeled using 15,000 Ci/mmol α -³²P-CTP (EasyTides) and the Prime-It RmT Random Primer Labeling Kit (Stratagene, #300392) following the manufacturer's protocol. The specific activity of the probe was measured to 1.3×10^9 dpm/ μ g. The blots were hybridized and rinsed as described previously (28) and exposed to X-ray film for 3 weeks.

35

Preparation of rat, rabbit, human and fetal pig brain tissue

Seven male Wistar rats (250-300 g) and 5 male New Zealand white rabbits (2.5-3.5 kg) were deeply anesthetized before transcardial perfusion with 0.5 l (rats) or 1.5 l (rabbits) phosphate buffered 4% paraformaldehyde (pH 7.4) at 4°C as approved by the Danish Council for Animal Research Ethics. The brains were immersed in the same fixative for 24 hrs and divided into 2-3 minor coronal tissue blocks by histOmer embedding and sectioning on a HistOtech slicer (29) before vibratome sectioning into 50 µm sections.

Human brain tissue was obtained from 2 donors who had donated their remains for educational and scientific purposes at the Institute of Anatomy, University of Aarhus. The brains were removed 24-72 hrs postmortem and briefly stored in 10% formalin before smaller tissue blocks containing the frontal cortex were paraffin-embedded and microtome sectioned into 10 µm sections.

Fetal pig brain tissue was obtained from pregnant sows anesthetized by carbon dioxide and sacrificed by bleeding at embryonic day 40, 60, 80, 100 and 115, respectively. The removed fetal brain tissue was briefly immersed in 10% formalin before tissue blocks containing the forebrain (cortex cerebri) and the hindbrain (lower brainstem and cerebellum) were embedded in paraffin and microtome sectioned into 2 µm sections.

For quantitative analysis of the mRNA expression, the hippocampus, cortex, basal ganglia, cerebellum and brain stem were dissected from the fetal brains and immediately frozen in liquid nitrogen. After thawing on ice, total RNA was extracted with RNeasyTM-LysR (Qiagen) according to the manufacturer's protocol.

Immunohistochemistry

The anti-BRD1 monoclonal antibody used was provided by Bryan Young and was identical to that used in the initial BRD1 cloning study (30). It was raised against the 11 amino acid peptide RRPFSWEDVDR corresponding to amino acids 692-702 of BRD1. Vibratome and cryostat sections were initially blocked for endogenous biotin before preincubation with 1% Triton X-100 and 0.2% milk (Biding, Denmark) in TBS for 30 min followed by incubation at 4°C with the primary monoclonal antibody (mouse anti-BRD1) diluted 1:500-1:1000 in TBS containing 1% Triton X-100 and 0.2% milk for 72 hrs. After rinsing with TBS and 1% Triton X-100 for 3x15 min, the sections were incubated for 1 hr at room temperature with the secondary antibody (sheep anti-mouse IgG biotin-labeled, Amersham, RPN 1001) diluted 1:200-1:400 in TBS containing 1% Triton X-100 and 0.2% milk. Endogenous peroxidase activity was blocked with a

solution of 80 ml TBS with 10 ml H₂O₂ and 10 ml methanol for 10 min. Avidin-peroxidase (Sigma: A 3151) diluted 1:200-1:400 with TBS containing 1 % Triton X-100 and 0.2 % milk was then applied for 1 hr at room temperature. After rinsing for 3 x 15 min in TBS + 1 % Triton X-100, the formed avidin-peroxidase complexes were visualized by incubation for 10 min with diaminobenzidine (DAB) made by dissolving a 10 mg DAB-tablet (Kem-En-Tec Diagnostics A/S) in 10 ml water and immediately before use adding 10 µl of 35 % H₂O₂. After mounting and coverslipping with Depex, the sections were analyzed using a light microscope and compared with Nissl-stained sections, securing systematic analysis of consecutive coronal brain levels.

10 Paraffin embedded sections were stained according to the above mentioned protocol after initial target retrieval by microwave boiling of the sections for 10 min in a citrate buffer (pH=6.0) and usage of the primary monoclonal mouse anti-BRD1 antibody diluted 1:150.

Double immunofluorescence staining was performed on vibratome sections incubated for 72 hrs at 4°C with the primary monoclonal BRD1 antibody diluted 1:500, followed by a 1 hr incubation at room temperature with the secondary antibody (goat anti-mouse Ig FITC-labeled, Abcam, ab6785) diluted 1:600. The sections were then treated with 500 µg RNaseA (Roche, 109 142) for 20 min, before nuclear staining with TO-PRO-3 (Molecular Probes, T3605) diluted 1:1000 for 5 min, followed by mounting and coverslipping with Vecta-shield H1000 and subsequent confocal microscopic analysis.

cDNA synthesis and real-time qPCR

cDNA was synthesized from 1 µg total RNA in 20 µl reactions using iScript™ cDNA synthesis Kit (Biorad). After synthesis the cDNA was diluted five times with double distilled water. The real-time RT-PCR reactions were made with DyNAmo™SYBR® Green qPCR kit (Finnzymes) in a total volume of 20 µl, using 6 pmol primers specific for pig *BRD1* (forward: 5'-GGGCCAAGTGCAGCGGCTAC-3', reverse: 5'-CTCCATCATCTTCAGCTTGTC). Amplification was carried out on a Biorad iCycler using the following conditions: 95°C 15 min, (94°C 10 s, 58°C 20 s, 72°C 30 s) 40 repeats, 72°C 10 min, cooled to 20°C. PCR specificity was controlled by melting curve analysis. All reactions were also run using primers specific for pig-*GAPDH* as an internal standard (forward: 5'-GGGGAATTGCGCCACCATGGTGAAGGTCGGAGTGAAC-3', reverse: 5'-GGGGAATTGATGACAAGCTTCCCATTCTC-3').

The relative standard curve for the real-time qPCR reactions were made from RT-PCR of cDNA from the 115-day-old fetus hippocampus, amplified with the above mentioned conditions. The product was run on a 1% agarose gel, cut out and purified using QIAquick Gel Extraction Kit Protocol (Qiagen). The DNA was diluted and used to produce a relative standard curve showing above 90% PCR efficiency in the area. The specificity of the PCR product was verified by sequencing showing the position of the product to span the last two exons of the gene. Relative quantification of the expression was determined as follows. Three replicates of the threshold cycle for *BRD1* ($C_{T,X}$) and *GAPDH* ($C_{T,R}$) were measured. Subtracting each $C_{T,R}$ replicate from each $C_{T,X}$ nine threshold differences (ΔC_T) were calculated. Under assumptions of normality and variance homogeneity the temporal change in expression was examined by analysis of variance contrasting to embryonic day 115 the mean threshold difference from each of the other days. P-values are adjusted for the number of contrasts tested using the Sidák method. Note this analysis is a variant of the $2^{-\Delta\Delta C_T}$ method (31) with $\Delta\Delta C_T$ equal to the contrasts and results are presented in terms of fold changes $2^{-\Delta\Delta C_T}$ (Fig. 5). Moreover, a polynomial regression model up to third order was determined using a forward inclusion stepwise procedure. The significance of the highest order term is indicated and the resulting regression curve with 95 % confidence bands is shown after back-transformation by $2^{-\Delta C_T}$ (Fig. 5).

20

Results

Association analysis

5 SNPs in *BRD1*, 9 SNPs in 4 neighboring genes and two microsatellite markers were selected and genotyped in a Scottish case-control sample, including 103 SZ cases, 162 BPD cases, and 200 controls (Table 1). Rs3752466 turned out to be constant and was consequently excluded from further analysis. No significant deviation from Hardy-Weinberg equilibrium was observed for any of the SNPs in the case or control groups. Single-marker and multi-marker haplotype analysis was performed comparing controls to SZ, BPD, and the two case groups combined. (Table 1 around here).

30

Significant single-marker associations were observed for two *BRD1* SNPs and D22S1169 (Table 2). The promoter SNP rs138880 showed association in SZ, BPD and the combined case group with p-values of 0.0061, 0.0274 and 0.0046,

respectively. The 3' UTR SNP rs4468 and D22S1169 showed significant association with SZ ($P = 0.0088$ and 0.0214 , respectively) but not with BPD. These allelic associations were also found in the analysis of the genotypic distribution with similar p-values (results not shown). (*Table 2 around here*).

5

The haplotype analysis considering the overall distribution of all haplotypes of 2-4 neighboring markers in a sliding window fashion showed comparable results in both disorders (Table 2). Highly significant overall p-values (as low as 0.00001) were observed in especially the 3- and 4-marker analysis involving the *BRD1* SNPs
10 rs138855, rs2239848, rs138880 and rs138881. Haplotypes involving the microsatellite markers and the proximal *BRD1* SNPs showed primarily association in SZ.

Analysis of the individual haplotypes revealed that the overall association could be
15 attributed to both "risk" and "protective" haplotypes (Table 3). A 3-marker core haplotype spanning *BRD1* SNPs rs138855, rs2239848 and rs138880 had a frequency of around 9% in cases against only 1% in controls, producing a haplotype specific p-value of 10^{-06} . Removing the rare middle SNP rs2239848 from the analysis resulted in a reduced 2-marker "risk" haplotype (G-C) showing a frequency of ~10%
20 in cases versus ~1% in controls (p-value of 2.8×10^{-07} in the combined case group). (*Table 3 around here*).

"Protective" haplotypes over-represented among controls included both microsatellites, all the *BRD1* SNPs and extended into *MLC1* (megalencephalic
25 leukoencephalopathy with subcortical cysts 1) (Table 3). These related and rather frequent haplotypes showed primarily significant results when compared to their frequencies in SZ.

Linkage disequilibrium

30 A high degree of inter-marker linkage disequilibrium (LD) between the SNPs in *BRD1* was found in controls and cases (Table 4). Likewise a high LD between the SNPs in *MLC1* extending distally to *MOV10L1* (Moloney leukemia virus 10-like 1, homolog (mouse)) was seen in both cases and controls. No significant LD was seen between *BRD1* and *MLC1*. (*Table 4 around here*).

35

In silico analysis

Using MatInspector, the two promoter SNPs in *BRD1* (Table 1) were analyzed for potential effects on binding sites for transcription factors. The C-allele of rs138880 introduced binding sites for two transcription factors: The zinc finger binding protein factor encoded by *ZNF202* which is thought to predominantly regulate genes participating in lipid metabolism (32), and hairy and enhancer of split homolog 1 (HES-1) (showing a high core and matrix similarity), which is a transcriptional repressor inhibiting neural differentiation (33). Rs138881 did not introduce any changes.

The synonymous *BRD1* SNP rs2239848 located in exon 1 was analyzed for effects on exon splicing enhancers (ESE) and exon splicing silencers (ESS). According to ESEfinder the presence of the rare A-allele eliminated a binding site for the SR protein SRp55. However, analysis of the splice site in exon 1 revealed a very strong donor site thus suggesting a very limited potential effect of exon 1 ESEs. Neither RESCUE-ESE nor FAS-ESS identified any effect of rs2239848 on splicing.

Analysis of intronic and 3'UTR SNPs did not suggest any differential effects of the alleles.

Northern blotting

A *BRD1* transcript of the expected size (approximately 4.6 kb) was observed in most of the human brain regions tested, i.e. whole brain, cerebellum, cerebral cortex, medulla, spinal cord, occipital pole, frontal lobe, caudate nucleus, corpus callosum, hippocampus and thalamus (data not shown). A further faint band of a slightly larger size was present in whole brain, cerebellum and cerebral cortex, which suggests alternative splicing of the pre-mRNA that seems to be differently regulated across the human brain.

Immunohistochemistry

BRD1-immunostaining showed similar neuronal staining patterns in the adult rat, rabbit, and human cortex cerebri (Fig. 1). The neurons in cortex layers I-VI displayed prominent BRD1 immunoreactivity in the perikaryal cytosol surrounding a weaker granular staining of the nucleus (Figs. 1-2). BRD1 immunoreactivity was likewise seen in the proximal part of the primary dendrites, whereas the distal dendrites and the axon seemed unstained (Figs. 1-2). Glial staining was not noted in the human cerebral cortex or any part of the rat and rabbit CNS. Consecutive sectioning and immunostaining of the rat and rabbit brain confirmed that neuronal BRD1

immunoreactivity was distributed throughout the adult nervous system e.g. the cerebrum, brainstem, cerebellum and spinal cord (Fig. 3). (*Figs. 1-3 around here*).

5 Fetal pig brain tissue of embryonic day 40, 60 (Figs. 4A-C), 80, 100, and 115 (Fig. 4D-F) revealed dense nuclear staining in the neuroepithelial cell layer and the early differentiated neuroblasts (Figs. 4A-B). Medium differentiated neuroblasts displayed an intense nuclear and perikaryal cytosolar staining pattern (Fig. 4C), while fully differentiated neurons generally stained more weakly and in particular had a very weak nuclear staining compared to the staining intensity seen in the perikaryal
10 cytosol (Figs. 4D-F). (*Fig. 4 around here*).

mRNA expression in fetal pig brain

The level of expression of *BRD1* mRNA was measured and normalized to the amount of *GAPDH* mRNA, which was constantly expressed. While the overall trend
15 in the five examined areas of fetal pig brain was the same, the differences in expression levels were most pronounced in cortex cerebri in particular but also in the regions of the brainstem and basal ganglia (Fig. 5). The maximum amount of mRNA was measured at embryonic day 60 in all samples. The two areas that were possible to dissect in the 40-day-old embryo (cortex and cerebellum) showed a lower level of
20 expression. For some areas the abundance of mRNA was also relatively increased at day 80 (brainstem, cortex and basal ganglia). Between embryonic day 80 and 115 *BRD1* mRNA expression leveled off. (*Fig. 5 around here*).

25

30

5

Table 1 Genotyped polymorphisms and allele frequencies.

Gene	SNP	Location (Mb) ^a	Type of SNP	Alleles ^b	MAF		
					Controls	BPD	SZ
<i>FAM19A5</i>	rs132234	47.424523	Intron	C/T	0.27	0.29	0.27
<i>FAM19A5</i>	rs3752466	47.466709	3' UTR	C/T	0.00	0,00	0.00
	D22S922	47.491607	Microsatellite				
	D22S1169	47.722917	Microsatellite				
<i>BRD1</i>	rs4468	48.488513	3' UTR	T/C	0.35	0.37	0.49
<i>BRD1</i>	rs138855	48.519343	Intron	G/C	0.16	0.14	0.13
<i>BRD1</i>	rs2239848	48.537615	Syn.	G/A	0.01	0.02	0.01
<i>BRD1</i>	rs138880	48.539472	Promoter	A/C	0.16	0.22	0.25
<i>BRD1</i>	rs138881	48.541343	Promoter	G/A	0.10	0.12	0.13
<i>MLC1</i>	rs6010260	48.818300	Nonsyn.	G/T	0.13	0.14	0.16
<i>MLC1</i>	rs137931	48.826710	Promoter	C/-	0.26	0.22	0.27
<i>MLC1</i>	rs137932	48.826857	Promoter	G/A	0.25	0.22	0.27
<i>MOV10L1</i>	rs3810971	48.849123	Nonsyn.	C/T	0.24	0.26	0.20
<i>MOV10L1</i>	rs2272843	48.901923	Nonsyn.	C/A	0.15	0.13	0.12
<i>MAPK8IP2</i>	rs715519	49.328513	Promoter	C/G	0.18	0.16	0.19
<i>MAPK8IP2</i>	rs916005	49.334387	Intron	G/A	0.04	0.04	0.05

^a According to the UCSC Genome Browser, May 2004 assembly (<http://www.genome.ucsc.edu>)^b Major allele/minor allele on the +strand (<http://www.genome.ucsc.edu>)

Nonsyn=Nonsynonymous SNP, MAF= Minor allele frequency, Syn.=Synonymous SNP

Table 2 Single marker and overall haplotype association analysis

		Empirical overall p-values ^a			
Gene	SNP	Single marker	2-marker	3-marker	4-marker
BPD+SZ					
FAM19A5	rs132234	0,6116			
	D22S922	0,5552	0,5738		
	D22S1169	0,0383	0,4270	0,4635	
BRD1	rs4468	0,0215	0,3317	0,4030	0,1853
BRD1	rs138855	0,4186	0.3182	0,1374	0,1149
BRD1	rs2239848	0,3983	0.6658	0.2092	0,2630
BRD1	rs138880	0,0046	0,0381	0,00005	0.0607
BRD1	rs138881	0,1727	0.0681	0.1550	0.0003
MLC1	rs6010260	0,1631	0.4043	0.2102	0,3086
MLC1	rs137931	0,5796	0.1604	0.1926	0.2358
MLC1	rs137932	0,6327	0.9210	0.4051	0.3061
MOV10L1	rs3810971	0,9874	0.9882	0.8355	0.2924
MOV10L1	rs2272843	0,3018	0.1966	0.4691	0,4295
MAPK8IP2	rs715519	0,7976	0.7516	0.4568	0.4123
MAPK8IP2	rs916005	1.0000	0.9552	0.9786	0,7152
BPD					
FAM19A5	rs132234	0.6271			
	D22S922	0,0885	0,8920		
	D22S1169	0,4042	0,6196	0,4032	
BRD1	rs4468	0.7814	0,8945	0,8753	0,7872
BRD1	rs138855	0.4377	0,0769	0,5369	0,5341
BRD1	rs2239848	0,2501	0.5005	0.1742	0,5517
BRD1	rs138880	0.0274	0.0797	0,00006	0.1877
BRD1	rs138881	0.2473	0.1822	0.2096	0,0013
MLC1	rs6010260	0.4011	0,6883	0.6108	0.5603
MLC1	rs137931	0.2842	0.2795	0.1011	0.4430
MLC1	rs137932	0.3229	0.8343	0.3976	0.2014
MOV10L1	rs3810971	0.5368	0.8155	0.8016	0.7716
MOV10L1	rs2272843	0.5064	0.1395	0.3280	0.3263
MAPK8IP2	rs715519	0.4675	0.6392	0.3063	0.4475
MAPK8IP2	rs916005	0.8427	0.8895	0.8328	0.3167
SZ					
FAM19A5	rs132234	0.7013			
	D22S922	0,3501	0,5950		
	D22S1169	0,0214	0,0194	0,0426	
BRD1	rs4468	0,0088	0,0184	0,0657	0,0228
BRD1	rs138855	0.5010	0.1138	0,0117	0,0839
BRD1	rs2239848	1.0000	0.5944	0.0470	0,0095
BRD1	rs138880	0.0061	0.1610	0,00001	0.0203
BRD1	rs138881	0.2896	0.0521	0.1358	0,0002
MLC1	rs6010260	0,0823	0.3103	0,0763	0.1593
MLC1	rs137931	0.8440	0.1836	0.3326	0.1100
MLC1	rs137932	0.7615	0,8969	0,2821	0.4183
MOV10L1	rs3810971	0.3431	0,8971	0.9061	0.1839
MOV10L1	rs2272843	0.2701	0.3417	0,7132	0.7849
MAPK8IP2	rs715519	0.7254	0.6423	0.5626	0.5239
MAPK8IP2	rs916005	0.8482	0.9765	0.9187	0.7463

Table 3 Distribution of selected individual haplotypes.

Haplotype											Haplotype frequency			Empirical p-values ^a		
S1	M1	M2	S2	S3	S4	S5	S6	S7	S8	S9	Controls	BPD	SZ	Combined	BPD	SZ
"Risk" haplotypes																
				G		C					0,0105	0.0937	0.0993	2.8x10⁻⁰⁷	1.6x10⁻⁰⁶	1x10⁻⁰⁶
				G		C	A				0.0080	0.0591	0.0667	4x10⁻⁰⁵	1x10⁻⁰⁴	2x10⁻⁰⁴
				G		C	G				0.0027	0.0258	0.0326	0,0022	0,0018	0.0015
				G	G	C					0.0105	0.0895	0.0951	1x10⁻⁰⁶	3x10⁻⁰⁶	4x10⁻⁰⁶
				G	G	C	A				0.0081	0.0611	0.0663	2x10⁻⁰⁵	8x10⁻⁰⁵	2x10⁻⁰⁴
	2	3	C								<0.0001	<0.0001	0,0577	0,0248	-	0,0120
"Protective" haplotypes																
			2	T							0,2743	0,0797	0,1253	1.5x10⁻⁴	0,0389	5x10⁻⁴
	1		2	T							0,1740	0,0442	0,0300	1.8x10⁻⁰⁴	0,0865	4.3x10⁻⁰⁴
	1		5	T							0,0585	0,0536	<0.0001	0,0134	0,6006	0,0042
	1		2	T	G						0,2120	0,0675	0,0104	4x10⁻⁰⁵	0,0811	8.4x10⁻⁰⁴
	1		5	T	G						0,0458	0,0387	<0.0001	0,0207	0,5625	0,0046
	1	2/5	T	G							0,2463	0,1672	0,0119	5.7x10⁻⁰⁶	0,0691	7.8x10⁻⁰⁶
			2	T	G	G					0,2930	0,1330	0,0843	9.2x10⁻⁰⁵	0,0461	2.5x10⁻⁰⁴
			5	T	G	G					0,0401	<0.0001	<0.0001	0,0044	0,0334	0,0090
						A	G				0.8378	0.7830	0.7640	0.0232	0,0658	0,0314
						A	G	G			0.7369	0.6672	0.6185	0,0114	0,0785	0,0061
						A	G	G	C		0.5678	0.5072	0.3913	0,0254	0,2936	0,0019
						A	G	G	C	G	0.5693	0.5038	0.3912	0,0211	0,2452	0,0014
			T	G	G	A	G	G	C	G	0.3898	0.2554	0.2402	0.0030	0,1486	0,0039
			T	G	G	A					0.6313	0.5795	0.4809	0,0189	0,4404	0,0100
			G			A					0,8315	0.7685	0.7666	0,0217	0,0414	0,0683

^a Empirical haplotype specific p-values based on 100 000 000 permutations using the HTR program.

P-values < 0.05 in bold

S2 to S9 correspond to SNP rs4468, rs138855, rs2239848, rs138880, rs138881, rs6010260, rs137931, and rs137932

M1 to M2 correspond to D22S922 and D22S1169.

Table 4 Intermarker linkage disequilibrium measured by D'.
Cases above and right of diagonal, controls below and left of diagonal.

Gene	SNP	S1	M1	M2	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13
FAM19A5	rs132234 (S1)		0,05	0,06	0,12	0,53	1,00	0,39	<0,01	0,09	0,08	0,10	0,04	0,03	0,14	0,27
	D22S922 (M1)	0,08		0,13	0,02	0,44	0,12	0,07	0,06	0,21	0,02	0,00	0,00	0,07	0,02	0,06
	D22S1161 (M2)	0,13	0,22		0,15	0,16	0,44	0,14	0,16	0,22	0,11	0,14	0,19	0,34	0,08	0,25
BRD1	rs4468 (S2)	0,14	0,25	0,28		0,84	1,00	0,95	1,00	0,06	0,42	0,42	0,04	0,15	0,06	0,15
BRD1	rs138855 (S3)	0,04	0,06	0,17	0,90		0,59	0,91	0,39	0,47	0,17	0,17	0,12	0,03	0,05	<0,01
BRD1	rs2239841 (S4)	1,00	<0,01	<0,01	1,00	1,00		1,00	0,69	1,00	0,11	0,11	0,18	0,85	0,20	1,00
BRD1	rs138880 (S5)	0,08	0,09	0,20	0,92	0,94	1,00		0,98	0,31	0,07	0,07	0,05	0,09	0,01	0,09
BRD1	rs138881 (S6)	0,20	0,01	0,40	0,85	0,89	1,00	1,00		0,04	0,19	0,19	0,07	0,07	0,12	1,00
MLC1	rs6010261 (S7)	<0,01	0,44	0,15	0,02	0,19	1,00	0,03	0,07		1,00	1,00	0,53	1,00	0,29	<0,01
MLC1	rs137931 (S8)	0,30	0,10	0,29	0,45	0,01	<0,01	0,06	0,23	0,97		0,99	0,93	1,00	0,06	0,03
MLC1	rs137932 (S9)	0,29	0,09	0,29	0,45	0,01	0,01	0,06	0,24	0,96	1,00		0,93	1,00	0,06	0,04
MOV10L1	rs3810971 (S10)	0,31	0,13	0,16	0,05	0,53	0,04	0,70	0,71	0,54	0,87	1,00		0,95	<0,01	0,16
MOV10L1	rs2272841 (S11)	0,32	0,18	0,20	0,06	0,56	0,28	0,79	1,00	1,00	0,81	1,00	1,00		0,03	0,19
MAPK8IP2	rs715519 (S12)	0,13	0,13	0,27	0,58	<0,01	1,00	0,16	0,15	0,22	0,06	0,08	<0,01	0,10		<0,01
MAPK8IP2	rs916005 (S13)	0,89	0,29	0,52	0,20	0,11	1,00	0,16	0,10	1,00	0,34	0,33	0,20	0,34	0,50	

Significant ($P < 0.05$) D' values > 0.7 in bold

5 Example 2. Analysis of distantly related patients and unrelated controls from the Faeroe Islands (with focus on GPR24)

Subjects

Two samples were analyzed, one from the Faeroe Islands and one from Scotland.

10 Informed consent was obtained from all patients prior to inclusion, and the study was approved by the local research ethical committees where the patients were recruited. The patients from the Faeroe Islands are well-documented cases of severe SZ or BPD treated at the Department of Psychiatry, National Hospital, Torshavn and thoroughly interviewed and diagnosed by experienced psychiatrists according to

15 ICD10 diagnostic criteria for research and DSMIV (Jorgensen et al. 2002a). The genealogy of the distantly related 17 individuals with BPD and the 11 individuals with SZ was deduced from information on birth, marriages, and deaths in church and civic records of the Faeroese, and could be tracked back to a common ancestor born around 1600 (Fig. 3). The average number of generations relating two patients in the

20 genealogically shortest possibly way through one of the parents were six for patients with SZ and seven for patients with BPD. The sample included the cases previously analyzed for shared chromosome 22 segments using microsatellite markers (Jorgensen et al. 2002a) and four additional cases, one patient with SZ and three with BPD. The control group consisted of 44 unrelated persons (22 couples each with

25 a single offspring) from the Faeroe Islands without a history of psychiatric disease.

Haplotypes for chromosomal segments consisting of two to four neighboring markers were determined for cases on the basis of either available parental genotypes or genotypes of spouse and a child when available. All controls had their haplotype reconstructed from the genotypes of their offspring. This method will reconstruct the majority of relatively short haplotypes correctly

(Fig. 3 about here)

The case-control sample from Scotland consisted of 103 patients with SZ, 162 patients with BPD and 200 ethnically matched controls. The cases were diagnosed using SADS-L interview and RDC and DSM-IV criteria (Borglum et al. 2001; Borglum et al. 2003). The controls were from the Blood Transfusion Service, Edinburgh, and were screened to exclude people with serious chronic illness. Genomic DNA was isolated from blood samples according to standard procedures.

Sequencing and genotyping

Sequencing was carried out using 100 ng of DNA to perform PCR amplification, JETquick PCR Purification kit (Genomed GmbH, www.genomed-dna.com) for purification of the PCR product, and ABI BigDye kit for direct sequencing on an ABI310 Genetic Analyzer (Applied Biosystems, Foster City, CA). Sequences were analyzed in both directions. Genotyping of the selected SNPs was performed using 40 ng of DNA per multiplex PCR, Exonuclease1 and Shrimp Alkaline Phosphatase were used for purification steps and the SNPs were genotyped by multiplex single base extension technology using the ABI SNaP-shot kit and an ABI 310 Genetic Analyzer or a 3100 Avant Genetic Analyzer (Applied Biosystems, Foster City, CA) according to the manufacturer's recommendations. The data was analyzed using ABI 310 GeneScan 3.1.2 (Applied Biosystems, Foster City, CA). Standard PCR conditions were used for both sequencing and genotyping. Scoring of genotypes was performed by two investigators independently and in case of disagreement the sample was re-analyzed.

In order to further control for genotyping errors all SNPs were analyzed twice in at least 50 individuals, the GPR24 SNPs were analyzed twice in 100 individuals. No discordant genotypes were observed, indicating that the error rate was very low (for allele calls less than 0.01).

The microsatellite marker D22S279 was analyzed in the four newly ascertained Faeroese cases and the Scottish sample using fluorescent primers, standard conditions for PCR amplification, and separation of allelic fragments on an ABI310 Genetic Analyzer (Applied Biosystems, Foster City, CA).

Statistical analysis

The Faroese population

The data from the 6 polymorphic SNPs genotyped in the Faeroese sample and the new D22S279 genotypes were merged with the D22S279 genotypes produced by Jorgensen et al. (Jorgensen et al. 2002a), and a test of association was performed as implemented in CLUMP (Sham and Curtis 1995). The test is a modification of a Chi-squared test simulating the distribution of the test statistic using a Monte Carlo approach. The program evaluates all alleles or haplotypes in one test and is therefore sensitive to situations where more than one allele or haplotype are more frequent in either of the groups analyzed. In addition the Monte Carlo approach counteracts the invalidation of the asymptotic sampling distribution of the Chi-squared statistics potentially introduced by polymorphic markers. The p-values derived from CLUMP presented in this paper are from the subtests T1 and T4, which are the most reliable parameters when analyzing extended haplotypes. T1 is the standard Pearson χ^2 statistics of the 2xN contingency table and T4 is obtained by reshuffling alleles or haplotypes of a 2x2 table until χ^2 has reached a maximum, thereby comparing any combination of alleles or haplotypes with the rest. In the analysis of specific haplotypes, test of associations were performed using Fisher's Exact test.

Classical case-control analysis might detect differences between cases and controls owing to ignored population substructure or improperly accounted relatedness among individuals not necessarily owing to true association between a marker and a trait.

In the present dataset genealogical information is available for cases only (one of several genealogical routes is shown in Fig. 3), while the genealogy for controls remains unknown. In order to get an idea of how related controls are and whether they fall into the same genealogy as cases, we calculated pair-wise estimates of genetic relatedness (r) for all pairs and average relatedness estimates (r^*) for pairs within the two groups (cases and controls), but also for case-control pairs. Relatedness or relationship coefficients are defined as the proportion of genes/loci in one individual with alleles identical to these of a reference individual. Estimates of genetic relatedness were calculated using the algorithms developed by Queller & Goodnight (Queller and Goodnight 1989) as implemented in SPAGeDi 1.2 (Hardy and Vekemans 2002). Average within and between group relatedness estimates

were obtained using 60 randomly selected unlinked markers, standard errors were obtained by jack-knifing over loci. Parametric t-tests were used to test whether there were significant differences in average relatedness. Pair-wise relatedness coefficients for each pair of individuals were estimated using 660 markers more or less randomly distributed through out the genome (made available from an unpublished study). Using 660 markers of which some would be linked non-independent markers would overestimate the effective number of loci, resulting in underestimation of the variance (standard error). The actual value of relatedness coefficient based on all 660 markers should, however, be very accurate. Estimates of pair-wise genetic distances between individuals were obtained using the algorithm developed by Rousset (Rousset 2000) as implemented in SPAGeDi1.2 (Hardy and Vekemans 2002). Pair-wise genetic distances were used in a multidimensional scaling algorithm (Alscal procedure - as implemented in SPSS 11.5) to map similarity between individuals relative to each other. The population structure including the genetic differentiation within the case-control sample was evaluated by Wright's F-statistics. The 60 unlinked markers were used to calculate the genetic distance between the case and control group using Wright's F_{ST} . Under the hypothesis of no differentiation between the individuals and populations a null distribution of F_{IT} , F_{IS} and F_{ST} values was obtained by performing 3,000 permutations of individual genotypes among all individuals (F_{IT}), among individuals within populations (F_{IS}) and among populations (F_{ST}) as implemented in SPAGEDi 1.2 (Hardy and Vekemans 2002).

The Scottish population

For each of the seven polymorphic SNPs genotyped in the Scottish sample Chi-square and Fisher's Exact test were used to assess allele and genotype distribution and the program Haplotype Trend Regression (HTR) was used to estimate the frequency and analyze the distribution of haplotypes (Zaykin et al. 2002). When comparing two groups HTR produces an overall p-value for the observed distribution of all the haplotypes of a given segment and in addition a haplotype-specific p-value describing the likelihood of the observed distribution of each of the specific haplotypes. The p-values from HTR presented in this study are empirical values based on 100,000 permutations. In both samples analyzed the controls were compared to individuals with BPD, to individuals with SZ and to the two groups

combined. P-values less than 0.05 are referred to as significant. No correction for multiple testing was performed.

Pair-wise linkage disequilibrium was tested using the Slatkin and Excoffier expectation-maximization algorithm (Excoffier and Slatkin 1995) as implemented in

5 Ldmax from the GOLD software package

(<http://www.sph.umich.edu/csg/abecasis/GOLD/index.html>).

In silico analyses

The impact of a promoter SNP on potential binding sites for transcription factors was

10 analyzed using the program MatInspector (www.genomatix.de) (Quandt et al. 1995;

Werner 2000). This program utilizes a library of matrix descriptions for transcription factor binding sites to identify potential sites in a sequence analyzed and assign a

quality rating of matches (core and matrix similarity) estimating the influence of a SNP on the binding of transcription factors. The possible effect of an intragenic SNP

15 on splicing was investigated using the programs ESEfinder release 2.0

(<http://rulai.cshl.edu/tools/ESE>) (Cartegni et al. 2003; Quandt et al. 1995; Werner

2000), RESCUE-ESE Web Server (<http://genes.mit.edu/burgelab/rescue-ese>)

(Fairbrother et al. 2002) and NNSPLICE (http://www.fruitfly.org/seq_tools/splice)

(Reese et al. 1997). ESEfinder identifies putative exon splicing enhancers (ESE)

20 responsive to the human SR proteins SF2/ASF, SC35, SRp40 and SRp55. RESCUE-ESE Web Server predicts which sequences have ESE activity by statistical

analysis of exon-intron and splice site compositions. NNSPLICE analyze the

structure of the donor and the acceptor sites using a neural network recognizer.

25 **Results**

SNPs

In search for potential susceptibility variants the coding region (1269 bp in 2 exons), intron-exon boundaries, and the promoter region (500bp upstream to the transcription initiation site) of *GPR24* were sequenced in five individuals with SZ, four

30 with BPD and one control person from the Faeroese sample. Two SNPs (rs133070

and rs133073) were identified. In addition the dbSNP database

(<http://www.ncbi.nlm.nih.gov/SNP>) and the genome browser of University of

California Santa Cruz (<http://www.genome.ucsc.edu/>) were used for selection of two

additional SNPs in *GPR24*, two SNPs in *ADSL* and two SNPs in *ST13* (Table 1).

35 Eight markers (including D22S279) were genotyped in the sample from the Faeroe

Islands and all 10 markers in the Scottish sample. Rs5757921 and rs133071 turned

out to be monomorphic and were excluded from further analysis. All SNPs were found to be in Hardy-Weinberg equilibrium in both samples (results not shown).

(Table 1 about here)

5 Association analysis of the Faeroese sample

The Faeroese sample was analyzed using CLUMP comparing the controls to individuals with BPD, to individuals with SZ and to the two groups combined, and significant associations in all 3 groups were observed (Table 2).

Several single markers showed significant association. The three *GPR24* SNPs
10 rs133068, rs133069 and rs133073 showed association with SZ yielding p-values of 0.008-0.02. In BPD and BPD/SZ combined, single marker association was observed for rs909669 (*ADSL*) and rs133070 (*GPR24*) with p-values between 0.0036 and 0.037.

Significantly skewed overall distribution of 2-, 3-, and 4-marker haplotypes involving
15 all four *GPR24* SNPs were found when comparing controls to BPD with T1 p-values as low as 0.0009. The strongest signal was centered on 2-marker haplotypes from rs133069 to rs133073 in *GPR24*. When comparing controls to SZ unequal haplotype distribution was observed for 2-, 3-, 4- and 5-marker haplotypes spanning all four SNPs in *GPR24* in addition to rs909669 in *ADSL* with a 5-marker haplotype showing
20 a T4 p-value of 0.0054. Similarly, comparing controls to the combined group of cases revealed significant associations with 2-, 3-, 4- and 5-marker haplotypes spanning the SNPs in *GPR24* as well as rs909669 in *ADSL*. The strongest signal was observed for the same 2-marker haplotypes as in BPD (minimal p-value of 0.004) and the 4-/5-marker haplotypes yielding maximum signals in SZ (p-values as low as
25 0.002).

(Table 2 about here)

The distribution of the specific haplotypes contributing to the overall signal is summarized in Table 3. The overall signal in BPD was strongest for 2- and 3-marker haplotypes involving rs133069 to rs133073. The specific 2-marker haplotype
30 containing the G and C alleles of the *GPR24* SNPs rs133070 and rs133073, respectively, was present in 9% of the chromosomes in bipolar cases and 39% in controls yielding a p-value of 0.0165. Another 2-marker haplotype A-C involving the same SNPs had a frequency of 23% in BPD and 0% in controls (p-value of 0.0048). In SZ the overall signal appeared strongest for 5-marker haplotypes spanning the
35 segment from rs909669 to rs133070. This signal was mainly due to the 5-marker haplotype C7CCA, which was overrepresented in cases. It was present in 90% of the

chromosomes in SZ against only 9% in the controls giving a p-value of 5×10^{-6} . A more modest overall signal appearing in 2-marker haplotypes covering rs909669 to rs133069 was due to part of the same specific 5-marker haplotype described above.

The overall signal in the combined group reflected the strongest of the signals in BPD and SZ and correlated to the same specific haplotypes. The specific 5-marker haplotype straddling rs909660 to rs133070 as seen in SZ (C7CCA) showed a p-value of 0.0006 in the combined sample, and together with a variant of this haplotype (C6CCA) the haplotypes C-6/7-CCA were overrepresented in cases with a frequency of 75% against only 22% in controls (p-value of 7×10^{-5}). Finally, the same specific 2- and 3-marker haplotypes covering rs133069 to rs133073 as seen in BPD showed p-values as low as 0.0183 in the combined sample (Table 3).

(Table 3 about here)

The within group estimates of relatedness among cases with bipolar affective disorder did not differ significantly from the between group relatedness estimates ($r_{bp} = -0.0272 \pm 0.0107$ vs. $r_{bp \leftrightarrow sz+con} = -0.0203 \pm 0.0040$, $t_{17,72} = 0.7102$, $p = 0.4795$), nor did the average relatedness among cases with schizophrenia differ significantly from the between group relatedness estimate ($r_{sz} = -0.0223 \pm 0.0126$ vs. $r_{sz \leftrightarrow bp+con} = -0.0152 \pm 0.0045$, $t_{11,72} = 0.5675$, $p = 0.5719$). Considering the combined dataset of cases; the within group relatedness estimate of cases for both disorder did not differ significantly from the estimated relatedness between the case and control group ($r_{bp+sz} = -0.0208 \pm 0.0063$ vs. $r_{bp+sz \leftrightarrow con} = -0.0189 \pm 0.0023$, $t_{28,72} = 0.3536$, $p = 0.7244$). Individuals within the two case groups (considered separately and together) are therefore not significantly more related to each other than they are to individuals outside the group. However, the within group estimate did differ significantly from the between group estimate for controls ($r_{con} = -0.0059 \pm 0.0041$ vs. $r_{con \leftrightarrow bp+sz} = -0.0189 \pm 0.0023$, $t_{44,72} = 2.9902$, $p = 0.0034$), indicating that controls are in fact on average more related to each other than they are to individuals outside the group.

Overall within group estimated relatedness did not differ significantly from the average between group relatedness estimate ($r_{within\ group} = -0.0143 \pm 0.0005$ vs. $r_{between\ groups} = -0.0185 \pm 0.0024$, $t_{72,72} = 1.7132$, $p = 0.0889$).

Multidimensional scaling of pair-wise genetic distances (Rousset 2000) between individuals did not reveal an overall clustering of cases in relation to controls (results not shown). Based on the pair-wise estimates of relatedness or genetic distance between each pair of individuals, some individuals appeared to be more related than the population average, thus sharing more alleles than expected based on the population allele frequencies. This was, however, not consistently within groups.

The amount of genetic differentiation (F_{ST}) between cases and controls was not statistically significant ($F_{ST} = 0.0014$, $P_{two-tailed} = 0.4595$, 3,000 permutations). Likewise there was no evidence for inbreeding within individuals neither relative to the total sample nor relative to subgroups ($F_{IT} = -0.0008$, $P_{two-tailed} = 0.9180$; $F_{IS} = -0.0022$; $P_{two-tailed} = 0.8121$, 3,000 permutations, when cases and controls are considered as two subpopulations). Combining the two case groups did, however, reveal genetic differentiation among the case groups and the control group ($F_{ST} = 0.0034$, $P_{two-tailed} = 0.0330$; $F_{IT} = 0.0004$, $P_{two-tailed} = 0.9540$; $F_{IS} = -0.0030$, $P_{two-tailed} = 0.7414$, 3,000 permutations), increasing the risk for false-positive findings when combining the two disorders.

Association analysis of the Scottish case-control sample

In the Scottish sample only D22S279 showed significant single marker association with SZ, while haplotype analysis revealed significant associations in both disorders (Table 4). In BPD a minimal overall p-value of 0.0003 was observed for haplotypes including all four *GPR24* SNPs. In SZ the maximal signal was slightly more proximal, including 2 *GPR24* SNPs and D22S279 ($p=0.0005$). In the combined group of cases similar but less significant associations were observed.

(Table 4 about here)

Some of the "risk" haplotypes identified in the Faeroese sample were also found over-represented among the Scottish patients (Table 5). These (C)7CCA(T) haplotypes were predominantly over-represented in Faeroese SZ and in the Scottish BP, while the related (C)2CCA haplotypes, which were not present in the Faeroese population, were over-represented among Scottish SZ patients. The rest of the individual haplotypes found associated in the Scottish sample differed from those identified in the Faeroese sample. For example the 2-marker haplotypes containing either the 2, 4 or 8 allele of D22S279 in conjunction with the *GPR24* rs133068 G-allele, which had a combined frequency of 9.3% in SZ versus 1.5% in controls ($p=9.8 \times 10^{-5}$).

(Table 5 about here)

Linkage disequilibrium

A high degree of intermarker linkage disequilibrium (LD) was observed especially between the closely located SNPs in *GPR24*, extending to some degree to the more distal *ST13* SNP rs710193 in the Scottish sample, and in the Faeroese sample centromeric to *ADSL* (Table 6).

(Table 6 about here)

In silico analysis

Using the Matinspector (www.genomatix.de) the different alleles of the four promoter SNPs (Table 1) were analyzed for potential effects on binding sites for transcription factors. The C allele of the *GPR24* SNP rs133068 introduced a binding site for the transcription factor ZBP-89 whereas the G allele introduced a binding site for X-box binding protein RFX1. In both cases a high core- and matrix similarity were seen. The C allele of rs133069 introduced binding sites for SP1, TGFbeta, RREB1, ZBP-89 and ZIC2 of which especially ZBP-89 and ZIC2 showed high core and matrix similarity. Particularly the transcription factor ZIC2 is interesting since the ZIC genes play an important role in neural development (Aruga 2004). For rs133070 and rs909669 alternative alleles did not cause any changes.

The synonymous *GPR24* SNP rs133073 was analyzed for effects on exon splicing enhancers (ESE) and donor and acceptor sites. ESEfinder identified a single potential exon splicing enhancer that only appeared when the T allele was present, creating a binding site for the splicing factor SRp40 with the score 3.3, which is just above the threshold of 2.67 indicating a significant score.

Table 1 Genotyped polymorphisms and allele frequencies in samples from the Faeroe Islands and Scotland.

Gene	Marker	Location (Mb) ⁱ	Type of marker	Alleles ^b	MAF, Faeroe Islands			MAF, Scotland		
					Controls	BPD	SZ	Controls	BPD	SZ
<i>ADSL</i>	rs909669	39.066917	Promoter	C/T	0.15	0.00	0.00	0.10	0.11	0.15
<i>ADSL</i>	rs5757921	39.067154	Nonsyn	G/A	NG	NG	NG	0.00	0.00	0.00
	D22S279	39.347314	Microsatellite		-	-	-	-	-	-
<i>GPR24</i>	rs133068	39.398907	Promoter	C/G	0.45	0.42	0.15	0.48	0.47	0.49
<i>GPR24</i>	rs133069	39.398962	Promoter	C/A	0.48	0.59	0.15	0.49	0.47	0.48
<i>GPR24</i>	rs133070	39.399273	Promoter	A/G	0.44	0.16	0.22	0.42	0.42	0.39
<i>GPR24</i>	rs133071	39.399732	Promoter	C/T	0.00	0.00	0.00	0.00	0.00	0.00
<i>GPR24</i>	rs133073	39.400195	Synonymous	T/C	0.50	0.41	0.19	0.41	0.45	0.41
<i>ST13</i>	rs710193	39.547690	Nonsyn	C/T	0.20	0.16	0.22	0.10	0.12	0.13
<i>ST13</i>	rs1573745	39.565335	Intron	G/A	NG	NG	NG	0.07	0.05	0.10

^a According to the UCSC Genome Browser, May 2004 assembly (<http://www.genome.ucsc.edu>)

^b Major allele/minor allele on the +strand (<http://www.genome.ucsc.edu>)

Nonsyn=Nonsynonymous SNP, MAF= Minor allele frequency, NG=Not genotyped

Table 2 T1 and T4 p-values from CLUMP association analysis of the Faeroese sample.

Gene	Marker	Single marker		2-marker		3-marker		4-marker		5-marker	
		T1	T4	T1	T4	T1	T4	T1	T4	T1	T4
BPD+SZ											
ADSL	rs909669	0.0036	0.0036								
	D22S279	0.0705	0.0268	0.0805	0.0501						
GPR24	rs133068	0.1461	0.1461	0.1818	0.1270	0.4823	0.3996				
GPR24	rs133069	0.0740	0.0740	0.0716	0.1107	0.2017	0.2023	0.2672	0.2064		
GPR24	rs133070	0.0116	0.0116	0.0145	0.0828	0.0525	0.1831	0.0064	0.0023	0.0059	0.0031
GPR24	rs133073	0.0546	0.0546	0.0040	0.0188	0.0268	0.0626	0.0479	0.1278	0.0063	0.0023
ST13	rs710193	0.7631	0.7631	0.4522	0.4361	0.1158	0.1697	0.1755	0.2417	0.1722	0.239
SZ											
ADSL	rs909669	0.1178	0.1178								
	D22S279	0.0322	0.0174	0.0447	0.008						
GPR24	rs133068	0.0197	0.0197	0.0542	0.0476	0.0990	0.0353				
GPR24	rs133069	0.0100	0.0100	0.0469	0.0353	0.1870	0.1727	0.1481	0.0495		
GPR24	rs133070	0.2107	0.2107	0.3299	0.3299	0.4456	0.4456	0.0426	0.0367	0.0156	0.0054
GPR24	rs133073	0.0083	0.0083	0.5104	0.5104	0.3269	0.3269	0.4486	0.4486	0.0432	0.0375
ST13	rs710193	0.7065	0.7065	0.3844	0.3368	0.3917	0.5842	0.3957	0.5862	0.3932	0.5852
BPD											
ADSL	rs909669	0.0373	0.0373								
	D22S279	0.3548	0.2390	0.4799	0.3873						
GPR24	rs133068	0.8365	0.8365	0.4893	0.5473	0.8745	0.7353				
GPR24	rs133069	0.6815	0.6815	0.6484	0.5936	0.5063	0.4663	0.7976	0.5424		
GPR24	rs133070	0.0195	0.0195	0.0030	0.0159	0.0316	0.0972	0.1531	0.1076	0.2957	0.1557
GPR24	rs133073	0.4217	0.4217	0.0009	0.0093	0.0042	0.0303	0.0433	0.0997	0.1521	0.1056
ST13	rs710193	1.0000	1.0000	0.8264	0.8600	0.0850	0.1940	0.1753	0.2821	0.1745	0.2831

P-values < 0.05 in bold

5

Table 3 Distribution of selected individual haplotypes from CLUMP analysis of the sample from the Faeroe Islands.

Haplotype						Haplotype frequency				P-values		
S1	M	S2	S3	S4	S5	Controls	BPD	SZ	BPD+SZ	BPD	SZ	BPD+SZ
"Risk" haplotypes												
C	7	C	C	A		0.09	0.28	0.90	0.50	0.1179	5×10^{-6}	0.0006
C	6	C	C	A		0.13	0.39	0.00	0.25	0.0406	0.5569	0.3176
C	6/7	C	C	A		0.22	0.67	0.90	0.75	0.0026	0.0002	7×10^{-5}
	7	C	C	A	T	0.15	0.39	0.64	0.44	0.2868	0.0013	0.0139
	6	C	C	A	T	0.12	0.28	0.14	0.28	0.0340	1.0000	0.2166
	6/7	C	C	A	T	0.26	0.67	0.78	0.72	0.0076	0.0013	0.0005
			A	C		0.00	0.23	0.00	0.14	0.0048	1.0000	0.0234
"Protective" haplotypes												
		A	G			0.38	0.08	0.22	0.16	0.0393	0.3279	0.0364
		A	G	C		0.38	0.10	0.21	0.15	0.0307	0.3279	0.0526
			G	C		0.39	0.09	0.21	0.14	0.0165	0.3286	0.0183

Table 4 Single marker and haplotype association analysis in the case-control sample from Scotland

		Empirical overall p-values ^a				
Gene	SNP	Single marker	2-marker	3-marker	4-marker	5-marker
<u>BPD+SZ</u>						
ADSL	rs909660	0,2189				
D22S279		0,4644	0,0490			
GPR24	rs133068	0,8538	0,1821	0,0319		
GPR24	rs133069	0,7126	0,1791	0,3066	0,0591	
GPR24	rs133070	0,6721	0,4592	0,5945	0,3363	0,3390
GPR24	rs133073	0,5632	0,0262	0,0155	0,0184	0,1747
ST13	rs710193	0,2317	0,1247	0,0715	0,0790	0,1895
ST13	rs1573745	0,7989	0,4429	0,0632	0,1346	0,2604
<u>SZ</u>						
ADSL	rs909660	0,0655				
D22S279		0,0277	0,0012			
GPR24	rs133068	0,8145	0,0054	0,0015		
GPR24	rs133069	0,8734	0,1238	0,0005	0,0012	
GPR24	rs133070	0,3638	0,0618	0,0691	0,0053	0,0497
GPR24	rs133073	0,8469	0,1879	0,0148	0,0135	0,0189
ST13	rs710193	0,2758	0,3269	0,2240	0,1069	0,1460
ST13	rs1573745	0,2879	0,1394	0,2105	0,2465	0,1871
<u>BPD</u>						
ADSL	rs909660	0,6833				
D22S279		0,5453	0,0580			
GPR24	rs133068	0,6757	0,6295	0,3278		
GPR24	rs133069	0,6748	0,3582	0,7233	0,5948	
GPR24	rs133070	0,9373	0,1855	0,0779	0,3612	0,4471
GPR24	rs133073	0,3470	0,1471	0,0053	0,0003	0,0410
ST13	rs710193	0,3292	0,1197	0,2251	0,0252	0,0018
ST13	rs1573745	0,1588	0,5903	0,1230	0,3998	0,1587

^a Empirical overall p-values based on 100,000 permutations using the HTR program.

P-values < 0.05 in bold.

96

Table 5 Distribution of selected individual haplotypes in the case-control sample from Scotland.

Haplotype							Haplotype frequency			Empirical p-values ^a		
S1	M	S2	S3	S4	S5	S6	Controls	BPD	SZ	BPD	SZ	BPD+SZ
"Risk" haplotypes												
C	7	C	C	A			0.1564	0.2459	0.1783	0.0520	0.8847	0.1120
	7	C	C	A			0.1560	0.2518	0.1660	0.0221	0.8869	0.1040
	7	C	C	A	T		0.1663	0.2470	0.1659	0.0499	0.9698	0.1966
C	2						0.0115	<0.0001	0.0909	0.2349	0.0005	0.1118
	2	C	C	A			0.0039	<0.0001	0.0604	0.4531	0.0022	0.1005
	8	G					0.0037	0.0173	0.0219	0.0691	0.0229	0.0559
	2	G					0.0074	<0.0001	0.0434	0.1632	6.90x10-4	0.1366
	4	G					0.0037	0.0038	0.0287	1.0000	0.0247	0.3235
	2/4/8	G					0.0148	0.0208	0.0934	0.6185	9.78x10-5	0.0152
	2/4/8	G	A				0.0147	0.0208	0.0942	0.6200	9.40x10-5	0.0150
T	5						<0.0001	0.0355	0.0405	0.0041	0.0025	0.0058
T	5	G					<0.0001	0.0180	0.0246	0.0164	0.0134	0.0181
T	5	G	A				<0.0001	0.0179	0.0250	0.0191	0.0136	0.0179
T		G	A	G			0.0169	0.0621	0.0380	0.0474	0.3964	0.0330
T		G	A	G	C		0.0129	0.0619	0.0236	0.0296	0.4308	0.0263
"Protective" haplotypes												
		G	A	G	T		0.0202	<0.0001	<0.0001	0.0257	0.0675	0.0006
			A	G	T		0.0202	<0.0001	<0.0001	0.0264	0.0683	0.0006
				G	T		0.0254	0.0041	<0.0001	0.0252	0.0279	0.0050
		G	A	A	T		0.0537	0.0082	0.0897	0.0021	0.0841	0.4945
		G	A	A			0.0522	0.0162	0.0977	0.0256	0.0442	0.9692
			A	A	T		0.0537	0.0165	0.0949	0.0247	0.0534	0.8646
C	3						0.2634	0.2675	0.1704	0.8912	0.0314	0.2608
C	5						0.1718	0.1531	0.0958	0.7527	0.0534	0.2421
C	3/5						0.4353	0.4203	0.2662	0.7444	0.0018	0.0750
C	5	G	A				0.1614	0.1307	0.0719	0.5996	0.0247	0.1250
C	3/5	G					0.3199	0.3208	0.2080	0.9939	0.0204	0.2432
	3/5	G					0.3339	0.3454	0.2256	0.8470	0.0351	0.3597

^a Empirical overall p-values based on 100,000 permutations using the HTR program.

P-values < 0.05 in bold.

S1 to S6 correspond to SNP: rs909669, rs133068, rs133069, rs133070, rs133073, rs710193.

M= D22S279.

Table 6 Intermarker linkage disequilibrium measured by D'.
Cases above and right of diagonal, controls below and left of diagonal.

Gene	SNP		S1	M	S2	S3	S4	S5	S6	S7
Faeroese sample										
ADSL	rs909669	(S1)		0.00	0.00	0.00	0.00	0.00	0.00	
	D22S279	(M)	0.30		0.73	0.73	0.82	0.84	0.54	
GPR24	rs133068	(S2)	0.68	0.43		1.00	1.00	0.92	0.41	
GPR24	rs133069	(S3)	0.72	0.52	1.00		1.00	0.91	0.41	
GPR24	rs133070	(S4)	1.00	0.46	1.00	1.00		1.00	0.36	
GPR24	rs133073	(S5)	0.75	0.58	0.94	1.00	1.00		0.39	
ST13	rs710193	(S6)	1.00	0.43	0.61	0,61	1.00	0.64		
Scottish sample										
ADSL	rs909669	(S1)		0,32	0.10	0.11	0.14	0.01	0.05	0.20
	D22S279	(M)	0,46		0,38	0,37	0,48	0,46	0,23	0,23
GPR24	rs133068	(S2)	0,19	0,46		0.95	0.93	0.92	0.58	0.60
GPR24	rs133069	(S3)	0.20	0,46	0,99		0.89	0.89	0.56	0.60
GPR24	rs133070	(S4)	0,58	0,55	0,94	0,95		0.99	0.51	0.85
GPR24	rs133073	(S5)	0,55	0,61	0,91	0,92	0.95		0.48	0.86
ST13	rs710193	(S6)	1.00	0,70	0.90	0,89	0.90	0.90		0.63
ST13	rs1573745	(S7)	0,28	0,55	0,42	0,43	0,49	0,61	1.00	

Significant (P<0.05) D' values >0.7 in bold

Example 3

Association in three different case-control samples

5

Subjects

Three Caucasian case-control samples from Scotland, UK and Denmark respectively were analyzed. All patients gave informed consent prior to inclusion, and study approval was obtained from the local research ethical committees where the patients were recruited.

10

The case-control sample from Scotland consisted of 103 patients with SZ, 162 patients with BPD and 200 ethnically matched controls. The subjects suffering from SZ or BPD were interviewed using the Schizophrenia and Affective Disorder Schedule-Lifetime Schedule (SADS-L)⁴⁰ (SADS-L) interview and diagnosed according to RDC and DSM-IV criteria.^{41,42} The controls were from the Blood Transfusion Service, Edinburgh, and were screened to exclude people with serious chronic illness.^{43,44}

15

The UK sample consisted of 300 individuals with BPD, 265 individuals with SZ and 314 screened normal controls. The BPD cases and controls were included if both

20

parents and all four grandparents were of Irish, Welsh, Scottish or English ancestry as defined by an ancestry checklist. In the selection of the SZ cases one of the grand parents was allowed to be of Caucasian European origin but not of Jewish or non-European Union ancestry (EU countries before the 1994 enlargement). All subjects, were interviewed with the SADS-L by a psychiatrist and diagnosed according to RDC criteria. Cases with bipolar disorder were all bipolar 1 disorder. In the SZ group patients with schizoaffective bipolar disorder or schizo-mania were not included.

The controls were selected on the basis of not having a family history of schizophrenia, alcoholism or bipolar disorder and for not having a past or present personal history of any defined mental disorder.⁴⁵

The Danish sample consisted of 124 bipolar patients, 115 individuals suffering from schizophrenia and 96 ancestrally matched unscreened controls. Only patients with an age of onset below 35 were included. Cases were interviewed using the semi-structured interview SCAN version 2.1.⁴⁶ and best estimate diagnoses according to ICD-10-DCR⁴⁷ and DSM-IV⁴⁸ were made by two psychiatrists.. All bipolar cases met the ICD-10 –DCR diagnostic criteria for bipolar affective disorder and bipolar I disorder (DSM-IV).

The 115 individuals suffering from schizophrenia were first ever admitted cases fulfilling DSM-IV and ICD-10-DCR criteria for schizophrenia.

Genotyping and sequencing

Genomic DNA was isolated from blood samples according to standard procedures. Sequencing was carried out using 100 ng of DNA to perform PCR amplification, JETquick PCR Purification kit (Genomed GmbH, www.genomed-dna.com) for purification of the PCR product, and ABI BigDye kit for direct sequencing on an ABI310 Genetic Analyzer (Applied Biosystems, Foster City, CA). Sequences were analyzed in both directions.

Genotyping was carried out using the ABI SNaP-shot kit (Applied Biosystems, Foster City, CA), an ABI 310 Genetic Analyzer or a 3100 Avant Genetic Analyzer (Applied Biosystems, Foster City, CA), and the program ABI 310 GeneScan 3.1.2 (Applied Biosystems, Foster City, CA). Primer sequences were obtained from dbSNP (www.ncbi.nlm.nih.gov/SNP). PCR was performed with up to 5 primer sets simultaneously using 40 ng of DNA and standard PCR conditions (primer sequences and PCR conditions are available on request). The PCR products were treated with Exonuclease1 and Shrimp Alkaline Phosphatase, and the multiplex single base

extension was carried out according to the manufacturer's recommendations (Applied Biosystems, Foster City, CA).

To control for genotyping errors all SNPs were scored independently by two investigators. Any discordances lead to re-analysis of the sample. Furthermore a
 5 number of SNPs (including *SERHL* rs881542) were analyzed twice in at least 85 individuals. No divergent genotypes were observed, indicating a very low error rate.

Statistical analysis

The data from the polymorphic SNPs genotyped in the samples from Scotland,
 10 Denmark, and the UK were analyzed using the program Haplotype Trend Regression (HTR), which estimate the frequency and the distribution of single markers and haplotypes.⁴⁹ When comparing two groups HTR produce an overall p-value for the observed distribution of all the different haplotypes. In addition HTR produce a haplotype-specific p-value describing the likelihood of the observed distribution of
 15 each of the specific haplotypes. In this study the empirical p-values presented are based on up to 100,000,000 permutations. In the analyses performed the controls were compared to individuals with BPD, to individuals with SZ as well as to these groups combined. P-values less than 0.05 are referred to as significant. No correction for multiple testing was performed. Hardy Weinberg equilibrium for the
 20 individual SNPs was analyzed using χ^2 .

Analyzes of linkage disequilibrium (LD) was performed using the program Ldmax from the GOLD software package
 (http://www.sph.umich.edu/csg/abecasis/GOLD/index.html).^{50,51}

25 *In silico* analyses

MatInspector (www.genomatix.de)^{52,53} was used to assess and analyze the impact of promoter SNPs on potential binding sites for transcription factors. The possible effect of intragenic SNPs on splicing was investigated using the programs ESEfinder release 2.0 (<http://rulai.cshl.edu/tools/ESE/>),⁵²⁻⁵⁴ RESCUE-ESE Web Server,
 30 (<http://genes.mit.edu/burgelab/rescue-ese/>),⁵⁵ FAS-ESS web server (<http://genes.mit.edu/fas-ess/>),⁵⁶ ExonScan Web Server (<http://genes.mit.edu/exonscan/>),⁵⁶
 and NNSPLICE (http://www.fruitfly.org/seq_tools/splice).⁵⁷

35 The effect of 3' UTR SNPs on microRNA binding sites were analyzed using the miRBase Targets Pre-release Version 1.0 (<http://microrna.sanger.ac.uk/targets/v1/>),

a web resource provided by the Wellcome Trust Sanger Institute containing computationally predicted targets for microRNAs across a number of species.⁵⁸

Results

5

SNPs

In search for susceptibility variants sequencing of the coding region, intron-exon boundaries, and the promoter region (500bp upstream to the transcription initiation site) of the gene *PIPPIN* was carried out in five individuals with SZ, four with BPD and one control person from the Faeroese sample. We identified a single polymorphism in the promoter region (rs6002408). In addition the dbSNP database (<http://www.ncbi.nlm.nih.gov/SNP>) and the genome browser of University of California Santa Cruz (<http://www.genome.ucsc.edu/>) were used for selection of 11 SNPs in *EP300*, *PIPPIN*, *NHP2L1*, and *SERHL* (table 1). A total of 12 SNPs were analyzed. All SNPs were found to be in Hardy-Weinberg equilibrium in all samples. Rs5758252 and rs1802521 were monomorphic and excluded from statistical analyses. Method problems, and uncertain genotyping caused rs2294976, rs60002408, rs132806, and rs5758405 to be excluded from the statistical analyses of the Scottish sample, and rs60002408 to be excluded from statistical analysis in the replication samples from Denmark and the UK.

20

Association analysis.

In the Scottish sample three SNPs showed a significant single marker association; rs20551 (*EP300*) was associated with BPD (p-value=0.0132), rs926333 (*SERHL*) was associated with SZ (p-value=0.0131), and rs1006407 (*PIPPIN*) was associated with SZ and BPD combined (p-value=0.049) (table 2). In the overall haplotype analysis the strongest association was observed when comparing controls to cases combined. 2-, 3-, 4-, and 5-marker haplotypes was associated with both SZ and BPD with the most significant signal centered on rs8779 (*NHP2L1*). Especially the specific 2-marker haplotype rs1006407-T / rs8779-A, was underrepresented in cases (0.7%) compared to controls (5%) yielding a p-value of 4.17×10^{-6} (table 3).

30

In the UK sample rs8779 showed significant association with BPD, SZ, and cases combined (p-values of 0.0044, 0.0149, and 0.0024 respectively). Rs926333 was found to be associated with BPD and cases combined (table 4). In the overall haplotype analysis the association was almost entirely found in BPD centered on the 2-marker haplotype rs1006407-rs8779 and 3-, 4-, and 5-marker centromeric

35

extensions of this haplotype (table 4). Analysis of the individual haplotypes revealed that rs1006407-T / rs8779-A (the same haplotype as original was identified in the Scottish sample), had a frequency of 2.8% in the controls, and a frequency of 13.6% in cases suffering from BPD (table 5) (which is skewed in the opposite direction compared to the Scottish sample) yielding a haplotype specific p-value of 1.36×10^{-9} . The haplotype rs1006407-T/rs8779-G in contrary was underrepresented in cases suffering from BPD (68.7%) compared to controls (80.2%) producing a haplotype specific p-value of 1.93×10^{-5} .

Additionally two modest overall significant results were found involving the SNPs in *NHP2L1* and in *SERHL* (table 4), with several individual haplotypes generating the association.

In the Danish sample rs881542 (*SERHL*) showed single marker association with SZ, BPD and cases combined whereas rs5758405 (*NHP2L1*) and rs1006407 was found to be associated with BPD and cases combined. Rs8779 showed a significant association with BPD (table 6).

The haplotype analysis produced the most significant associations when comparing BPD to controls and less though still significant association when comparing SZ to controls. The strongest signal was centered on 2-, to 5-marker haplotypes associated with BPD, SZ, as well as theses disorders combined, involving primarily the three SNPs rs5758405 (*NHP2L1*), rs881542 and rs926333 (both *SERHL*). The specific haplotypes causing this association were rs5758405-A /rs881542-C/rs926333-G (and derivates of this haplotype), which was overrepresented in cases (72%) compared to controls (56%), and rs881542-G/rs926333-G (and derivates of this haplotype), which was overrepresented in controls (23%) compared to cases (13%), yielding haplotype specific p-values of 1.49×10^{-4} and 1.2×10^{-3} respectively.

As in the samples from Scotland and UK the overall 2-marker haplotype rs1006407/rs8779 showed association with BPD and SZ. The specific haplotype rs1006407-T/rs8779-A was in this sample significantly overrepresented in SZ (12.8%) compared to controls (2.12%)(p-value= 7.82×10^{-4}) whereas the haplotype was not found in BPD at all (table 7).

Combining the original case-control sample from Scotland with the replication samples from Denmark and UK produced overall and haplotype specific associations similar to the signals found in the samples separately. Furthermore a number of haplotypes were found to have a consistently skewed frequency across all three samples (table 8).

In silico analysis

Using MatInspector (www.genomatix.de) the *SERHL* promoter SNP rs881542 (table 1) was analyzed for effects on potential binding sites for transcription factors. The program suggested that the rare G-allele introduced new binding sites for the transcription factors Collagen krox protein (zink finger protein zfp67) shown to be involved in regulation of type 1-, and 2 collagen gene transcription^{59,60} and Myc associated zing finger protein (MAZ), a transcription factor closely related to SP1 and involved in regulation of numerous genes.⁶¹

The non-synonymous *EP300* SNP rs20551 located in exon 15 and the *SERHL* SNP rs926333 located in exon 2 were analyzed for effects on exon splicing enhancers (ESE), exon splice silencers (ESS) and donor and acceptor sites. FAS-ESS identified a site for a potential exon splice silencer introduced by the G-allele of rs20551 but analysis of the splice sites revealed strong donor and acceptor site suggesting a very limited potential effect on splicing of this site. The ExonScan analysis of rs20551 confirmed this by finding no other potential splicing variation of exon 15 and hardly any difference in the summed splicing score introduced by the G-allele of rs20551.

According to NNSPLICE the *NHP2L1* intron SNP rs5758405 had no effect on potential donor or acceptor sites in intron 1. Likewise the *EP300* intron SNPs rs2294976 (*EP300*) and rs2076578 (*EP300*) had no effect on splicing.

MiRBase Targets Pre-release Version 1.0 identified no changes in miRNA binding sites in the 3'UTR sequences of *PIPPIN* or *NHP2L1*.

Linkage Disequilibrium

In general the same LD pattern was observed in cases and in controls in the three different samples (table 9). A high degree of LD was observed between SNPs located in *EP300*, *PIPPIN*, and *NHP2L1*. A limited level of LD was observed between SNPs located in *SERHL* and SNPs in *EP300*, *PIPPIN* and *NHP2L1* indicating that *SERHL* could be located in another haplotype block. The results was confirmed by data retrieved from The International HapMap Project (<http://www.hapmap.org/>) in which *EP300*, *PIPPIN* and *NHP2L1* in the CEU population were located in a 780 kb loosely defined haplotype block, while *SERHL* was located in a clearly distinctive more telomeric block.

References:

1. 22, S.C.L.G.f.C. A transmission disequilibrium and linkage analysis of D22S278 marker alleles in 574 families: further support for a susceptibility locus for schizophrenia at 22q12. Schizophrenia Collaborative Linkage Group for Chromosome 22. *Schizophr Res* **32**, 115-21 (1998).
2. Aamantidis A, Thomas E, Foidart A, Tyhon A, Coumans B, Minet A, Tirelli E, Seutin V, Grisar T, Lakaye B. 2005. Disrupting the melanin-concentrating hormone receptor 1 in mice leads to cognitive deficits and alterations of NMDA receptor function. *Eur J Neurosci* **21**(10):2837-44.
3. Abecasis, G.R. & Cookson, W.O. GOLD--graphical overview of linkage disequilibrium. *Bioinformatics* **16**, 182-3 (2000).
4. Als TD, Dahl HA, Flint TJ, Wang AG, Vang M, Mors O, Kruse TA and Ewald H (2004) Possible evidence for a common risk locus for bipolar affective disorder and schizophrenia on chromosome 4p16 in patients from the Faroe Islands. *Mol Psychiatry*, **9**, 93-8.
5. An, W.; Kim, J.; Roeder, R. G. Ordered cooperative functions of PRMT1, p300, and CARM1 in transcriptional activation by p53. *Cell* **117**: 735-748, 2004.
6. Arany, Z.; Newsome, D.; Oldread, E.; Livingston, D. M.; Eckner, R. A family of transcriptional adaptor proteins targeted by the E1A oncoprotein. *Nature* **374**: 81-84, 1995.
7. Arany, Z.; Sellers, W. R.; Livingston, D. M.; Eckner, R. E1A-associated p300 and CREB-associated CBP belong to a conserved family of coactivators. (Letter) *Cell* **77**: 799-800, 1994.
8. Asherson P, Mant R, Williams N, Cardno A, Jones L, Murphy K, Collier DA, Nanko S, Craddock N, Morris S et al. (1998) A study of chromosome 4p markers and dopamine D5 receptor gene in schizophrenia and bipolar disorder. *Mol Psychiatry*, **3**, 310-20.
9. Association, A.P. *Diagnostic and statistical manual of mental disorders.*, (American Psychiatric Association, Washington, 1994).
- 10 Auga J. 2004. The role of Zic genes in neural development. *Mol Cell Neurosci* **26**(2):205-21.
- 11 Badner J. A., and E. S. Gershon, 2002, Meta-analysis of whole-genome linkage scans of bipolar disorder and schizophrenia: *Mol Psychiatry*, v. 7, p. 405-11.
- 12 Berrettini W (2003) Evidence for shared susceptibility in bipolar disorder and schizophrenia. *Am J Med Genet C Semin Med Genet*, **123**, 59-64.
- 13 Berrettini W. H., 2000, Are schizophrenic and bipolar disorders related? A review of family and molecular studies: *Biol Psychiatry*, v. 48, p. 531-8.

- 14 Bjarkam CR, Pedersen M and Sorensen JC (2001) New strategies for embedding, orientation and sectioning of small brain specimens enable direct correlation to MR-images, brain atlases, or use of unbiased stereology. *J Neurosci Methods*, 108, 153-9.
- 15 Blackwood DH, He L, Morris SW, McLean A, Whitton C, Thomson M, Walker MT, Woodburn K, Sharp CM, Wright AF et al. (1996) A locus for bipolar affective disorder on chromosome 4p. *Nat Genet*, 12, 427-30.
- 16 Blouin, J.L. et al. Schizophrenia susceptibility loci on chromosomes 13q32 and 8p21. *Nat Genet* **20**, 70-3 (1998).
- 17 Borglum AD, Hampson M, Kjeldsen TE, Muir W, Murray V, Ewald H, Mors O, Blackwood D and Kruse TA (2001) Dopa decarboxylase genotypes may influence age at onset of schizophrenia. *Mol Psychiatry*, 6, 712-7.
- 18 Borglum AD, Kirov G, Craddock N, Mors O, Muir W, Murray V, McKee I, Collier DA, Ewald H, Owen MJ et al. (2003) Possible parent-of-origin effect of Dopa decarboxylase in susceptibility to bipolar affective disorder. *Am J Med Genet B Neuropsychiatr Genet*, 117, 18-22.
- 19 Borowsky, B.; Durkin, M. M.; Ogozalek, K.; Marzabadi, M. R.; DeLeon, J.; Heurich, R.; Lichtblau, H.; Shaposhnik, Z.; Daniewska, I.; Blackburn, T. P.; Branchek, T. A.; Gerald, C.; Vaysse, P. J.; Forray, C. Antidepressant, anxiolytic and anorectic effects of a melanin-concentrating hormone-1 receptor antagonist. *Nature Med.* 8: 825-830, 2002.
- 20 Bttencourt JC, Frigo L, Rissman RA, Casatti CA, Nahon JL, Bauer JA. 1998. The distribution of melanin-concentrating hormone in the monkey brain (*Cebus apella*). *Brain Res* 804(1):140-3.
- 21 Bttencourt JC, Presse F, Arias C, Peto C, Vaughan J, Nahon JL, Vale W, Sawchenko PE. 1992. The melanin-concentrating hormone system of the rat brain: an immuno- and hybridization histochemical characterization. *J Comp Neurol* 319(2):218-45.
- 22 Cartegni L, Wang J, Zhu Z, Zhang MQ and Krainer AR (2003) ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res*, 31, 3568-71.
- 23 Castella P, Wagner JA and Caudy M (1999) Regulation of hippocampal neuronal differentiation by the basic helix-loop-helix transcription factors HES-1 and MASH-1. *J Neurosci Res*, 56, 229-40.
- 24 Castiglia D, Scaturro M, Nastasi T, Cestelli A, Di Liegro I. PIPPin, a putative RNA-binding protein specifically expressed in the rat brain. *Biochem Biophys Res Commun*. 1996 Jan 5;218(1):390-4.
- 25 Chaki S, Funakoshi T, Hirota-Okuno S, Nishiguchi M, Shimazaki T, Iijima M, Grottick AJ, Kanuma K, Omodera K, Sekiguchi Y and others. 2005. Anxiolytic- and antidepressant-like profile of ATC0065 and ATC0175: nonpeptidic and orally active melanin-concentrating hormone receptor 1 antagonists. *J Pharmacol Exp Ther* 313(2):831-9.

- 26 Chambers, J.; Ames, R. S.; Bergsma, D.; Muir, A.; Fitzgerald, L. R.; Hervieu, G.; Dytko, G. M.; Foley, J. J.; Martin, J.; Liu, W.-S.; Park, J.; Ellis, C.; Ganguly, S.; Konchar, S.; Cluderay, J.; Leslie, R.; Wilson, S.; Sarau, H. M. Melanin-concentrating hormone is the cognate-ligand for the orphan G-protein-coupled receptor SLC-1. *Nature* 400: 261-265, 1999.
- 27 Chang, M.S. et al. HRad17 colocalizes with NHP2L1 in the nucleolus and redistributes after UV irradiation. *J Biol Chem* 274, 36544-9 (1999).
- 28 Chen S, Smith DF. 1998. Hop as an adaptor in the heat shock protein 70 (Hsp70) and hsp90 chaperone machinery. *J Biol Chem* 273(52):35194-200.
- 29 Cojocaru, V., Nottrott, S., Klement, R. & Jovin, T.M. The snRNP 15.5K protein folds its cognate K-turn RNA: a combined theoretical and biochemical study. *Rna* 11, 197-209 (2005).
- 30 Coon H, Jensen S, Holik J, Hoff M, Myles-Worsley M, Reimherr F, Wender P, Waldo M, Freedman R, Leppert M et al. (1994) Genomic scan for genes predisposing to schizophrenia. *Am J Med Genet*, 54, 59-71.
- 31 Coon, H. et al. Analysis of chromosome 22 markers in nine schizophrenia pedigrees. *Am J Med Genet* 54, 72-9 (1994).
- 32 Corfas G, Roy K and Buxbaum JD (2004) Neuregulin 1-erbB signaling and the molecular/cellular basis of schizophrenia. *Nat Neurosci*, 7, 575-80.
- 33 Craddock N, O'Donovan MC and Owen MJ (2005) The genetics of schizophrenia and bipolar disorder: dissecting psychosis. *J Med Genet*, 42, 193-204.
- 34 Craddock N, O'Donovan MC and Owen MJ (2006) Genes for schizophrenia and bipolar disorder? Implications for psychiatric nosology. *Schizophr Bull*, 32, 9-16.
- 35 DeLisi, L.E. et al. A genome-wide scan for linkage to chromosomal regions in 382 sibling pairs with schizophrenia or schizoaffective disorder. *Am J Psychiatry* 159, 803-12 (2002).
- 36 Detera-Wadleigh, S.D. et al. A high-density genome scan detects evidence for a bipolar-disorder susceptibility locus on 13q32 and other potential loci on 1q32 and 18p11.2. *Proc Natl Acad Sci U S A* 96, 5604-9 (1999).
- 37 Devaney JM, Donarum EA, Brown KM, Meyer J, Stober G, Lesch KP, Nestadt G, Stephan DA and Pulver AE (2002) No missense mutation of WKL1 in a subgroup of probands with schizophrenia. *Mol Psychiatry*, 7, 419-23.
- 38 Devlin B, Roeder K. 1999. Genomic control for association studies. *Biometrics* 55(4):997-1004.

— SUBSTITUTE SHEET

- 39 Eckner, R.; Ewen, M. E.; Newsome, D.; Gerdes, M.; DeCaprio, J. A.; Lawrence, J. B.; Livingston, D. M. Molecular cloning and functional analysis of the adenovirus E1A-associated 300-kD protein (p300) reveals a protein with properties of a transcriptional adaptor. *Genes Dev.* 15: 869-884, 1994.
- 40 Edenberg, H.J. et al. Initial genomic scan of the NIMH genetics initiative bipolar pedigrees: chromosomes 3, 5, 15, 16, 17, and 22. *Am J Med Genet* 74, 238-46 (1997).
- 41 Enright, A.J. et al. MicroRNA targets in *Drosophila*. *Genome Biol* 5, R1 (2003).
- 42 Etchegaray, J.-P.; Lee, C.; Wade, P. A.; Reppert, S. M. Rhythmic histone acetylation underlies transcription in the mammalian circadian clock. *Nature* 421: 177-182, 2003.
- 43 Excoffier L and Slatkin M (1995) Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol Biol Evol*, 12, 921-7.
- 44 Fairbrother WG, Yeh RF, Sharp PA and Burge CB (2002) Predictive identification of exonic splicing enhancers in human genes. *Science*, 297, 1007-13.
- 45 Galera, P., Musso, M., Ducy, P. & Karsenty, G. c-Krox, a transcriptional regulator of type I collagen gene expression, is preferentially expressed in skin. *Proc Natl Acad Sci U S A* 91, 9372-6 (1994).
- 46 Gayther, S. A.; Batley, S. J.; Linger, L.; Bannister, A.; Thorpe, K.; Chin, S.-F.; Daigo, Y.; Russell, P.; Wilson, A.; Sowter, H. M.; Delhanty, J. D. A.; Ponder, B. A. J.; Kouzarides, T.; Caldas, C. Mutations truncating the EP300 acetylase in human cancers. *Nature Genet.* 24: 300-303, 2000.
- 47 Ghayor, C. et al. Regulation of human COL2A1 gene expression in chondrocytes. Identification of C-Krox-responsive elements and modulation by phenotype alteration. *J Biol Chem* 275, 27421-38 (2000).
- 48 Gill M., H. Vallada, D. Collier, P. Sham, P. Holmans, R. Murray, P. McGuffin, S. Nanko, M. Owen, S. Antonarakis, D. Housman, H. Kazazian, G. Nestadt, A. E. Pulver, R. E. Straub, C. J. MacLean, D. Walsh, K. S. Kendler, L. DeLisi, M. Polymeropoulos, H. Coon, W. Byerley, R. Lofthouse, E. Gershon, C. M. Read, and et al., 1996, A combined analysis of D22S278 marker alleles in affected sib-pairs: support for a susceptibility locus for schizophrenia at chromosome 22q12. Schizophrenia Collaborative Linkage Group (Chromosome 22): *Am J Med Genet*, v. 67, p. 40-5.
- 49 Grossman, S. R.; Deato, M. E.; Brignone, C.; Chan, H. M.; Kung, A. L.; Tagami, H.; Nakatani, Y.; Livingston, D. M. Polyubiquitination of p53 by a ubiquitin ligase activity of p300. *Science* 300: 342-344, 2003.

- 50 Hamshere, M.L. et al. Genomewide linkage scan in schizoaffective disorder: significant evidence for linkage at 1q42 close to DISC1, and suggestive evidence at 22q11 and 19p13. *Arch Gen Psychiatry* **62**, 1081-8 (2005).
- 51 Hardy OJ, Vekemans X. 2002. SPAGEDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* **2**:618-620
- 52 Harrison PJ and Weinberger DR (2005) Schizophrenia genes, gene expression, and neuropathology: on the matter of their convergence. *Mol Psychiatry*, **10**, 40-68.
- 53 Hasan, S.; Hassa, P. O.; Imhof, R.; Hottiger, M. O. Transcription coactivator p300 binds PCNA and may have a role in DNA repair synthesis. *Nature* **410**: 387-391, 2001.
- 54 Hasan, S.; Stucki, M.; Hassa, P. O.; Imhof, R.; Gehrig, P.; Hunziker, P.; Hubscher, U.; Hottiger, M. O. Regulation of human flap endonuclease-1 activity by acetylation through the transcriptional coactivator p300. *Molec. Cell* **7**: 1221-1231, 2001.
- 55 Ida, K.; Kitabayashi, I.; Taki, T.; Taniwaki, M.; Noro, K.; Yamamoto, M.; Ohki, M.; Hayashi, Y. Adenoviral E1A-associated protein p300 is involved in acute myeloid leukemia with t(11;22)(q23;q13). *Blood* **90**: 4699-4704, 1997.
- 56 Ishibashi M (2004) Molecular mechanisms for morphogenesis of the central nervous system in mammals. *Anat Sci Int*, **79**, 226-34.
- 57 Jensen TG, Andresen BS, Bross P, Jensen UB, Holme E, Kolvraa S, Gregersen N and Bolund L (1992) Expression of wild-type and mutant medium-chain acyl-CoA dehydrogenase (MCAD) cDNA in eucaryotic cells. *Biochim Biophys Acta*, **1180**, 65-72.
- 58 Jorgensen TH, Borglum AD, Mors O, Wang AG, Pinaud M, Flint TJ, Dahl HA, Vang M, Kruse TA and Ewald H (2002) Search for common haplotypes on chromosome 22q in patients with schizophrenia or bipolar disorder from the Faroe Islands. *Am J Med Genet*, **114**, 245-52.
- 59 Jorgensen TH, Degn B, Wang AG, Vang M, Gurling H, Kalsi G, McQuillin A, Kruse TA, Mors O, Ewald H. 2002b. Linkage disequilibrium and demographic history of the isolated population of the Faroe Islands. *Eur J Hum Genet* **10**(6):381-7.
- 60 Kaganovich M, Peretz A, Ritsner M, Bening Abu-Shach U, Attali B and Navon R (2004) Is the WKL1 gene associated with schizophrenia? *Am J Med Genet B Neuropsychiatr Genet*, **125**, 31-7.
- 61 Kasper, L. H.; Boussouar, F.; Ney, P. A.; Jackson, C. W.; Rehg, J.; van Deursen, J. M.; Brindle, P. K. A transcription-factor-binding surface of coactivator p300 is required for haematopoiesis. *Nature* **419**: 738-743, 2002.

- 62 Kelsoe J. R., M. A. Spence, E. Loetscher, M. Foguet, A. D. Sadovnick, R. A. Remick, P. Flodman, J. Khristich, Z. Mroczkowski-Parker, J. L. Brown, D. Masser, S. Ungerleider, M. H. Rapaport, W. L. Wishart, and H. Luebbert, 2001, A genome survey indicates a possible susceptibility locus for bipolar disorder on chromosome 22: *Proc Natl Acad Sci U S A*, v. 98, p. 585-90.
- 63 Kmoch S, Hartmannova H, Stiburkova B, Krijt J, Zikanova M, Sebesta I. 2000. Human adenylosuccinate lyase (ADSL), cloning and characterization of full-length cDNA and its isoform, gene structure and molecular basis for ADSL deficiency in six patients. *Hum Mol Genet* 9(10):1501-13.
- 64 Kolakowski, L. F., Jr.; Jung, B. P.; Nguyen, T.; Johnson, M. P.; Lynch, K. R.; Cheng, R.; Heng, H. H. Q.; George, S. R.; O'Dowd, B. F. Characterization of a human gene related to genes encoding somatostatin receptors. *FEBS Lett.* 398: 253-258, 1996.
- 65 Kolodrubetz, D.; Burgum, A. Sequence and genetic analysis of NHP2: a moderately abundant high mobility group-like nuclear protein with an essential function in *Saccharomyces cerevisiae*. *Yeast* 7: 79-90, 1991.
- 66 Lachman, H.M. et al. Linkage studies suggest a possible locus for bipolar disorder near the velo-cardio-facial syndrome region on chromosome 22. *Am J Med Genet* 74, 121-8 (1997).
- 67 Ladurner AG, Inouye C, Jain R and Tjian R (2003) Bromodomains mediate an acetyl-histone encoded antisilencing function at heterochromatin boundaries. *Mol Cell*, 11, 365-76.
- 68 Lakaye, B.; Minet, A.; Zorzi, W.; Grisar, T. Cloning of the rat brain cDNA encoding for the SLC-1 G protein-coupled receptor reveals the presence of an intron in the gene. *Biochim. Biophys. Acta* 1401: 216-220, 1998.
- 69 Lewis CM, Levinson DF, Wise LH, DeLisi LE, Straub RE, Hovatta I, Williams NM, Schwab SG, Pulver AE, Faraone SV et al. (2003) Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: Schizophrenia. *Am J Hum Genet*, 73, 34-48.
- 70 Li T, Ma X, Sham PC, Sun X, Hu X, Wang Q, Meng H, Deng W, Liu X, Murray RM, Collier DA. Evidence for association between novel polymorphisms in the PRODH gene and schizophrenia in a Chinese population. *Am J Med Genet B Neuropsychiatr Genet.* 129B:13-5, 2004
- 71 Liang, S.G. et al. A linkage disequilibrium study of bipolar disorder and microsatellite markers on 22q13. *Psychiatr Genet* 12, 231-5 (2002).
- 72 Lin, C. H.; Hare, B. J.; Wagner, G.; Harrison, S. C.; Maniatis, T.; Fraenkel, E. : A small domain of CBP/p300 binds diverse proteins: solution structure and functional studies. *Molec. Cell* 8: 581-590, 2001.
- 73 Livak KJ and Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, 25, 402-8.

- 74 Lukasik SM, Cierpicki T, Borloz M, Grembecka J, Everett A and Bushweller JH (2006) High resolution structure of the HDGF PWWP domain: a potential DNA binding domain. *Protein Sci*, 15, 314-23.
- 75 Manning, E.T., Ikehara, T., Ito, T., Kadonaga, J.T. & Kraus, W.L. p300 forms a stable, template-committed complex with chromatin: role for the bromodomain. *Mol Cell Biol* **21**, 3876-87 (2001).
- 76 Marie S, Cuppens H, Heuterspreute M, Jaspers M, Tola EZ, Gu XX, Legius E, Vincent MF, Jaeken J, Cassiman JJ and others. 1999. Mutation analysis in adenylosuccinate lyase deficiency: eight novel mutations in the re-evaluated full ADSL coding sequence. *Hum Mutat* 13(3):197-202.
- 77 Marsh, D. J.; Weingarth, D. T.; Novi, D. E.; Chen, H. Y.; Trumbauer, M. E.; Chen, A. S.; Guan, X.-M.; Jiang, M. M.; Feng, Y.; Camacho, R. E.; Shen, Z.; Frazier, E. G.; et al. Melanin-concentrating hormone 1 receptor-deficient mice are lean, hyperactive, and hyperphagic and have altered metabolism. *Proc. Nat. Acad. Sci.* 99: 3240-3245, 2002.
- 78 McCullagh, P.; Chaplin, T.; Meerabux, J.; Grenzeliass, D.; Lillington, D.; Poulson, R.; Gregorini, A.; Saha, V.; Young, B. D. The cloning, mapping and expression of a novel gene, BRL, related to the AF10 leukaemia gene. *Oncogene* 18: 7442-7452, 1999.
- 79 McGuffin P., M. J. Owen, and A. E. Farmer, 1995, Genetic basis of schizophrenia: *Lancet*, v. 346, p. 678-82.
- 80 Meyer J, Huberth A, Ortega G, Syagailo YV, Jatzke S, Mossner R, Strom TM, Ulzheimer-Teuber I, Stober G, Schmitt A et al. (2001) A missense mutation in a novel gene encoding a putative cation channel is associated with catatonic schizophrenia in a large pedigree. *Mol Psychiatry*, 6, 302-6.
- 81 Moises, H.W. et al. Potential linkage disequilibrium between schizophrenia and locus D22S278 on the long arm of chromosome 22. *Am J Med Genet* **60**, 465-7 (1995).
- 82 Mowry BJ, Holmans PA, Pulver AE, Gejman PV, Riley B, Williams NM, Laurent C, Schwab SG, Wildenauer DB, Bauche S et al. (2004) Multicenter linkage study of schizophrenia loci on chromosome 22q. *Mol Psychiatry*, 9, 784-95.
- 83 Muraoka, M.; Konishi, M.; Kikuchi-Yanoshita, R.; Tanaka, K.; Shitara, N.; Chong, J.-M.; Iwama, T.; Miyaki, M. p300 gene alterations in colorectal and gastric carcinomas. *Oncogene* 12: 1565-1569, 1996.
- 84 Myles-Worsley, M. et al. Linkage of a composite inhibitory phenotype to a chromosome 22q locus in eight Utah families. *Am J Med Genet* **88**, 544-50 (1999).
- 85 Nakashima, K.; Yanagisawa, M.; Arakawa, H.; Kimura, N.; Hisatsune, T.; Kawabata, M.; Miyazono, K.; Taga, T. Synergistic signaling in fetal brain by STAT3-Smad1 complex bridged by p300. *Science* 284: 479-482, 1999.

- 86 Nastasi T, Muzi P, Beccari S, Bellafigliore M, Dolo V, Bologna M, Cestelli A, Di Liegro I. Specific neurons of brain cortex and cerebellum are PIPPin positive. *Neuroreport*. 2000 Jul 14;11(10):2233-6.
- 87 Nastasi T, Scaturro M, Bellafigliore M, Raimondi L, Beccari S, Cestelli A, di Liegro I. PIPPin is a brain-specific protein that contains a cold-shock domain and binds specifically to H1 degrees and H3.3 mRNAs. *J Biol Chem*. 1999 Aug 20;274(34):24087-93
- 88 Pissios P, Trombly DJ, Tzamelis I, Maratos-Flier E. 2003. Melanin-concentrating hormone receptor 1 activates extracellular signal-regulated kinase and synergizes with G(s)-coupled pathways. *Endocrinology* 144(8):3514-23.
- 89 Plomann, M.; Lange, R.; Vopper, G.; Cremer, H.; Heinlein, U. A. O.; Scheff, S.; Baldwin, S. A.; Leitges, M.; Cramer, M.; Paulsson, M.; Barthels, D. PACSIN, a brain protein that is upregulated upon differentiation into neuronal cells. *Europ. J. Biochem*. 256: 201-211, 1998.
- 90 Polymeropoulos MH, Coon H, Byerley W, Gershon ES, Goldin L, Crow TJ, Rubenstein J, Hoff M, Holik J, Smith AM and others. 1994. Search for a schizophrenia susceptibility locus on human chromosome 22. *Am J Med Genet* 54(2):93-9.
- 91 Potash J. B., and J. R. DePaulo, Jr., 2000, Searching high and low: a review of the genetics of bipolar disorder: *Bipolar Disord*, v. 2, p. 8-26.
- 92 Potash J. B., P. P. Zandi, V. L. Willour, T. H. Lan, Y. Huo, D. Avramopoulos, Y. Y. Shugart, D. F. MacKinnon, S. G. Simpson, F. J. McMahon, J. R. DePaulo, Jr., and M. G. McInnis, 2003b, Suggestive linkage to chromosomal regions 13q31 and 22q12 in families with psychotic bipolar disorder: *Am J Psychiatry*, v. 160, p. 680-6.
- 93 Potash J. B., Y. F. Chiu, D. F. MacKinnon, E. B. Miller, S. G. Simpson, F. J. McMahon, M. G. McInnis, and J. R. DePaulo, Jr., 2003a, Familial aggregation of psychotic symptoms in a replication set of 69 bipolar disorder pedigrees: *Am J Med Genet*, v. 116B, p. 90-7.
- 94 Potash JB, Willour VL, Chiu YF, Simpson SG, MacKinnon DF, Pearlson GD, DePaulo JR, Jr., McInnis MG. 2001. The familial aggregation of psychotic symptoms in bipolar disorder pedigrees. *Am J Psychiatry* 158(8):1258-64.
- 95 Prapapanich V, Chen S, Nair SC, Rimerman RA, Smith DF. 1996. Molecular cloning of human p48, a transient component of progesterone receptor complexes and an Hsp70-binding protein. *Mol Endocrinol* 10(4):420-31.
- 96 Pulver A. E., M. Karayiorgou, P. S. Wolyniec, V. K. Lasseter, L. Kasch, G. Nestadt, S. Antonarakis, D. Housman, H. H. Kazazian, D. Meyers, and et al., 1994, Sequential strategy to identify a susceptibility gene for schizophrenia: report of potential linkage on chromosome 22q12-q13.1: Part 1: *Am J Med Genet*, v. 54, p. 36-43.

97. Puri, V. et al. Failure to Confirm Allelic Association Between Markers at the CAPON Gene Locus and Schizophrenia in a British Sample. *Biol Psychiatry* (2005).
98. Qu D, Ludwig DS, Gammeltoft S, Piper M, Pelleymounter MA, Cullen MJ, Mathes WF, Przypek R, Kanarek R, Maratos-Flier E. 1996. A role for melanin-concentrating hormone in the central regulation of feeding behaviour. *Nature* 380(6571):243-7.
99. Quandt K, Frech K, Karas H, Wingender E and Werner T (1995) MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res*, 23, 4878-84.
100. Queller DC, Goodnight KF. 1989. Estimating Relatedness Using Genetic-Markers. *Evolution* 43(2):258-275.
101. Ragvin A, Valvatne H, Erdal S, Arskog V, Tufteland KR, Breen K, AM OY, Eberharter A, Gibson TJ, Becker PB et al. (2004) Nucleosome binding by the bromodomain and PHD finger of the transcriptional cofactor p300. *J Mol Biol*, 337, 773-88.
102. Rapoport JL, Addington AM, Frangou S and Psych MR (2005) The neurodevelopmental model of schizophrenia: update 2005. *Mol Psychiatry*, 10, 434-49.
103. Reese MG, Eeckman FH, Kulp D and Haussler D (1997) Improved splice site detection in Genie. *J Comput Biol*, 4, 311-23.
104. Ritter, B.; Modregger, J.; Paulsson, M.; Plomann, M. PACSIN 2, a novel member of the PACSIN family of cytoplasmic adapter proteins. *FEBS Lett.* 454: 356-362, 1999.
105. Roelfsema, J.H. et al. Genetic heterogeneity in Rubinstein-Taybi syndrome: mutations in both the CBP and EP300 genes cause disease. *Am J Hum Genet* 76, 572-80 (2005).
106. Rousset F. 2000. Genetic differentiation between individuals. *Journal of Evolutionary Biology* 13(1):58-62.
107. Rubie C, Lichtner P, Gartner J, Siekiera M, Uziel G, Kohlmann B, Kohlschutter A, Meitinger T, Stober G and Bettecken T (2003) Sequence diversity of KIAA0027/MLC1: are megalencephalic leukoencephalopathy and schizophrenia allelic disorders? *Hum Mutat*, 21, 45-52.
108. Sadusky, T. J.; Kemp, T. J.; Simon, M.; Carey, N.; Coulton, G. R. : Identification of Serhl, a new member of the serine hydrolase family induced by passive stretch of skeletal muscle in vivo. *Genomics* 73: 38-49, 2001. Note: Erratum: *Genomics* 74: 251 only, 2001.

109. Sadusky, T.J., Kemp, T.J., Simon, M., Carey, N. & Coulton, G.R. Identification of Serhl, a new member of the serine hydrolase family induced by passive stretch of skeletal muscle in vivo. *Genomics* **73**, 38-49 (2001).
110. Saito, H.; Fujiwara, T.; Shin, S.; Okui, K.; Nakamura, Y. Cloning and mapping of a human novel cDNA (NHP2L1) that encodes a protein highly homologous to yeast nuclear protein NHP2. *Cytogenet. Cell Genet.* **72**: 191-193, 1996.
111. Saito, Y.; Nothacker, H.-P.; Wang, Z.; Lin, S. H. S.; Leslie, F.; Civelli, O. Molecular characterization of the melanin-concentrating-hormone receptor. *Nature* **400**: 265-269, 1999.
112. Saleem Q, Dash D, Gandhi C, Kishore A, Benegal V, Sherrin T, Mukherjee O, Jain S, Brahmachari SK. 2001. Association of CAG repeat loci on chromosome 22 with schizophrenia and bipolar disorder. *Mol Psychiatry* **6**(6):694-700.
113. Schmitz G, Heimerl S and Langmann T (2004) Zinc finger protein ZNF202 structure and function in transcriptional control of HDL metabolism. *Curr Opin Lipidol*, **15**, 199-208.
114. Schwab SG and Wildenauer DB (1999) Chromosome 22 workshop report. *Am J Med Genet*, **88**, 276-8.
115. Segurado, R. et al. Genome scan meta-analysis of schizophrenia and bipolar disorder, part III: Bipolar disorder. *Am J Hum Genet* **73**, 49-62 (2003).
116. Sham PC, Curtis D. 1995. Monte Carlo tests for associations between disease and alleles at highly polymorphic loci. *Ann Hum Genet* **59** (Pt 1):97-105.
117. Sinibaldi L, De Luca A, Bellacchio E, Conti E, Pasini A, Paloscia C, Spalletta G, Caltagirone C, Pizzuti A, Dallapiccola B. Mutations of the Nogo-66 receptor (RTN4R) gene in schizophrenia. *Hum Mutat.* **24**:534-5, 2004.
118. Skibinska M, Hauser J, Czerski PM, Leszczynska-Rodziewicz A, Kosmowska M, Kapelski P, Slopian A, Zakrzewska M, Rybakowski JK. Association analysis of brain-derived neurotrophic factor (BDNF) gene Val66Met polymorphism in schizophrenia and bipolar affective disorder. *World J Biol Psychiatry* **5**:215-20, 2004
119. Smith DG, Tzavara ET, Shaw J, Luecke S, Wade M, Davis R, Salhoff C, Nomikos GG, Gehlert DR. 2005. Mesolimbic dopamine super-sensitivity in melanin-concentrating hormone-1 receptor-deficient mice. *J Neurosci* **25**(4):914-22.

120. Song, J. et al. Transcriptional regulation by zinc-finger proteins Sp1 and MAZ involves interactions with the same cis-elements. *Int J Mol Med* **11**, 547-53 (2003).
121. Spitzer R & J, E. *The Schedule for Affective Disorders and Schizophrenia, Lifetime Version, 3rd edition.*, (The Schedule for Affective Disorders and Schizophrenia, Lifetime Version, 3rd edition, New York, 1977).
122. Spitzer R, J, E. & E, R. *Research Diagnostic Criteria for a selected group of functional disorders, 3rd edition*, (New York, New York State Psychiatric Institute, New York, 1978).
123. Stober G, Saar K, Ruschendorf F, Meyer J, Nurnberg G, Jatzke S, Franzek E, Reis A, Lesch KP, Wienker TF et al. (2000) Splitting schizophrenia: periodic catatonia-susceptibility locus on chromosome 15q15. *Am J Hum Genet*, 67, 1201-7.
124. Strakowski SM, Delbello MP and Adler CM (2005) The functional neuroanatomy of bipolar disorder: a review of neuroimaging findings. *Mol Psychiatry*, 10, 105-16.
125. Sumoy, L.; Pluvinet, R.; Andreu, N.; Estivill, X.; Escarceller, M. PACSIN 3 is a novel SH3 domain cytoplasmic adapter protein of the pacsin-syndapin-FAP52 gene family. *Gene* 262: 199-205, 2001.
126. Takahashi S, Faraone SV, Lasky-Su J and Tsuang MT (2005) Genome-wide scan of homogeneous subtypes of NIMH genetics initiative schizophrenia families. *Psychiatry Res*, 133, 111-22.
127. Takahashi, K.; Totsune, K.; Murakami, O.; Sone, M.; Satoh, F.; Kitamuro, T.; Noshiro, T.; Hayashi, Y.; Sasano, H.; Shibahara, S. Expression of melanin-concentrating hormone receptor messenger ribonucleic acid in tumor tissues of pheochromocytoma, ganglioneuroblastoma, and neuroblastoma. *J. Clin. Endocr. Metab.* 86: 369-374, 2001.
128. Tini, M.; Benecke, A.; Um, S.-J.; Torchia, J.; Evans, R. M.; Chambon, P. Association of CBP/p300 acetylase and thymine DNA glycosylase links DNA repair and transcription. *Molec. Cell* 9: 265-277, 2002.
129. Vallada H, Curtis D, Sham PC, Murray RM, McGuffin P, Nanko S, Gill M, Owen M, Collier DA. 1995. Chromosome 22 markers demonstrate transmission disequilibrium with schizophrenia. *Psychiatr Genet* 5(3):127-30.
130. Verma R, Mukerji M, Grover D, C BR, Das SK, Kubendran S, Jain S and Brahmachari SK (2005) MLC1 gene is associated with schizophrenia and bipolar disorder in Southern India. *Biol Psychiatry*, 58, 16-22.
131. Wang Z, Rolish ME, Yeo G, Tung V, Mawson M and Burge CB (2004) Systematic identification and analysis of exonic splicing silencers. *Cell*, 119, 831-45.

132. Weaver, B. K.; Kumar, K. P.; Reich, N. C. Interferon regulatory factor 3 and CREB-binding protein/p300 are subunits of double-stranded RNA-activated transcription factor DRAF1. *Molec. Cell. Biol.* 18: 1359-1368, 1998.
133. Werner T (2000) Computer-assisted analysis of transcription control regions. Matinspector and other programs. *Methods Mol Biol*, 132, 337-49.
134. WHO. *Schedules for Clinical Assessment in Neuropsychiatry*, (World Health Organization, Geneva., 1996).
135. WHO. *The ICD-10 Classification of Mental and Behavioural Disorders. Diagnostic Criteria for Research.*, (World Health Organization, Geneva., 1993).
136. Wildenauer DB, Schwab SG, Maier W and Detera-Wadleigh SD (1999) Do schizophrenia and affective disorder share susceptibility genes? *Schizophr Res*, 39, 107-11.
137. Wing, J.K., Sartorius, N. & Üstün, T.B.E. *Diagnosis and clinical measurement in psychiatry. A reference manual for SCAN*, (New York, 1998).
138. Xi ZR, Qin W, Yang YF, He G, Gao SH, Ren MS, Peng YW, Zhang Z, He L. Transmission disequilibrium analysis of the GSN gene in a cohort of family trios with schizophrenia *Neurosci Lett.* 372:200-3, 2004
139. Yao, T. P.; Oh, S. P.; Fuchs, M.; Zhou, N.-D.; Ch'ng, L.-E.; Newsome, D.; Bronson, R. T.; Li, E.; Livingston, D. M.; Eckner, R. Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. *Cell* 93: 361-372, 1998.
140. Yeo G and Burge CB (2004) Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol*, 11, 377-94.
141. Zaykin DV, Westfall PH, Young SS, Karnoub MA, Wagner MJ and Ehm MG (2002) Testing association of statistically inferred haplotypes with discrete and continuous traits in samples of unrelated individuals. *Hum Hered*, 53, 79-91.

Table 1-9

Table 1 Genotyped polymorphisms and allele frequencies.

Gene	SNP	Location (Mb) ^a	Type of SNP	Alleles ^b	MAF SCOTLAND			MAF UK			MAF DK		
					Controls	BPD	SZ	Controls	BPD	SZ	Controls	BPD	SZ
<i>EP300</i>	rs20551	39.872508	Nonsyn	A/G	0.38	0.30	0.38	0.30	0.32	0.29	0.28	0.23	0.27
<i>EP300</i>	rs2294976	39.889208	Intron	C/A	-	-	-	0.08	0.07	0.11	0.10	0.11	0.06
<i>EP300</i>	rs2076578	39.894109	Intron	C/T	0.34	0.33	0.31	0.34	0.32	0.33	0.36	0.35	0.37
<i>EP300</i>	rs5758252	39.898735	Nonsyn	A/T	0.00	0.00	0.00	-	-	-	-	-	-
<i>PIPPIN</i>	rs6002408	40.210909	Promoter	G/A	-	-	-	-	-	-	-	-	-
<i>PIPPIN</i>	rs1006407	40.296532	3' UTR	T/C	0.15	0.20	0.21	0.17	0.18	0.18	0.20	0.12	0.16
<i>NHP2L1</i>	rs8779	40.394776	3' UTR	G/A	0.21	0.18	0.19	0.19	0.26	0.26	0.22	0.12	0.34
<i>NHP2L1</i>	rs132806	40.395293	3' UTR	C/T	-	-	-	0.38	0.34	0.39	0.34	0.41	0.36
<i>NHP2L1</i>	rs1802521	40.400816	Nonsyn	T/G	0.00	0.00	0.00	-	-	-	-	-	-
<i>NHP2L1</i>	rs5758405	40.400881	Intron	A/C	-	-	-	0.18	0.19	0.20	0.21	0.13	0.16
<i>SERHL</i>	rs881542	41.220498	Promoter	C/G	0.20	0.16	0.23	0.23	0.17	0.17	0.23	0.12	0.14
<i>SERHL</i>	rs926333	41.222243	Nonsyn	G/A	0.11	0.14	0.20	0.01	0.02	0.03	0.05	0.09	0.03

^a According to the UCSC Genome Browser, May 2004 assembly (<http://www.genome.ucsc.edu>)^b Major allele/minor allele on the +strand (<http://www.genome.ucsc.edu>)

Nonsyn=Nonsynonymous SNP, MAF= Minor allele frequency, Syn.=Synonymous SNP

- =no data available for the SNP in this sample.

Table 2 Single marker and overall haplotype association analysis in the case-control sample from Scotland.

		Empirical overall p-values ^a				
Gene	SNP	Single marker	2-marker	3-marker	4-marker	5-marker
BPD+SZ						
<i>EP300</i>	rs20551	0.1072				
<i>EP300</i>	rs2076578	0.7960	0.0158			
<i>PIPPIN</i>	rs1006407	0.0487	0.2046	0.0047		
<i>NHP2L1</i>	rs8779	0.3639	4.0x10⁻⁵	0.0016	7.5x10⁻⁵	
<i>SERHL</i>	rs881542	0.7099	0.5484	0.0006	0.0020	5.9x10⁻⁵
<i>SERHL</i>	rs926333	0.0565	0.1627	0.2433	0.0007	0.0109
SZ						
<i>EP300</i>	rs20551	0.7312				
<i>EP300</i>	rs2076578	0.7049	0.6775			
<i>PIPPIN</i>	rs1006407	0.0899	0.3323	0.0285		
<i>NHP2L1</i>	rs8779	0.6759	0.0315	0.2056	0.0957	
<i>SERHL</i>	rs881542	0.2902	0.7437	0.1035	0.0922	0.0164
<i>SERHL</i>	rs926333	0.0131	0.0406	0.1300	0.0122	0.0426
BPD						
<i>EP300</i>	rs20551	0.0132				
<i>EP300</i>	rs2076578	0.9154	0.0011			
<i>PIPPIN</i>	rs1006407	0.0963	0.3200	0.0016		
<i>NHP2L1</i>	rs8779	0.3100	0.0007	0.0176	0.0001	
<i>SERHL</i>	rs881542	0.1884	0.3666	0.0063	0.0308	0.0008
<i>SERHL</i>	rs926333	0.3943	0.5573	0.6804	0.0468	0.2464

^a Empirical overall p-values based on 100,000,000 permutations using the HTR program.
P-values < 0.05 in bold.

Table 3 Distribution selected individual haplotypes in the case-control sample from Scotland.

Haplotype						Haplotype frequency				Empirical p-values ^a		
S1	S3	S4	S5	S8	S9	Controls	Cases	BPD	SZ	Combined	BPD	SZ
		T	A			0.0639	0.0069	0.0038	0.0117	4.2x10-6	7.7x10-5	0.0065
		T	A	C		0.0588	0.0032	<0.001	0.0087	2.2x10-6	5.4x10-5	0.0053
G	T					<0.001	0.0435	0.0625	<0.001	0.0024	0.0002	-
G	T	T				<0.001	0.0382	0.0589	<0.001	0.0096	0.0005	-
G	C	C				0.0874	0.1606	0.1607	0.1649	0.0039	0.0069	0.0234
G	C	C	A			0.0927	0.1478	0.1454	0.1555	0.0244	0.0483	0.0519
		C	A	C	A	0.0161	0.0582	0.0475	0.0684	0.0036	0.0444	0.0012
				G	A	0.0073	0.0306	0.0142	0.0549	0.0799	0.8565	0.0071

^a Empirical overall p-values based on 100,000,000 permutations using the HTR program.
S1 to S9 correspond to SNP: rs20551, rs2294976, rs2076578, rs1006407, rs8779, rs132806, rs5758405, rs881542, rs926333.
- =no p-value value calculated by HTR.

5

10

15

20

Table 4 Single marker and overall haplotype association analysis in the case-control sample from UK.

Gene	SNP	Single marker	Empirical overall p-values ^a			
			2-marker	3-marker	4-marker	5-marker
<u>BPD+SZ</u>						
EP300	rs20551	0.9165				
EP300	rs2294976	0.8545	0.9540			
EP300	rs2076578	0.6831	0.9145	0.6174		
PIPPIN	rs1006407	0.7295	0.9554	0.7203	0.1967	
NHP2L1	rs8779	0.0024	1.0x10⁻⁶	4.6x10⁻⁵	0.0002	7.1x10⁻⁵
NHP2L1	rs132806	0.7380	0.0411	0.2390	0.5064	0.4782
NHP2L1	rs5758405	0.5405	0.5010	0.0218	0.0536	0.1800
SERHL	rs881542	0.1777	0.5431	0.7254	0.1253	0.3031
SERHL	rs926333	0.0378	0.0106	0.0825	1.0000	0.0497
<u>SZ</u>						
EP300	rs20551	0.5960				
EP300	rs2294976	0.2095	0.4662			
EP300	rs2076578	0.8892	0.5282	0.5771		
PIPPIN	rs1006407	0.7103	0.8425	0.5255	0.3510	
NHP2L1	rs8779	0.0149	0.0827	0.2246	0.1940	0.1058
NHP2L1	rs132806	0.7226	0.0546	0.1928	0.3120	0.4166
NHP2L1	rs5758405	0.4549	0.3086	0.0459	0.0926	0.1729
SERHL	rs881542	0.3350	0.8203	0.5076	0.1430	0.2796
SERHL	rs926333	0.3211	0.1065	0.3183	1.0000	1.0000
<u>BPD</u>						
EP300	rs20551	0.5202				
EP300	rs2294976	0.3364	0.4251			
EP300	rs2076578	0.5903	0.4775	0.3769		
PIPPIN	rs1006407	0.8198	0.9425	0.9041	0.2412	
NHP2L1	rs8779	0.0044	2.4x10⁻¹¹	4.7x10⁻⁹	1.3x10⁻⁸	1.1x10⁻⁸
NHP2L1	rs132806	0.2669	0.0038	0.0501	0.2802	0.3318
NHP2L1	rs5758405	0.7431	0.4534	0.0081	0.0144	0.1489
SERHL	rs881542	0.1882	0.3103	0.7406	0.0852	0.1502
SERHL	rs926333	0.0155	0.0157	0.0677	0.4050	0.0714

^a Empirical overall p-values based on 100,000,000 permutations using the HTR program.
P-values < 0.05 in bold.

Table 5 Distribution of selected individual haplotypes in the case-control sample from UK.

Haplotype									Haplotype frequency				Empirical p-values ^a		
S1	S2	S3	S4	S5	S6	S7	S8	S9	Controls	Cases	BPD	SZ	Combined	BPD	SZ
A	C		T	A					0.0127	0.0718	0.0890	0.0493	7.0x10⁻⁶	6.5x10⁻⁸	0.0061
A			T	A					0.0133	0.0860	0.1080	0.0614	6.2x10⁻⁷	2.1x10⁻⁹	0.0016
	C		T	A					0.0280	0.0882	0.1153	0.0530	1.3x10⁻⁵	2.9x10⁻⁸	0.0402
		C	T	A					0.0178	0.0601	0.0835	0.0251	6.3x10⁻⁵	2.3x10⁻⁷	0.0650
			T	A					0.0283	0.1026	0.1362	0.0645	1.6x10⁻⁶	1.4x10⁻⁹	0.0137
			T	G					0.8021	0.7197	0.6874	0.7562	0.0005	1.9x10⁻⁵	0.0933
				A	T				<0.001	0.0267	0.0432	0.0071	0.0088	0.0028	0.0295
						A	C	A	0.0345	0.0108	0.0097	0.0256	0.0068	0.0091	0.1274
							C	A	0.0371	0.0138	0.0106	0.0290	0.0126	0.0077	0.2231
							C	G	0.7687	0.8162	0.8221	0.7862	0.0209	0.0167	0.1399

^a Empirical overall p-values based on 100,000,000 permutations using the HTR program.
S1 to S9 correspond to SNP: rs20551, rs2294976, rs2076578, rs1006407, rs8779, rs132806, rs5758405, rs881542, rs926333.
- =no value calculated by HTR.

Table 6 Single marker and overall haplotype association analysis in the case-control sample from Denmark.

Gene	SNP	Single marker	Empirical overall p-values ^a			
			2-marker	3-marker	4-marker	5-marker
<u>BPD+SZ</u>						
EP300	rs20551	0.4570				
EP300	rs2294976	0.6981	0.8096			
EP300	rs2076578	0.8966	0.0599	0.8173		
PIPPIN	rs1006407	0.0488	0.2144	0.0462	0.3234	
NHP2L1	rs8779	0.3981	0.0797	0.2749	0.0335	0.6045
NHP2L1	rs132806	0.2982	0.5614	0.1823	0.2319	0.0437
NHP2L1	rs5758405	0.0453	0.1958	0.2549	0.3628	0.2453
SERHL	rs881542	0.0005	0.0015	0.0113	0.0461	0.0357
SERHL	rs926333	0.4982	0.0010	0.0015	0.0217	0.0492
<u>SZ</u>						
EP300	rs20551	0.8907				
EP300	rs2294976	0.1734	0.6296			
EP300	rs2076578	0.7536	0.0582	0.6455		
PIPPIN	rs1006407	0.2771	0.7759	0.4973	0.7063	
NHP2L1	rs8779	0.2156	0.0065	0.0557	0.0693	0.1549
NHP2L1	rs132806	0.7415	0.3238	0.0269	0.0872	0.0905
NHP2L1	rs5758405	0.1829	0.4890	0.0221	0.0621	0.1038
SERHL	rs881542	0.0086	0.0237	0.1012	0.0322	0.0190
SERHL	rs926333	0.3166	0.0148	0.0391	0.1481	0.0367
<u>BPD</u>						
EP300	rs20551	0.2597				
EP300	rs2294976	0.6625	0.6913			
EP300	rs2076578	0.9466	0.2034	0.6708		
PIPPIN	rs1006407	0.0219	0.0463	0.0182	0.2915	
NHP2L1	rs8779	0.0034	0.0078	0.0236	0.0928	0.1696
NHP2L1	rs132806	0.1508	0.0251	0.0347	0.0247	0.0123
NHP2L1	rs5758405	0.0302	0.0972	0.0390	0.1322	0.0389
SERHL	rs881542	0.0008	0.0016	0.0082	0.0248	0.0152
SERHL	rs926333	0.8064	0.0011	0.0031	0.0144	0.0214

^a Empirical overall p-values based on 100,000,000 permutations using the HTR program.
P-values < 0.05 in bold.

Table 7 Distribution of selected individual haplotypes in the case-control sample from Denmark

Haplotype									Haplotype frequency				Empirical p-values ^a		
S1	S2	S3	S4	S5	S6	S7	S8	S9	Controls	Cases	BPD	SZ	Combined	BPD	SZ
	C	C							0.1735	0.1046	0.0700	0.1333	0.0423	0.0093	0.3027
		T	G						0.7736	0.8093	0.8780	0.7195	0.3215	0.0036	0.2499
	C	T	G	T					0.1112	0.2130	0.2532	0.1676	0.0583	0.0041	0.4447
		T	G	T	A	C			0.2151	0.3232	0.3451	0.2961	0.0303	0.0046	0.2928
		T	A						0.0212	0.0554	<0.001	0.1279	0.1129	0.0214	0.0008
			G	C	A	G			0.0950	0.0382	0.0444	0.0304	0.0058	0.0296	0.0078
			A	C	A				<0.001	0.0442	<0.001	0.0950	0.0530	-	0.0006
					A	C			0.5902	0.7504	0.7643	0.7345	0.0002	0.0001	0.0045
						A	C	G	0.5563	0.7119	0.7227	0.6984	0.0005	0.0008	0.0058
							C	G	0.7120	0.8298	0.8349	0.8235	0.0007	0.0030	0.0042
						A	G	G	0.2003	0.1017	0.0955	0.1096	0.0037	0.0046	0.0284
							G	G	0.2337	0.1318	0.1159	0.1512	0.0012	0.0013	0.0252

^a Empirical overall p-values based on 100 000 000 permutations using the HTR program.
S1 to S9 correspond to SNP: rs20551, rs2294976, rs2076578, rs1006407,
rs8779, rs132806, rs5758405, rs881542, rs926333.
- =no value calculated by HTR.

5

10

15

20

Table 8 Selected individual haplotypes and SNPs having a consistent distribution inbetween samples.

Haplotype										Haplotype frequency				Empirical p-values ^a		
Sample	S1	S2	S3	S4	S5	S6	S7	S8	S9	Controls	CASES	BPD	SZ	Combined	BPD	SZ
DK								2		0.2300	0.1330	0.1200	0.1400	0.0005	0.0008	0.0086
UK								2		0.2300	0.1700	0.1700	0.1700	0.1777	0.1882	0.3350
Scottish								2		0.2000	0.1907	0.1650	0.2300	0.7099	0.1884	0.2902
Combined								2		0.2000	0.1700	0.1600	0.1800	0.0050	0.0017	0.1295
DK				1	2					0.0054	0.0072	0.0041	0.0121	0.8689	0.8511	0.6778
UK				1	2					0.0144	0.0374	0.0557	0.0181	0.0254	0.0015	0.7145
Scottish				1	2					0.0184	0.0271	0.0298	0.0230	0.3366	0.2798	0.6337
Combined				1	2					0.0140	0.0279	0.0355	0.0179	0.0291	0.0028	0.5858
DK						1	1	1		0.0000	0.0442	0.0000	0.0950	0.0530	-	0.0006
UK						1	1	1		0.0118	0.0196	0.0078	0.0296	0.0324	0.3299	0.0166
DK								1	2	0.2808	0.1700	0.1650	0.1760	0.0007	0.0030	0.0042
UK								1	2	0.2310	0.1840	0.1780	0.2140	0.0209	0.0167	0.1399
Scottish								1	2	0.3150	0.3180	0.2890	0.3680	0.5508	0.5385	0.6798
Combined								1	2	0.2660	0.2120	0.2040	0.2220	0.0011	0.0005	0.0716
DK								2	2	0.2337	0.1318	0.1159	0.1512	0.0012	0.0013	0.0252
UK								2	2	0.1941	0.1636	0.1642	0.1629	0.1437	0.1745	0.2537
Scottish								2	2	0.1970	0.1624	0.1558	0.1760	0.3795	0.1802	0.7788
Combined								2	2	0.2011	0.1551	0.1509	0.1605	0.0020	0.0016	0.0670

^a Empirical overall p-values based on 100,000,000 permutations using the HTR program.

S1 to S9 correspond to SNP: rs20551, rs2294976, rs2076578, rs1006407, rs8779, rs132806, rs5758405, rs881542, rs926333.
- =no p-value calculated by HTR.

5

10

15

20

Table 9 Intermarker linkage disequilibrium measured by D'.
Cases above and right of diagonal, controls below and left of diagonal

DK sample			S1	S2	S3	S4	S5	S6	S7	S8	S9
Gene	SNP										
<i>EP300</i>	rs20551	(S1)		0.60	0.62	0.70	0.76	0.01	0.67	0.17	0.12
<i>EP300</i>	rs2294976	(S2)	0.52		0.74	1.00	0.57	0.84	1.00	0.01	0.10
<i>EP300</i>	rs2076578	(S3)	1.00	1.00		0.42	0.62	0.14	0.39	0.16	0.10
<i>PIPPIN</i>	rs1006407	(S4)	0.43	1.00	0.57		0.95	0.63	0.85	0.14	0.23
<i>NHP2L1</i>	rs8779	(S5)	0.52	0.91	0.74	0.85		0.55	0.87	0.34	0.20
<i>NHP2L1</i>	rs132806	(S6)	0.01	0.75	0.12	0.75	0.76		0.45	0.25	0.17
<i>NHP2L1</i>	rs5758405	(S7)	0.68	1.00	0.87	0.87	0.94	0.89		0.68	0.16
<i>SERHL</i>	rs881542	(S8)	0.19	0.19	0.14	0.11	0.06	0.09	0.29		0.00
<i>SERHL</i>	rs926333	(S9)	0.35	0.24	0.37	0.00	0.06	0.29	0.02	0.71	
UK sample			S1	S2	S3	S4	S5	S6	S7	S8	S9
Gene	SNP										
<i>EP300</i>	rs20551	(S1)		1.00	0.99	0.67	0.28	0.21	0.69	0.02	0.02
<i>EP300</i>	rs2294976	(S2)	1.00		1.00	1.00	0.27	0.84	0.73	0.20	0.04
<i>EP300</i>	rs2076578	(S3)	0.98	1.00		0.47	0.09	0.14	0.53	0.16	0.14
<i>PIPPIN</i>	rs1006407	(S4)	0.73	0.63	0.53		0.72	0.97	0.95	0.00	0.29
<i>NHP2L1</i>	rs8779	(S5)	0.69	0.78	0.51	0.90		0.66	0.66	0.02	0.65
<i>NHP2L1</i>	rs132806	(S6)	0.05	0.70	0.10	0.94	1.00		0.92	0.10	0.40
<i>NHP2L1</i>	rs5758405	(S7)	0.75	0.81	0.60	0.93	0.90	1.00		0.00	0.13
<i>SERHL</i>	rs881542	(S8)	0.02	0.05	0.14	0.03	0.04	0.13	0.07		0.18
<i>SERHL</i>	rs926333	(S9)	0.01	0.01	0.03	0.25	0.30	0.07	0.57	1.00	
Scottish sample			S1	S3	S4	S5	S8	S9			
Gene	SNP										
<i>EP300</i>	rs20551	(S1)		1.00	0.67	0.46	0.06	1.00			
<i>EP300</i>	rs2076578	(S3)	1.00		0.31	0.07	0.09	0.04			
<i>PIPPIN</i>	rs1006407	(S4)	0.70	0.63		0.93	0.09	1.00			
<i>NHP2L1</i>	rs8779	(S5)	0.69	0.61	0.97		0.11	1.00			
<i>SERHL</i>	rs881542	(S8)	0.01	0.32	0.24	0.17		0.03			
<i>SERHL</i>	rs926333	(S9)	0.21	0.47	0.12	0.07	1.00				

Significant (P<0.05) D' values >0.65 in bold and italic

SEQUENCE LISTING

<110> Aarhus Universitet

<120> Method for diagnosis and treatment of a mental disease

<130> P999PC00

<140> PCT/DK2006/000254

<141> 2006-05-11

<160> 96

<170> PatentIn version 3.3

<210> 1

<211> 18976

<212> DNA

<213> homo sapiens

<400> 1

taacaagtga cagagacaat catggtgctg agggaagaag ggtgaagggtg tatggggccc	60
actcaccttg atgacattcc ttgtgctgaa gcgggtcctg gtgacctggt gccaacagtg	120
ccggacctgc cagaccagaa agaggcagag gacaagcagg ttgccacagc actgaggggtc	180
cctgcagctg tgattctggg agagcaggcc ctgcactggc tcccactcag ggccatggaa	240
tggccccacc acagcacagg caccacagcag tgacaacaac atggcaaagc ttagatcagg	300
cagatggtgt tgggagctag agggcaggcc tggcacatct gatgcagcca gctgcagtag	360
gatagccgcc acctcatagc acccctgtca ctgaggccac ctcacaaagg gggctgatgc	420
cccagcctca actctgacct tgggtgccagg cctgggttagc atgctcttgg tgcaggtaag	480
gacagtattg gggaattgtt caaaagtaat taggaggaag aacttgtttt aaacaaaccc	540
atcctctgga atgaaaagat taagtccct gaatgtgtga tggggtgata taacctgcgg	600
tctgcttgat ttctttcttc ctatgaaata cttggatctg cctgtccttc tattatcaaa	660
catttaatga aactgtatt gaatacatgt attcttgttt aaccactcaa catgcctttc	720
tcgttgcttg cagtatcctg atttcctgtg ggggaattac ttttcccatc ttaggatcag	780
tgttggtggg aaaataaatg ttaccctgaa gactctttta cccaaaagtt ctattcaaaa	840
tattaaatat ctgagcaggc agggacgtca gccaaattata atgagccagg ccttaactgt	900
ttagtggctg gaatTTTTTT TTTTTTTTTT gagaccgagt ttccactctt gttgcccagg	960
ctggagtgca atggcgcgat ctcggctcac tgcaacttcc acctcccaag ttcaagcgat	1020
tctcctgcct caacctccc agtagctggg attacaggca tgcgccacca cgcccggcta	1080
atTTTgtatt ttagtagag acggggggtt ctccatattg gtcaggctgg tatcgaactc	1140
ccaacctcag gtgatccgcc cgcctcagct tcccaaagtg ctgggattac aggcgtgagc	1200
caccgcaccc ggccctcttt tttttcttct tttttttgag acggagtctc actctgttgc	1260
ccaggctgga gtgcagtgtg cagtgggtgcg atctcggctc actgcaagct ccgcctccgg	1320

ggttcacggc attctcctgc ctcggcctcc cgagtagctg ggactacagg cgcccgccac	1380
catgcccggc taatTTTTTg tatttttagt agagacgggt cactgtgtta gccaggatgg	1440
tctcgatctc ctgacctagt gatccgcctg cctcagcctc ccaaagtgct gcgattacag	1500
gcgggagcca ccgcacccgg ccggcctttt tttttttttt tttgagagtc tcgctctgtc	1560
gcctcagctg gaggctaga gtgcagtggc atgatctcaa cttactgtag cctccacctc	1620
cgggggtcaa gcgattctcc tgcttcagcc tcctaagtag ctgggggttac aggcgcccg	1680
caccacgcct ggctaatttt tgtatttttt agtagggacg aggtttcgcc atgttggcca	1740
ggctgggtct gaactcctgt cctcaggtga tgcgcccggc tcggcctccc aaagtgccgg	1800
gattacagcc ataagccgct gcggccggcc ggtgggttac tttttgaatg tttaattcat	1860
tgtttctgcc aataatttgt actaggcagc tcataagaac gccaaagtac gttccttctt	1920
cgacacaggg ccagtttccg cgccagtgcc cccgtagagg gcgggggcca ctcgctcgcg	1980
tagcgtcacc ttcgcttcct gccgcgcggg gccatttccg ctgctgcttc tgtgagtttt	2040
tccggtgcac gcgagtgtt ctgaaacgtc agctgcgctc ccctaggagt gctgagccc	2100
cggaaccgca gccatgggtga gtgcgtgcgc cggccgaggg accgggaagc gggagactca	2160
gggttgcggt cactccctag cctgttagat ggaggcttag agcctgcgcc gtctcatccg	2220
ccggcgttgg gtcttaaccg ccggggccctt agttagaacc ccagcagcgg cgggggtcca	2280
gcagcctggc ccgaggttag aactcaggaa accttaggat ccccgagct gtgttgcttt	2340
gccgcacggt ctgtttgctt ggcaaaaggc ctgcagcgag cggagggggc atggcccagc	2400
ctgggagaaa gagaccaacc cacgtgggag agggcaaaaa gttgggtgaa gaaaacttg	2460
ctccaacttt tttttttcct tgtttctttt catacgaaat ctacgatggc gtgggaccct	2520
gggattcgga tcaacctctc ctaccaccc tacctcacc caccctcggt gctgggggttc	2580
aggaggagat ctcccatggg tctgcgctta gtgcccggc cgagtcacac gcgccaggga	2640
cctagaggag cctttggcat ttgggggctg tggggaacat gaagtgagcg cgactcacgg	2700
cctggagtgc aggaaaggct tctcctgcaa ggcggtggct tcaaaaacca gaagtctggt	2760
agcaccacgt gcggcgtctc taaactcttg gtcgtctctg acttttaaac atctagggaa	2820
agttctaaaa cgtgtgtcct ggcgtagtag ttttggttgg cacttggtga ggagggaaga	2880
atttgtaaag cactgtaact tctatagtta atccttccag tggagcttgg atcccggcgg	2940
tagcactccc cacaaacctc ctcccatgca ttttgaagct tagtacgtgg ctgagcctag	3000
aacttgaagc ttccaaactc agtgctcgct ttaacgcagc tccctgaaac gttttggaat	3060
cggatggtct caggttgaac tttgcagttt tcctcaactg taaaatgggg atgacagtgc	3120
atgagggtgt tgtgaggatc aaatgaaata aagcatggta aacatctagt cctgatgctt	3180
ggtacagagt aagcacgcag tatttggtgt tcaatcaaaa ccttgcaagt tttggctgtc	3240

tggtgtacag	cgaaaccaa	ctgatcatgc	taggtaggga	tttaggttgg	ctggtggtat	3300
gatatcattc	gttaatttat	ttaaccctg	aagaaattga	aaggcaaata	aatacacct	3360
aagcaagttt	caattagtgt	ttataagcta	gtttggtttt	aaaattgagt	ttaacttttt	3420
tgttgactaa	tagtgaaaca	tttaggtatt	taatgattgt	tcagattcct	gaagtcttcc	3480
tcttttatga	atttgatgtt	tattgagatt	tgagaatcct	gatttccaga	gaatgggttg	3540
ttggaaacaa	tggcaatttt	taaagttttt	ttaaattgac	atatcatagt	tgtacatact	3600
ttgggggttc	catgtaatat	ttctaagtat	atctatttta	aaacatgcaa	cagggccagg	3660
cgcggtggct	catgcttgta	atcccagcac	tttgggaggc	tgaggcgggt	ggatcatggg	3720
gtccggagtt	tgagaccagc	ctggccaaca	tagtgaaacc	ctgtttctac	taaaaattca	3780
aaaaaattag	ccaggcgtgg	tggtgggcac	ctgtaatccc	agctacttgg	gaggctgagg	3840
caaggagaat	cgcttgaacc	tgggaggcgg	gggttgcagt	gagccaagat	cacaccactg	3900
cactccagcc	cgtgtgacag	tttgagactc	cgtctcaaaa	aaataaataa	ataaaaacttg	3960
tgataaatta	ttaaccaaaa	tttccagctg	ggcacggtgg	ctcacgcctg	taatcccagc	4020
actttgggag	gctgaggcgg	gtggatcatg	aggtcaggag	ttcaagacca	gcctgaccaa	4080
catggtgaaa	ccccatctct	actaaaaata	caaaaattag	ccaggcgtgg	tggtgcgtgc	4140
ctgtaatccc	agctacttag	gaggctgagg	caggagaatc	actataaccg	gggaggcgggt	4200
gtttgcagtg	agccgagatc	atgccactgc	actacagcct	gggtgacaaa	gcaagactgt	4260
gtcaaagaaa	aaaaaaatcc	ctactgtagt	atcaaataaa	acatttgtat	ggccggacgt	4320
ggtagctcac	acctgtaatc	ccagcacttt	gggaggccga	ggcgggtgga	tcaccttagg	4380
tcaggagttc	aacctggcca	acatggtgaa	acccccatct	ctactaaaaa	tacaaaaatt	4440
agctgggcgt	ggtggtgcat	gcctgtagtc	ccagttactc	gggaggctga	ggcaggagaa	4500
ttccttgaat	atggaaggca	gaggtttcgg	tgagccgaga	ttgcgccact	gcactccagc	4560
ctgggcgaca	gagggagatg	ccgtctctta	aaaaaaaaaa	aaaaagataa	aagaataaac	4620
tccaaatcag	tagaggggaa	aatagtgcaa	atgcaatgat	aatccagag	aaaggaaaga	4680
aagaaaaaga	aacagaacaa	gtggaacaaa	tagaaggcac	aaaagaagga	aaatttgaaa	4740
ccattatacc	aataacttaa	tatgtaaagg	gactgagtgc	tccagttaaa	agacgggatg	4800
tggccaagca	ctgtggctca	agcctgcaat	cccagaaatt	gggaggctga	ggcgagagaa	4860
tctcttttagc	ccaggaattg	gagaccaccc	agggcaaaat	agtgagatcc	tgtctctaca	4920
aattttttttt	tttttttttt	tgagacagag	tctcgctctg	tcaccaggc	tgggtgtgca	4980
gtggcacgat	cttgactcac	tgaaacctct	gtctccttgg	ttcaagcgat	tctcctgttt	5040
cagcctcccg	agtagctggg	attacaagca	tgtgccacca	tgcccggcta	ttttttgtat	5100
tttttttagta	gagacgggggt	tttgccgtgt	tggtcaggct	ggtctccaac	tcctgacctc	5160

aagtgatctg	cctgcctcag	cctcccaaag	aggtgggatt	acaggcgtga	gccaccacac	5220
ccagcctaaa	caatTTTTTT	tttaattagc	caggagttgg	ccgggtgcag	tggtcatgc	5280
ctgtaaacc	agcgctttgg	gaggctgagg	tgggcggatc	acaaggtcag	gagttcaaga	5340
ccagcctggc	caacatgggtg	aaaccccgctc	tttactaata	atacaaaaat	tagccgggca	5400
tggtggcgca	tgcttgtaat	cccagctact	caggacgctg	aggcaggaga	attgcttgaa	5460
cctcggaggc	agaagttgcg	gtgagccgag	attgtgccac	tgactccag	cctgggcaac	5520
aagagcaaaa	ctccgtctca	aaaaataaaa	aataaaaata	aataaaaatta	gccaggagtg	5580
gtggcccaca	cttgtagtct	cagttactga	ggaggctgag	gcaggaggat	cgcttgagcc	5640
ctggaggctg	aggctgcagt	gagctattgt	catgccactg	cactctggct	tcagtgatag	5700
agtgagaccc	tgtctcaaca	acaaaaaggg	tgtaattctt	TTTTTTTTTT	TTTTtgagat	5760
ggagtttcac	tctttcgccc	aggctggagt	gaagtggcat	catctcaacc	tactgcaatc	5820
tccgcttccc	gggttcaagc	cattctcctg	cctcaacctc	caagtagctg	ggattacagg	5880
tgcttgccac	cacaccaga	taatTTTTgt	attcttagta	gagacggggg	tttcaccatg	5940
tcatccaggc	tggtcttgaa	ctcctgacct	caagagagcc	accaccttg	gcctcccaa	6000
gtgctgggat	tataggtgtg	agccactgtg	cctggactgc	cccagctaag	TTTTtaaata	6060
TTTTtagaga	tgggggtctc	actgtgtggc	ccaggctggg	ctcaaacttc	aggcctcaat	6120
cagtccccag	cctcagtttc	ccaagttggg	gggattatag	gcctgagcca	ctgcagccag	6180
ttcagaatac	actcttcaaa	catggaacgt	TTTTgaaaat	tggccataaa	gcaagccctg	6240
caaattttta	cattggtaat	attacatacc	atttttgcct	aacccaaaag	gaaataagtt	6300
aaaaattagt	aacaaagtag	tttgaaatga	aaaatataaa	ctatcagaga	gatttaaaaa	6360
ttgtagtagg	aattcaaaaa	tagaacagaa	caataatgaa	aataccacat	accagagcat	6420
aagtgacaag	ggcagagggg	ggaagaaaca	aatgaaagga	tataatgaaa	gtattacttg	6480
gagcctaaga	tgctaattgct	agggagaaac	gggataaaaa	ggtagaaaaa	aacactagca	6540
caaaataacc	caaagaaag	tagaaggaaa	gaaaaaaaac	aaaaggagaa	attgctgaaa	6600
tagaaagtag	agcatactgc	tgggtctttt	cctttccgaa	gcccctcttt	ctcaccagag	6660
agagctgttc	tttctctttc	ttttgcctat	taaacctctg	ctcccggcca	ggcgagtg	6720
ctcacacctg	taatcccagc	actttgggag	gtcgaggcag	gtagatcacc	tgaggtcagg	6780
agttcgagac	cagcctggcc	aacatggcaa	aaccagctct	ctactaaaaa	atacaaaaat	6840
tagcccgggtg	tggtgacgcg	cacctgtaat	cccagctact	ggtgggggct	gaggtaggag	6900
aattgcttga	accaggagg	cagaggttgc	agtaagccaa	gatcgacta	ctacactcca	6960
gcctgggcaa	cagagcaaaa	ctccatctca	aaaaaaacaa	caaaaaaaa	ccctgctcct	7020
aaactcctgt	gtgtccatgt	cctaaatttt	cttggcatga	gacaacaaac	cctggatatt	7080

taccccaggc	aacgaagcta	gttcagtttg	acactgtttgt	caaagttttg	gggaattagt	7140
cactgtcatg	cattgctggt	gggtgtgcaa	gttgttaaac	cctatatagg	gtcaatttag	7200
aatagatata	aaagttacaa	atgcatatta	tatcttttga	ctcgttaatc	ccacttctag	7260
gagtttatct	tacagatata	ctgaatgaaa	cataatatag	attaggtata	gatgagatta	7320
gacattatag	catttacata	atagtgaag	attaaaaaat	atataaatgg	gctgggcaca	7380
gtggctcaca	cctgtaatcc	cagtactttg	ggatgcccag	gcaggaagat	tgcttgaggc	7440
caggagttca	cgaccagcct	ggacaacata	ataagacagt	gcctctaaaa	ttaaaaaatt	7500
agccagacgt	ggtggtgcat	gcatgtagtc	ccagctactt	gggagtctga	ggcaggagga	7560
ccacttgagc	ccaggaaagt	ccccctcaga	aagattaatc	gaaaaaataa	ggccaggcac	7620
ggtggctcac	gcctgtaatc	ccaacacttt	gggaggctaa	ggcaagcgga	tcacttgagg	7680
tcagcaattg	ggagaccaga	ctggccaaca	tggtgaaaga	ccttctctac	taaaaataca	7740
gaaaaattag	ccgggcgtgg	ttgtgcatgc	ctgtagtacc	agctgctctg	gaggctgagg	7800
ctggagaatc	gcttgaaccc	cggaggaggt	tgcaagtgagc	caagatgcca	ccactacact	7860
ccagcctggg	tgacagagcg	agaccctgtc	tcaaaaaaga	aatgttttaa	gtaggtaaca	7920
tgtgaaggag	acatactcta	tagacattaa	aaagaaagga	tggccagggtg	cggtggctca	7980
tgcttgtaat	cccagcactt	tgggaggctg	aggcaggcag	atgacctgag	atcaggagtt	8040
cgagaccagc	ctggccaaag	tggtgaaacc	ccgtctctac	taaaaataca	aaaatcagcc	8100
ggccgtgatg	gcggtcacct	gtaatcctag	gtactcagga	ggctgaggct	gcagtgagcc	8160
gagatcgcg	cactgcactc	caacctaggc	gacagaggga	gactccgtct	ctaaataaat	8220
aatgttatga	ggaagagaaa	aataatatgg	aggtcttccc	ttgacctccc	ccaaccccct	8280
tccttttagag	aactcctcca	tccttcacct	aattcttcct	agttgtcctg	ggaccaagac	8340
aaattcgcct	aagataagga	gttacgtaga	agacagggcg	tcctgggtgta	cgggctggcc	8400
ctggacgtca	cttccgctgg	cggcagtggc	gtcacttcct	gctcttgggg	gcggggcatc	8460
cgtgtccttg	cggtgctggg	cagcagaccg	tccaaaccga	cacgcgtggt	atcctcgcgg	8520
tgtccggcaa	gagactacca	agacagacgc	tatggtgagt	gttcctactg	ggccattatc	8580
tccgggtcac	ctcccagagag	ctgtgggtgcc	gcgagcgtgg	tcccgctgtg	tcagaaaagc	8640
tgtaaaaaaa	cggaaggagg	aaaaggcagg	tgcgcggtgt	cgggcaggag	accaccgagg	8700
caggcgctat	agatgagtgg	tcccaccagg	ccatctccgt	ctccagcgac	tgccgtccac	8760
cctatgtcca	gcctccaagg	ccgaggttag	gtggtagccg	cagtgaatca	ctcgcgattc	8820
gcacgttaca	cacgatggcg	acatgtgggg	ccgcacctgg	gttataatat	gaatatatcc	8880
ttgtatataa	atatgtgttt	gtaaaaacaa	aaaaaaaaag	agaaaaatta	atgaaaatga	8940
tttactggag	cagtgggaat	gagatggaag	gagcggaata	ggcaggccac	tcaccaata	9000

tgctccctac	atgggtgtgt	atttatttat	ttattgagac	ggagtctcgc	tctgtcgccc	9060
acgctggagt	gcagtggcgc	catctccgct	cattgcatcc	tccgcctcct	gggttcaagc	9120
tattctgcgt	cagcctctcg	agtagctagg	actacaggcg	tgagacaccg	cgctcgtccc	9180
ctacatgggt	tattttgaca	cagatctgaa	tgacatcatt	ttatccataa	gtattacagt	9240
atatatctat	aaagtaagaa	tttttaacat	aactgcaaatt	gccattaatt	agttaataag	9300
ggttgcaagt	tggtgatatt	taactgtacc	attccctctt	tattttattag	gcggaagagt	9360
gatacaagag	gtactttctca	tcttctacct	ggttaccgcc	taatacagca	tttaggaaag	9420
gcacagaaga	tgattgtttc	cctttcccag	ttttcaaaat	attgagttgc	tatcacttgc	9480
atagttttga	cttttaaac	atgaagatat	ttacatat	tgaaaaataa	aatttaaagc	9540
aagctgacta	caaaataaat	ctacttgtat	accaaattgc	taaccacaag	gaaaatttta	9600
atttgaataa	cttttgaaca	aagtcctgcc	actggatggc	tttgatagga	tatatcctaa	9660
gaacaaaaat	aaattttgac	tttggtgagt	ttggtggtag	aggtggtata	ttggtgtagc	9720
aattttggaa	ccattttgtg	tgtattgtag	aggagggtgag	gaaatacttt	aatgtgttgg	9780
aaaccatggg	tttgaacaga	agatacgc	atggagtggg	gaatggaaag	aaaactttgt	9840
gctacattta	ctgtaaatta	tatcttattg	attcagtaaa	ttcagggtgga	atacgggaagt	9900
tcaaatttaa	agattaccca	tggactcctg	acctcagggtg	atccacccgc	ctcggccgcc	9960
caaagtgctg	tgattacagg	cgtgagccag	gagtttgaga	ccagcctggc	caacacgggtg	10020
aaaccctgtc	tccacaaaaa	atacaaaatt	agcctggcat	ggtggtgcac	gcttatagtc	10080
ccagctactt	gggaggctga	ggcacaagaa	tcatttgaac	tcgggaggca	gagattgcag	10140
taagctgaaa	tcgcaccact	gcactccagc	gtgggcaaca	gagcaagact	ccatctaaaa	10200
aataaaaaaca	gaaatggatt	agccatggac	tcaaatggat	gcggagccat	gggaacccgt	10260
ttgtaatcca	tttgctgtaa	tattatcaaa	cattcggctg	aagaaagaca	aaagtacatg	10320
agcttttagta	tactttactg	aactgtgtaa	ttcctaatag	tgtatttcct	ttttttcgaa	10380
ggtaggagtg	gaaagtccat	tgttaactac	ctggtgtttt	gtgtgtttga	gggttaggga	10440
gcagggcagc	tgccccctgt	gtctcaattg	ttgattagta	ttttgtgact	caggtaggaa	10500
aacttctggc	tcaattttgtc	tgatatTTTT	atgtcccccc	catagactga	ggctgatgtg	10560
aatccaaagg	cctatcccct	tgccgatgcc	cacctcacca	agaagctact	ggacctcggt	10620
cagcagtc	gtaactataa	gcagcttcgg	aaaggagcca	atgagggtaa	ggcagccaca	10680
ggatgctctg	ggggaatggt	aaaggcacct	gtttccatct	ggggcctctg	atctgatttc	10740
cagtttggga	ggagggctag	ctgctctcga	ctacaaccag	tgggaggact	ttggcattag	10800
atgctaaagc	tgatgggcta	agaatttttt	ttagaccac	caattgttaa	cattttgcca	10860
catttacctt	ctttttcact	ctataagtga	attttatttt	tgctgagcca	tttgaaagca	10920

gttgtggaga	ctgtatagtg	gctgtatgac	ctcaccagag	ccatcagaag	gaaagggata	10980
ttaatgcttt	ttagcatccc	ttatacgggtg	gatacttaac	attcatgaat	ttcattttggt	11040
tcccagaaac	attcccctacc	cctctccctt	tttttttttg	atgcaattat	tcaaaaacat	11100
ggaacgtttc	acgaatttgc	atgtcatcct	tgtgcagggg	tcattgcta	cttctctgta	11160
ttgttccaat	tttagtatat	gtactgctga	agcgagcact	acctcccttt	gtaaaaaagg	11220
aaaaaaaaaa	agctgttttt	cagtagaata	attatgtatg	catgatgctt	ctgcccagag	11280
gtgatatagg	atactctcac	tttattcatt	attatTTTTT	taattttatt	ttattttattt	11340
attttattttt	ttattttggag	agttttgctc	ttgttgccca	agctggagtg	ccatggcagg	11400
atctcagctt	actgcaacct	ccgcttcccc	ggttcaaagt	attctcctgc	ctcaacctcc	11460
caagtagctg	ggattacagg	catgcaccac	cacgtccagc	taatttttga	tttttagtaa	11520
agatgggggt	tcactatgtt	ggtcagtctg	gtctcgaact	cctgatctca	ggtgatctac	11580
ctgccttggc	ctcccaaagt	gctgctggga	ttacaggcgt	gagacactgt	gcccagcctt	11640
attttttattt	ttatttttttt	tcaagtagag	acgaggtttc	accatgttgg	ccaggctggt	11700
ctcgaactcc	tgacctgaag	tgatccgcct	gccttggcct	cccaaaattc	tgggattaca	11760
ggcgtgagcc	actgtgcccc	gctgcccagc	cttatttttta	ttttattttta	ttttttttgga	11820
gacagaatct	cgctctgtcg	cccaggctgg	tcttaaactc	ctgagctcag	gcaatctgcc	11880
tgccttggcc	tcccaaagtg	ttaggataac	agcaggcgtg	agctaccatg	cccagccaga	11940
atactctcat	tttaatagat	gctgggtagt	ttgatgtggt	agagctcttt	tgtcgttgaa	12000
gattttttaat	ggaaacatca	gagattttgt	caactgtggt	tgagatttag	gagaactatt	12060
acaatcgctt	ttacctgtgt	tgctgaagga	agataaagta	gactctgtgg	gtctcgtcgt	12120
agtgcctctt	cctccctgct	cgtcacactg	cagtaactgc	agtgcctggt	atagcagtgg	12180
ccagtggggc	attgtagggg	gaccaccag	ctctcctggg	aggtctctct	tctcacttga	12240
tgcactcagg	tcagcttggc	tgggcaggga	tgtgacttct	tacacttcag	gtagatgagg	12300
actcatggac	acagagcaac	cttattttcct	cttgccatac	tttgcacggg	gataagcaca	12360
caggtccttg	aatggagatg	ctcagtctca	gcgacttaat	tcagaatatg	cttactggga	12420
gttcagggag	ggagatgcaa	caattttccg	tacttctgcc	tcagtgtcag	gtattctttc	12480
ttgccatatt	tattgattgc	ctgctatgtg	ccaggcgcct	gtatatatgc	tgatgatgca	12540
ctggtgaaca	aaatgaagtc	tctgccttct	tggagtatac	attccagtgg	agaaagacat	12600
aggtttttatg	cctggtggta	aagaatgctg	taagaggccg	ggcgcggtgg	ctcacacctg	12660
taatcccagc	actttgggag	gccgagggtg	gcagatcatg	aggtcaggag	atcgagacca	12720
tcctggctaa	cacggtgaaa	cctcatctct	acaaaaata	caaaaaatta	gccgggctg	12780
gtggcgggcg	cctatagtcc	cagctgctca	ggaggctgag	gcaggaaaat	ggtgtgaacc	12840

cgggagggcgg agcttgcagt gagccgagat ctcaccactg cactctagcc tgggcgacag 12900
 agtgagactc cgtctcaaaa aaaaaaaaaa gaatgctgta aaacagaata aagtagagaa 12960
 gggggatagg aatggcagtg gtgggtcttt ttagatggag tattcagaga aggtctgtct 13020
 gaggagtttg agcagagaac tgaatagatg aggggtgttg aaagaaacgt ttttgggcat 13080
 ggtgtaaagg catgcttgag ggattctaag gaggctgggtg tgtggctgga actaagtgtg 13140
 gggatgagag gtactaggag atcacatgag accatgtagg ccactgttag cagtgagtac 13200
 aatggtaaata gagtagaagg attttgaaca gcaagattgc tatgatctta cttaacactt 13260
 ataaaagagt cactcctatg actttttagt ggtgagtaag ctatagtaat atcaatagaa 13320
 atgaacatgc tttgcatttg ccatgtgtca ggtattatta ttattattta ttttactttt 13380
 ttttgagata gggctttact ctgcagccca ggctggagtg cattggcatt gatcatggct 13440
 cactgcagcc acgacctcct gggcttaggt gatcctcca cctcagcgtc ccaggtagct 13500
 ggaactacag gtgtacgcca ccacacctgg ctaatttttt atatttttag tagagatggg 13560
 gttttgccat gttgcccagg ctggtctcaa actctgagtt caaacaatct acctgcctca 13620
 gcctcccagt gggctgggat tacagggtgtg agccaccatg cccagcctca tcattcttat 13680
 taactggttt aatcctttca ataatcctat taagtagaat tattaggtaa ttagaattag 13740
 gttaaaaaga gctgaggtgt ggggtgttcgt ttctcaggtg aaacatggct aaaagcttac 13800
 ggagtaagtg gaaaagaaag atgcgtgctg aaaagagaaa aaagaatgcc ccaaaggagg 13860
 ccagcaggct taaaagtatt ctcaaactag acggtgatgt tttaatgaaa gatgttcaag 13920
 agatagcaac tgtgggtggtg cccaacatt gccaaagaaa aatgcaatgt gaggtaaaag 13980
 atgaaaaaga tgacatgaaa agggagactg atactaagag aaacaaaaag actcttctag 14040
 accagcatgg acagtaccca atatggatga accaaaggca aagaaaaagg ctgacggcaa 14100
 agcgagagaa aagaaagggg aaaagcaaag caaaagcagt gaaagtggca aagggttttg 14160
 cgtggtagac tcttaaaacc ttggaaaatg ccacatggga tagatgatgg attagaatgt 14220
 atatacatgt atactttgat ttcagctgcc accaaatgaa aaccatttgt tttcctat 14280
 taacttggcc ttttttcttc aattcaaacc cagcataact cctgaagttt gttttgggaa 14340
 cttgaataaa atattttctt tgatacaaaa aaaaaaaga attaggttaa aaagaattag 14400
 gtaattcttt ttttattagg tagaatcctg ttaggtagaa ttattttata gatgataaaa 14460
 gtaagggaag ttaaacacct tgctcaaggt catttgtgta acaagaatga gtgccagcat 14520
 ttgcactcag acagtgtagc tctagagtgt gtgcccttct caagtactcc agcatgaatg 14580
 cagagtattc gcactgagcc acagaggcaa gcagtcagga gagtacgcat ctgaaacggg 14640
 gctgtggggg gttgaacttg catgggtgagt cgcagtgtgg atcagctttc tagaagtatt 14700
 gattatggca gattgagatg gtaaaactct gactctcagt cacagatacc ttgagaaata 14760

ctatatattatt	aataatactac	tacttgctgg	tggggggtttg	tggaagctgt	tgcagttgac	14820
aaggtaagag	atgatggtga	cttggacgag	tatagcagag	gagtagaaag	aagtgggtcag	14880
ggccaggcgt	ggtggctcat	gcctgtaatc	acagcacttt	gggaggccaa	ggtgggtgga	14940
tcacaagggt	aggagtttga	gaccagcctg	gccaacatag	tgaaacaaaa	aattagctgg	15000
gcatggtggc	aggcacctgt	aatcccagct	acttgggagg	ctgaggcagg	agaatcactt	15060
gaacctggga	ggcagagggt	gcaatgagcc	aagattgcgc	cattgcactc	cagcccgggc	15120
gacagtgtga	gactccatct	caaaaaaaga	aaaaaaaaaa	aaaaagaaag	aaaaaagttg	15180
tcagattagg	acaggcacag	tggctcatac	ctgtaatttc	agcacttttg	gaggccgagg	15240
tgggtggatc	acttgaggtc	aggagtttga	gaccagcctg	gacaacatgg	tgaagccctc	15300
tctctactaa	aaatacaaaa	attagctaaa	cacagtggca	cgtgcctgta	atcccagctg	15360
cttgggagac	tgaggcagga	gaatcactta	aaccggggaa	gcggagggtt	aagtgagcca	15420
agatcgcacc	actgcactcc	agctgtgtga	cagagcaaga	ctccatctca	aaaaaaaaaa	15480
agaagaagaa	gaagtgggtca	gattcaggac	gtgtttttgag	gataagtggg	ggattttaatg	15540
tagggcttga	gaaaatgcc	ttgtgtatga	ctgactgggt	tttgaatcct	ttatcatatc	15600
ccagggccgg	tctaggacac	ataattgtgt	gtgcctaggc	attcttccca	gtttgttgtg	15660
gctgccctgc	ccaagcccac	atcacctatt	agcttcacgt	ccttctctcc	gcagccacca	15720
aaaccctcaa	caggggcatc	tctgagttca	tcgtgatggc	tgcagacgcc	gagccactgg	15780
agatcattct	gcacctgccg	ctgctgtgtg	aagacaagaa	tgtgccctac	gtgtttgtgc	15840
gctccaagca	ggccctgggg	agagcctgtg	gggtctccag	gcctgtcatc	gcctgttctg	15900
tcaccatcaa	agaaggctcg	cagctgaaac	agcagatcca	atccattcag	cagtccattg	15960
aaaggctctt	agtctaaacc	tgtggcctct	gccacgtgct	ccctgccagc	ttccccctg	16020
aggttgtgta	tcatattatc	tgtgttagca	tgtagtattt	tcagctactc	tctattgtta	16080
taaaatgtag	tactaaatct	ggtttctgga	tttttgtgtt	gtttttgttc	tgtttttacag	16140
ggttgctatc	ccccttcctt	tcctccctcc	ctctgccatc	cttcattcctt	ttatcctccc	16200
tttttggaac	aagtgttcag	agcagacaga	agcagggtgg	tggcaccggt	gaaaggcaga	16260
aagagccagg	agaaagctga	tggagccagg	acagagatct	ggttccagct	ttcagccact	16320
agcttcctgt	tgtgtgcggg	gtgtgggtgga	attaaacagc	attcattgtg	tgtccctgtg	16380
cctggcacac	agaatcattc	atacgtgttc	aagtgatcaa	ggggtttcat	ttgctcttgg	16440
gggattaggt	atcatttggg	gaggaagcat	gtgttctgtg	aggttgttcg	gctatgtcca	16500
agtgtcgttt	actaatgtac	ccctgctgtt	tgcttttggg	aatgtgatgt	tgatgttctc	16560
cccctaccca	caaccatgcc	cttgagggtga	gcagggcagc	agcataccaa	agagatgtgc	16620
tgcaggactc	cggaggcagc	ctgggtgggt	gagccatggg	gcagttgacc	tgggtcttga	16680

aagagtcggg	agtgacaagc	tcagagagca	tgaactgatg	ctggcatgaa	ggattccagg	16740
aagatcatgg	agacctggct	ggtagctgta	acagagatgg	tggagtccaa	ggaaacagcc	16800
tgtctctggg	gaatgggact	ttctttgggt	gacacttggc	accagctctg	agagcccttc	16860
ccctgtgtcc	tgccaccatg	tgggtcagat	gtactctctg	tcacatgagg	agagtgctag	16920
ttcatgtgtt	ctccattctt	gtgagcatcc	taataaatct	gttccatttt	gatgacagaa	16980
tggcagtgcc	tgattcttag	taagtatccg	gctaaaattg	gaatgaagac	atacttagct	17040
ttactagagt	ggtgatagct	gcttaagact	ttttttttct	tctttgtgaa	agagtttcac	17100
tcttgttgcc	taggctggag	tgcaatggca	cgatctcggc	tcactgcagc	ctctgcctcc	17160
tgggttcaag	tgatgctcct	gcctcagccc	ccaaagtagc	tgggattaca	ggcgtgcgtc	17220
accacgcca	gctaattttg	tatttttagt	agagacaggg	tttcaccatg	ttggccaggc	17280
tggctttgaa	ctcctgacct	caggtaatct	gcctgtctca	gcctcgcaaa	gcgctgggat	17340
tacaggtgtg	agccaccacg	cctggccact	tgcttagact	tttttttacc	agaaggctat	17400
ggaagagctg	tctcctccta	ccaggtagtt	tctaccttgg	cagttcaaaa	tgtcgtctta	17460
tacacactca	actggagcat	tgctatcact	ggaactggct	ttctcaacct	ccgtgctggt	17520
gacatttggg	gctgggttaat	ttgttgtgag	gactgccctg	tgtgttgtat	gtttagtagc	17580
atcactggct	tctaccgct	aggtggcaat	agcagcacc	ccccacaacc	aaaaaatggt	17640
tccggacatt	gttaaagtcc	cccggcagta	aatcacccc	caattgagaa	tcactgttgg	17700
agaatcaggg	tttctagtaa	cccttggcac	atgcaggatt	ggtattgggt	actaaagcca	17760
ctgccccctc	acttgcctct	tctacttatt	atcttaagga	taagtccctt	gagagcctaa	17820
taactcagct	tgagaaatct	tccaagggtg	aagtaactgg	gtaccatagt	gatcccctct	17880
ttgccagctt	gggccaatat	accctgctc	tagcaagaag	cacacctggg	ggaacctggc	17940
agccacctta	gccagtcgga	cctgttctgt	gcctcctggg	ctggtcattt	ctcccagctt	18000
actccagtgt	tctcatcacc	accctgaag	cagttttata	tttttggatg	gtggaatttt	18060
acacagtgt	tactttttca	cagtcatgca	ttcagtgttt	ttgggtcaatg	acagaccaca	18120
tatatgacag	tgggtccccta	agattgtaat	actgtgtttt	tactgtacct	ttcctacgtt	18180
aagatgcaca	aatgccattg	tgttacaatt	gcctacggta	ttcagtaaca	tgcagtacag	18240
gtttgtagcc	taggtgtatg	tatagtaggc	tgttccatct	agggttttgt	ttgtttgttt	18300
gagacagtct	cactctgtcg	cccaggctgg	agtgcagtgg	tataatctca	gctcactgca	18360
acctccgcct	cctgggttca	agcgattctt	ttgcctcagc	ctcctgagta	gctgggatta	18420
caggagcacg	ccaccacacc	cggctaattt	ttgtattttt	agtagagaca	gggtttcccc	18480
atgttggtca	ggctgggtctt	gaactcctga	tctcttgatc	tgcccacctc	ggcctcccaa	18540
agtgtgggga	ttacaagcgt	gagccaccgc	accagccgg	tttttgtaag	tataactctat	18600

ggtgttttaca cagtgacgaa attgcctaac aatgcattttt tcagaactta cctctgtcat 18660
 taaatgacac attactttttt ttttttttttg agacagtctc gctctgtcac caagggttga 18720
 gtgcagtga gcaatctcag ctactgcaa cctccgcctc ctgggtttta gcgatgctcc 18780
 ttcctcagcc tcctgaatag ctgggattac agacacacgc caccacgccc agctaacttt 18840
 tgtgtttttta gtagagacgg ggtttttacca tgttgccag gctggtcggg aactcctaac 18900
 ctcaagtgat ccaacctgcc tcagcctccc aaagtgctgg gattacaggc gtgagccatc 18960
 ccgcccagcc aacgta 18976

<210> 2
 <211> 149375
 <212> DNA
 <213> homo sapiens

<400> 2
 gcggagtctc actctgttgc ccgggctgga gtgcagtggc gcgatatcag ctactgcaa 60
 gctccgcctc ccgggttcac gccattctcc tgcctcagcc tcccagtag ctgggaccac 120
 aggcgaccgc caccacgcct ggctaatttt ttgtattttt agtagagaca gggtttcacc 180
 atgttagcca ggatgggtctt gatctcctga cctcgtgatc cggccgtctc ggcctcccaa 240
 agtgctagga ttacaggcgt gagccacagc gcccgacaag gctctatatt tttatagggc 300
 atggaattgc gatgtaagcg gaacttacc agatatgaaa gtgctgagag ctgatggctg 360
 agttctatct atacaaatgg tccacaggta ttgtgggtgg caaagaggtc cctgttggct 420
 cagtcctgga gaatgaggag agacttcttg gcagtggttt gtcctggcac tgtgtgttgg 480
 tttttttttt ttttttttga gatggagtct cgcactgtcg ccagggtg agtacagtgg 540
 cgtgatcttg gctcactgca acctccgcct cccagggttca agtgattctc ctgcctcaga 600
 ctccaagta gctgggatta cagggtgcatg ccaccacacc cagctaattt tttgtatttt 660
 tagcagagac aggctttcac tgtgttggcc aggctggctt caaactcctg accttgatgat 720
 ctgcccgcct cggcctccca gagtactggg attccaggcg tgagccaccg tgcctagccg 780
 gttcttttgt ttctttctca attattgtca taaggacaat gccacagagc agtctagaaa 840
 gtaaacaacc actagatttc tttgtggact cttttctcat acctatctat atgtagcagc 900
 aattctacag tttctacat tttgtatcta gctttttctt ttattattat tttcttttat 960
 tttacttttt agagaactgg gtcttgctat gtggcccagg ttggcctcaa actcctgggc 1020
 tgaagcgatt ctctgcctc ggcgtcccaa gaagctggga ttccagggtg atgtcaccat 1080
 gcctggctct atctggcttt ttcttttaac gtttttagttt aagcattttc ccatagtgtc 1140
 acacagtcac cattattatt ctcaagggtt gggctgacct gggtttctaa tcctggctct 1200
 gccatttact ctctttactc cacatcacc ccagcaagtt acctggctca cccctcacct 1260

WO 2006/119775

PCT/DK2006/000254

ggggttcctc	tcaccatcct	ctccacctgg	atgacatgaa	gcagtgcctg	gacagcgcg	1320
aggacagtgc	tcaggaagtg	ctctccattt	tccttcaaga	gtatttccct	ctgcagaacg	1380
gggctaatac	cctatcatgc	aagggtgattg	tgagaattgg	aaagaatgcg	tgtgcattgc	1440
ctggtgcctt	gccatccaaa	tatatatttat	tagagagagc	agcaacatac	gcccccccaa	1500
ccccccgttt	aaactaagcg	gggtccacgtg	gaatgacgct	tttggcgcaa	gtcggaggag	1560
ctcctaacga	cattagcttc	tttgagtccc	ttgggagagg	aaactttctg	aaagactgta	1620
ggattacctt	ccctgaactt	ccgggtcaggt	gcgggtgctgc	cctgccgcgc	cctcttacct	1680
ctggatgccg	acgaaaccgc	aaaactccag	ccccgggaag	cgcgctctcaa	ggaccaccac	1740
cagcgagcgg	ctttgggaga	gtaatctgcg	ccaccctcgc	agcccattaa	ataggcactc	1800
cccgtttgat	tgacaggaga	agtcaccaat	gcgcgtcctg	gggctttccc	tcgtcctcac	1860
gcgggggtgag	ccgcctctct	actcatctgg	cccaatcggg	ggcgccgggg	aggctgtggg	1920
gcgggggctcg	gtgtccggta	gccaatggac	agagcccagc	gggagcggcg	ggggcgggtgc	1980
ggtcgcggct	ggcggcagaa	gcggccgcag	gggtctgggc	agggctgggc	agtgcctgccg	2040
gagcaaaagc	ggtagcggga	gcccggccgg	agctgggtct	ggagacgccg	tggcagcctg	2100
aacggagtgt	gcgacggatt	gggaggtgag	tggccgccga	ggtcaggggt	cgccgcgacg	2160
gttggggcgc	cgcgggcggg	acgcgggggc	gctgggtccgc	ggccggggccc	gcgccccgtc	2220
agggccgccg	cgccgagctc	cgcgccctga	gggaagcggg	tcgggcccgcg	cggaaacttcc	2280
tgtctctgtg	ggagggggcg	agtgcctcggg	gccccggggg	gcgcctggag	aagacatctt	2340
ggggaggacg	agggcgagtg	gggtcggctc	ggggaagggg	cgggggcccga	tgggggaggc	2400
gctgtgggag	cggggcagtt	ttccgggggtc	cggggcctgc	tggagaggga	agtgcaggag	2460
cagggccgctc	gggggttaag	agctgctaata	gagtggggga	ggtgtagaca	gctgcagaag	2520
ggggagctgg	ggcgactttt	ggagggacgc	gatttaggag	aaagctggga	atccgagacg	2580
tgcagagggtt	aggcgaggggt	gccggagcaa	cgctggaggg	aaggcgtggc	tgggacgcgc	2640
cgggtgagagg	gcagggctta	gtcctgggggt	ggcctggggg	gcgacgcagg	acacacgcgg	2700
gtgcgctgat	ttcggggcgc	ttgatgagcg	gcctccattt	ccttccttcc	caccctccgt	2760
gtcccggggg	gggggggggtg	gtgtgtgtgt	gtctgtgtgt	gtgtgtgtgt	gtgtgtgtgt	2820
gtgtagggcg	gggggagcaa	agcccttacc	gcagcacgctc	ctagggctgg	ggccaagcag	2880
gagaggccgg	ggaggtttgc	ccttggtgacc	ctcctcagag	tttccccaga	ctcactgaat	2940
cagaaatccg	aactccttac	tcctctaggg	ctttcaagcc	tggccaggca	tcctgagagg	3000
tccgcaattg	gagaggtagt	gattgctgct	gctgttgaat	ttcattgatt	cattggggtt	3060
cggctccatc	agagaaggctc	gatttttctt	tcagaaggag	ttggttgggtc	aggggagcaa	3120
gctgtgctat	tgctttgggg	cccacacggc	tgcagccttt	gacgtataac	aggacctggg	3180

WO 2006/119775

PCT/DK2006/000254

gatgccactc	ttggcttggg	tgccctgaggg	aaggtggcct	ttggctgggt	gtgtgcacca	3240
ttcctggatc	ctgtcttctc	tttttccctc	cgctgtccg	gtgtttttgt	ggttaagggg	3300
gtaggctgct	cagagtctcc	agggctactt	ggtacactgt	ttaccaaata	ctgcctggct	3360
gtcttgatga	aatcaggaa	accagctcaa	agagtttgaa	cctcagctgc	agatgggtgag	3420
ctcacacctt	cttccttccc	tccttttctg	tgcccttctc	tgcggttatt	cccgtgtgct	3480
gaagttctgt	tttaaccctc	ctgggattgc	caggttggag	cccagggcaa	gggaaccgac	3540
caggcaggca	gtgttgaatg	ccagccctcc	tgcttctgcc	caagatgttg	tcgttccact	3600
tgattcatga	gcctgcctgc	atgcagcctc	ctcatggaga	cattgaagct	gtgccagtgc	3660
aacttggatc	tctcttgatt	cacctggaaa	actttcactt	gtgtggactt	aaacttaatt	3720
ctagagtttg	gtagcctctt	tggtgaggtt	atgctggttg	agatatgagg	ccactacatc	3780
aaaaacattg	gtgtgttttt	cttttgacaa	tagacttaaa	gctttaaaga	aagtggccat	3840
taggctgggtg	atatttatta	cttgggaaag	taacataaca	cagagaagta	gatgggtgat	3900
taggtctgca	ctctttatcc	actttttggg	agaagaataa	caactcattt	gaatggaaat	3960
ggtatcctta	gctcgtgctg	ctttgggact	tggaacttta	tcacttttaa	tctcaaattgt	4020
atatgaagtt	gtgatgtgat	atttaaataa	tctaggctgg	gcgggggtggc	tcacgcctgg	4080
gatcccagca	cttgggaagg	ccgaggcggg	cagatcacct	gaggtcagaa	cttagagacc	4140
agcctggcta	acatggtgaa	accccgtctc	tactaaagat	acaaaaatta	aggccgggcg	4200
tggtggctca	cgctgtaat	ctccagcact	ttgggaggca	gaggcgggcg	gatcaccaga	4260
ggtcgggagt	tcgagaccag	cttgaccaac	atggagaaac	cccatgtcta	ctaaaaatac	4320
aaaattagct	gggtgtgggtg	gtgctcatac	ctgtaatccc	agctacttgg	gaggctgagg	4380
aaggagaatt	gcttgaaccc	aggaggcgga	ggttgcggtg	agccaagatc	cggccggttgc	4440
actccagcct	aggcaacaag	agcaaaactc	ctgtttcaaa	aaaaaaaaaa	aattagtctg	4500
gcgtagtggt	gcacacgtgt	aatctcagcc	tcagctactt	gggaggctgt	ggcatgataa	4560
ttgcttgaac	ctgggagatg	gaagttgcag	tgacttgaga	tcacaccact	gcactctagc	4620
ctgggtgaca	gagtgagact	ctgttttaaaa	ataaataaat	ctaaaagtat	gcatatatatt	4680
aaaaaaaaatt	ctggatttcc	accaacatag	cattctgcag	ggagataaca	tgaacttttc	4740
cctaatacctg	gatacccatc	cacatctctg	atggccaggc	atttctggag	tgggagcaga	4800
ggggcatttg	tccttgcttt	tttttttttt	ttttttttga	gagagagtct	ctgcccagggt	4860
tcaaatgatt	cctctgcctc	agcctcctca	gtagttggga	ttacaggcgc	ctgccacctt	4920
gcctggctag	tttttgtatg	tttgatatgta	tgtatgtatg	tatgtatgta	tgtatgtatg	4980
tatgtatgta	tttattttatt	ttgagacaga	gtctcgctct	gtcgcccggg	ctggaggcag	5040
tggtgcatc	tcagctcact	gtaacctctg	cctcccagggt	tcaagtgatt	cccctgcctc	5100

WO 2006/119775

PCT/DK2006/000254

agcctcctga gtaactggga ctacaggcat ccaccaccac gcctggctaa tttttgtatt	5160
tatagtagag acgggatttc accgtgttgg ccagtatggt ctcgatttcc tgaccttgtg	5220
atccacctgc cttggcctcc caaagtgctg ggattacagg cacctgccac cttgcctggc	5280
tagtttttgt atttttaata gagatgggggt tttaccatgt tggccaggct gatctggaac	5340
tcctgacctc aggtgatctg cccacctcag cttcccaaag tgctgggatt acaggcgcct	5400
gccaccttgc ctggttagtt tttgtatttt taatagagat ggggtttcac agtgttggcc	5460
aggctgatct ggaactcctg acctccagt atctgcccgc ctcggcctcg cagatgtcct	5520
tgctttttgt ttcttctcta taccttgaaa gtctgtgctg tgtaagggc gtcattgctca	5580
ggcttactga ttctgaagg gttatcagt gatgattccg gatgttgtgg aatatcaaac	5640
agatcgaaaa gtttctttaa taatagtgat ggcaccttgt gtttagagtt cgcagtggca	5700
tttaccagat gctctctggc caccttcatg tgaacctccc agctgactcg ggcctgcttt	5760
atctgccatc agattgtaat ggcagctttg tggaggccgc cctgtagatt atgcgattaa	5820
tgaccacatt catgtgagca atagtgttg cttagttttt ttccagttag gttgtttgga	5880
atggtagctg catacctgag ggcagtgcag caatgcctta gttacgaagg cagtttgcat	5940
tccacaggta gagtctttcc ttatggccag ggcactgatc cctcacatac cagaaagtat	6000
cattcgtccc tgccccagac agactccaga tgcgtcccag tgggtggttct tgggatagtt	6060
ttgtaggaat ttctgcagaa ttttagtggt cagatgagta tgctctgtct gtggactgga	6120
gacaggcttc aaagaattct caaattattc tgcacaatac atcatttttt ggtatctctg	6180
cagttccaaa ggtaactact tgagggacag tggccaccaa agagcacaag ttttgacatt	6240
ttgatgaaaa gcacacataa aacaaataaa aaccaccttt ccttaagtta cctattgaat	6300
tgtgcttatt ccataaatg catgttcac tgtttgctta ataaaactta aagcaccata	6360
gagtctttct tgtttcctag aataaggctc tttttctttt cttttcttgt ctttttttgt	6420
tgttgttgtt gttgagacgg agtctcgctt tgttgcccag gctggaagtg ccctgtggtg	6480
tgatcttggc tctactgcagg ctgcgcctcc cgagttcatg ccattctcct gcctcagcct	6540
ccaagtagc tgggactaca ggcgcccgcc gccacacttg gctaattgttt tgtatttttta	6600
gtagagacgg ggtttcacta tgttagccag gatggtcttg atctcctgac cttcgtgatc	6660
tgcccgctc tgcctcccaa agtgctggga ttacaggcgt gagccaccgt gcctggccgg	6720
tcttttttat tttcaattaa aaaaaatat atatatatat gtatttttaa acagagatgg	6780
gattccgcca tgtttgccag gctggtctca aaccctggc ctcaagtgat cctcccatct	6840
cgggttcca aagtgctagg attacagggtg tgagctacca cacctggcct attttcaatc	6900
ctaaaaaat catttgctg tttcttgatt ttaggcactg aactagacgg agatgagaat	6960
atagaaataa taaaatgggg ctgggcgctg tggctcacac ctgtaatccc agcactatgg	7020

WO 2006/119775

PCT/DK2006/000254

gaagtaaagg agcagatcac ttcagggtcag gagttcgaga ccagcctggc caacatgggtg	7080
aaaccccatc tctgctgaaa atacaaaaat tagtcgggta tgggtggcgtg cgcctataat	7140
ctcagctact cgggaggctg aggcaggaga atcactagaa cccgggaggc agaggctgca	7200
gtgagccgag attgtgccac tgcactccag ccttgggtga cagagcaaga ctctgtctca	7260
aaaaaaaaaa aaaaaaaaaa tagaaaaaat aaaatgatgt ctcaactatc aaagacctcg	7320
caagaagcct taaaaatgat aaataaggct agagattttg ctgtaccttt ggaaagtgca	7380
tgtgtgggtg tttatagtac atatttctgtg gccagataaa atttatatga ttatttttcc	7440
tccaagaat atgagatgac agagggggag gagcgtggat ggagactagg agccttgggc	7500
tcttgtccag cctttccctg accaaccctg aggtctggga caatcagcac agaactccag	7560
gtctcagttt tcatatttgt aaaatgaagg acttagagga ccttcacctt ttgcaaactct	7620
gatagcaact gtgaccctt cccacgaagg gggacagctt cataggcgtg gaccagatt	7680
aagaactgga ttattccatc tctaagctcc ctctcacctg tgatattctg taatcctatt	7740
ttcctgggtca tctttgagct aataataatg atgacttgta tttatagaag atgcacattc	7800
caaagcactt tcacctctca ttcttttgtc ctcaaaacat ttctgagtgg tacagggtggc	7860
aatagatatt ccttcctgtg atgaagaaag tgggccaggc ttggtcagca cgttgatgga	7920
gcaaacagcc agaggccgtc ctgggaccag aactggggct cctggttgcc atcctggggc	7980
ttttccatta ggttttgctt ccttggtggt tgtcattatg agttatgttt ggcagttcaa	8040
atatcccaga atgtgtcact cgtagaggaa actatttggt catctgccgg aaaggactgt	8100
ttttctctct gaatttggtat cactaaacac cttttctgta tgtgaaaagg ctacttattt	8160
gtttttgtta ttttttcttg ccacattttt catagggaca gtttgtaggg cttgagtgtt	8220
ttgagcagac ttattttaatg cttattgatg aaagagtaca ttaaaatgac tacttctcag	8280
ttgcttagca cttgagaaag ttgtgaaata aatgttaagt aaaaattcca aatgtctcac	8340
ttgtcctatc caagtatgtc ctcacagatg agatgctttc tgcagtctag gaggcccctc	8400
cctgggcca gagccgctgg gtgggtttga aagttgggtg tggattgtgc cagtggccag	8460
gactgtccaa cagtggctgc ttgtggcctt ctggctagat tcctaacatt ttatggacta	8520
atTTTTTTTgg tttattgaaa gagtacacat tctgagctat ataaagagct catatatgta	8580
aaagaacact gaacttaggc cgggcgtggt ggctaacgcc tgtaattcca gcactttggg	8640
aagctgaggt gggcagatca caaggtcagg agtttgagac cagcctggcc aatatgggtga	8700
aaccccatct ctactaaaaa tataaaaact agctgggtcat ggtgggtgct gcctgtaatc	8760
ccagctactg gggaggctga ggcaggcgaa tggcttgaac ccaggaggcg gaggttgag	8820
tgagccgaga ttgtgccact gcactccagc ctgggcgaca gagtgagact ctgtctcaaa	8880
aaaaaaagaa gaccactgaa ttttagagcc tcacagatgt taggtggctt atttctgatg	8940

WO 2006/119775

PCT/DK2006/000254

aaaataatca taataaagat gtttttttatt tttattttttt aaaagaatac ttctgatgta	9000
tttctgaaag ctgtttttgaa tactcagaca tggaagctga tagtactgca cacagcaggc	9060
atcactcaat gtggagtgtg tgtggagaat ccacctgtct tcctgttatg tcatgtgtga	9120
tgtgactatt gatagagttg cagccagctg gggcactcca ttaagctgcc cgtggaagta	9180
cctgctttttt attttcattt cactccctgg actgagagtg tttggaggat atctcagtg	9240
ctcatgtgga gggagcatct gtgggtggaa ggcaacaggc agggtagagtc agaggtagtg	9300
gtcctggcca ctgcatgcct ggtgcccctg gctctgcctg cagctctgga atcctgagat	9360
aaagatgatg gctgctctgt aaaggcagct cagtataggc agcaggcctg gaaacttctc	9420
ttttcattgt ggcggtacca gatagtagtg tgtacaaagt actgcacgac agaggagggc	9480
gggtctgact gtcctcagga cactacagag agctgacctt cagggaagtg acttgggagg	9540
gattgtgcca gagccacagg tgctctccag gtggggctga cgtgaaggac cattctgggc	9600
agagctacag agctctgaag aacaaggagt ggctgggaga gaggggtcac gcctgtaatc	9660
ccaacacttt gggaagctga ggtgggcgga tcacttgagg ccaggagttt gagaccagcc	9720
tggccaacat agcaaaacc tgtctctact aaaagtacaa aaattagctg ggcgtggtg	9780
tttgtgactg tgaaccagagg tactcgggag gctgaggcag gagaatcgct tgaacctggg	9840
ggcagagatt gcagtgagcc aggatagtg tacttcattc cagcctgggt gacagagcga	9900
gactgtctca aaaaaaaaaa aaaaaagaac aagaggtgtg tgtgagaggg gctgctaggg	9960
gcagggaagg ctccagggtta tgacctgcct tatattctcc tggaggacag ttcactgggg	10020
agcagtttct gggatatttaa aataccggga tgaagatctt cagagaagca agatgtcagt	10080
tgagaagtca gggatgcaga tgggggaggg gtgattcagg tgactcaatt tcacaaaggt	10140
cacctctgtc tctgggagtc tcaaggattg ccatctgtgt tcctaggaag cagtggctcc	10200
actctaataga gaagggtttt tcgagtggga atggcctcag ccttgactca ttctcttaac	10260
ctgcattttac gtgttttggtt gagaatgcag tagctgtgtg gggtctgctt ctctttcttt	10320
tttctttttt gagacggaat tttgctcttg tcgcccaggc tggagtacaa tgggtgcgatt	10380
tcggctcact gcaccctcca cctcccaggt tccagtgagt ctctgcctc agcctcccaa	10440
gtagctggga ttataggcat gcgccatcat gccgggctaa tttttgtatt tttggtagag	10500
atgggggtttt accatgttgg ccaggctggt ctcaaactcc tgacctcaaa cggcccgcct	10560
gcctctgcct cccaaagtgc tgggatttca ggcatgagcc accgtgcctg gctgtttctc	10620
tttccatatg gtttatgctt acagatcctt gaggttaata tgtcactttc actgttagtt	10680
cctacagaag aaaggtaatt ttgcaataag gttttggagg ttaggaacat ggtttttgag	10740
tcatacagcc cttgggtgtg gggttcgagct ccatcactag gtgactgtgc ggcttttggc	10800
caaatgtttc aatgtctcca aaacttgagt gtcttagtca gttggggctg ctataacaaa	10860

WO 2006/119775

PCT/DK2006/000254

gatgcggtgg	cttgtaaaca	acagaaattt	ccttcttaca	gttctgtagg	ctgcaagtct	10920
gagatcaggg	taccagcatg	gttgggttct	gaggaggggtc	ctttttcatg	ttgcggattg	10980
ctgacttctt	atagcttcac	atgacagaaa	gaaggcgaga	gagctctcag	gtacttttta	11040
taaaggcatg	agtggctgag	tgcagtggct	tatgcctgta	atctcaacac	ttttggaggc	11100
tgaggcagga	ggatcacttg	atgccaggag	ttcatggcca	gcctgggtaa	catagtgaga	11160
ccctatctca	acaaaaacta	aaaaaaaaaa	attgtaagga	cactagtccc	attcatgagg	11220
gccctgccta	attagcttcc	aaaggcccta	cttcctaata	ccatcaaact	gaggggtaga	11280
atttcaacat	atgaattttg	aggggatcca	tacatgcagt	cctttgcacc	aggatactga	11340
gagcccttcc	tcataggatt	ttggcggggc	agcatgaggt	gatggacccc	aggagcccag	11400
catgtgtgcc	tggtcacagt	gctcattcag	tcactctcct	atcagtctac	gtaagtgaat	11460
ttgccagcaa	actatacctt	taaagtgccc	aaagtgattt	tatccagttc	tatttaaaaa	11520
tccttccaga	actttttgga	cctttccttg	gtgactcggg	gtcacgggtg	tagtcgcagg	11580
tcattggtgt	gggtgaggcc	tcctgccctg	gccctgcatg	gactcccctc	atatacctggc	11640
acttacaggg	cctgtctttc	tccccaccac	cttttgtaag	ctgctgttgg	cacctggccc	11700
ttttgttggg	gtaagggcct	gatgatggtc	atatacctccc	tgggttttga	gggcagcctt	11760
gccatgcaga	gaggccttgc	tctcttctgg	gttggttaatt	ctctgcctga	gtttaatgtg	11820
atggattcta	gaaagcctga	ccctgggggtg	aagggatggt	gcttgactac	acgtagtgac	11880
tcaggattgg	aaggagtagc	tggtaatgat	agcttatccc	taggtgatgg	gaaataggac	11940
ttagaaagca	aatgataaag	gtgtgagcac	tgctcacctg	tgtgaaattg	tattaagata	12000
acacgcagca	gtgatagtta	aggcttttaga	attataggaa	gtttctcctt	gttttccaca	12060
cagtaccatt	gccctttttac	tttttagatta	atgttttaagt	aataactttg	taattaaaat	12120
aagtataaat	aaaaagagct	tggttgtcta	ggcatttgca	ttttaccaag	aagctttcca	12180
tagccctcat	tttgttttca	ggacagcctt	gttgatttgt	ttcccttttc	attgttttaga	12240
ggagaaagag	agtcagagct	aaaggggtga	gtgggttttgc	ctgaggccac	ctgccaggac	12300
cccatatggg	cctctgagcc	caagcccaga	gcgcccaggg	aggcctgtga	cggggcctct	12360
gcctggacct	gcagggtggg	cttttagatag	gatcaactgt	gtattctgaa	gtttcctatt	12420
gaatgaaatg	aatcttctt	aatgtaaagt	gatacccaaa	ccgtagtttg	tgagcatttt	12480
catgttttaa	actgcatttg	tctttcatta	tattaaatta	tgttgtcgcc	aaccgagcgt	12540
gccaagatca	tgaattgtgg	tcttcagagg	aaccacagag	caggcagtgc	tccctgtaca	12600
gacgtccaag	cccgttatta	gtgggttctg	tagctgtctg	attggctgga	gctggctgcc	12660
tgctcgccac	ttagctcgca	cgtgggtgatg	ctttaattaa	gcaattgcga	cctttgctcc	12720
tgcaaatgtg	tcatcagata	aggagactgg	gcatgtgcca	ccaaggagga	tgtgcgtggg	12780

ggggtgcagag tcagtgggtga ggggctgtct gcacaggccc ccatgttctg aagtttctga 12840
 gggaccattc aagaagcttt acattcctgc ccatttaatt cttgcaacca gatagtgttc 12900
 tctccacttg atccatcagt aaacctaagc tcagagagat tcagtggcct tcagaggggtc 12960
 accatcagag ctggtgctct gcctaccgaa gcccacactc atttcacctc ctctggctgg 13020
 ctggaataag taagttgtga atcactctca gatgtgaaca tctttaaaag ataattaatc 13080
 tttctcccc acctctgaga aacatagaca tgagggagat tatgtgtgtc ctttgcttaa 13140
 gacggcatct atgtctgtaa ctaacttggg ttcaataatc tgaaagttta caaccattta 13200
 cacagcaggg agataagggc acaaaagggg gagctgctgg ggtgcgggtca aggccgtggg 13260
 tgactcgctg gtacaccac cccatgatgc atcccacgtg ggacaggagg aagattccag 13320
 aggccctggc ttaggagtgc ctcactctcg gcagatgcat tttaaaaggc atcttatgct 13380
 gttaaagagc aggtagacaa gggagaagaa attgtgtttc tgtagaggac tgcactttgt 13440
 aaaattgcga aaggtcttca cttttgtca acgttgat atcacataag agcatttgct 13500
 caggtatatt agagcagata tttgttttgt tttgttttat tttgagacag agtcttgctc 13560
 tgtcgcccag gctggagtgc agtggcgca tctcggtca ctgcaagctc cgcctcctgg 13620
 gttcacacca ttctcctgcc tcagcctccc aagtagctgg gactacaggt gcccgccacc 13680
 acgcccagct aattttttgt atttttagta gagacggggg ttcaccgtgt tagccaggat 13740
 ggtctctatc tcctgacgtg atccgcctgc cttggcctcc cgaagtgtg ggattacagg 13800
 cgtgagccac cgcgcccagc cgatatttgt ttttatgctt gggataatca tcacttttgt 13860
 gtcaaatgac ttcttttttg ctttctttta atcttttaaa aattgaagg tcaacataga 13920
 aaacaatgtt ctgagtaaaa aactctttgt agtttacctt ccacgcacaa agtggttttt 13980
 gtgtatcttg atcaagtaac ctattccctg ccaactttgt accaagcttc agaagaggca 14040
 agagaggtac gttgtctctg tcctgcctca aggcaacact tcgaggttgt tagtttgaat 14100
 gtgaggaagc tgtaggtagt gagatcaggt gcctagttca aggccatcca gtttgtagat 14160
 gggagagatc tggagcttga actgacctg ctgtttcctt ttcaataact gacattgtta 14220
 taattgtatt attcagattc acttaaata tttctcaagt attttccatg ttttaccttc 14280
 atcttcaaaa agattattaa tggttatatt atcctgttga cgtaatgaga ttaaaacatt 14340
 ccctagtgtt ggacatttta ggttatttcc acttttttgc taatgtagat gaatttgcag 14400
 tttataattt catccttttt catctgtgta attagttcct ggggataaat tcccgagaat 14460
 agaaatcttg tattaaaggg tatgaacacc ttaagagtgt tttgaaagcc tgtcttgctg 14520
 cttgtaagtg tatttttctg attttttggg tgaaaaaatg gggattcttt ctctttaaag 14580
 tgcctgcctt ggcctttaga acttccaagt atctgtgaga taagaaagtt ctgtgagggt 14640
 ttcttaaatg agtcacattt ccttcctaag tttctggata tcattttact tgaattgggt 14700

WO 2006/119775

PCT/DK2006/000254

gccatacttt gatgagcaag gaacgctggt ttagaaaatc cttactactt gatgattaat 14760
 ccttggagct aatgttccac gttcccaagg ttgtgagtgc ttatggccac attcctgcat 14820
 cagctcagcc atcgcggtag gtatctcaca gaccttctct agctttcttt tcctcactta 14880
 agggatgcat attggaaaga ctctgcaggg ttctttctgc caggcagcct gggggctcgtg 14940
 tggtccttta gagttcatct ctagcttcta gatgctgctc tcatgggggt tcctggctct 15000
 gttcatctcc agagacaagc atctcaaagg gaacctggga agtttatagc actggtgctt 15060
 gcttgccctg tttgcttggt tgtttttctg ttctgggtctg ctgtggaact tgtcaccatg 15120
 tccccctccc atggactgac tttcctaccc tccactgtcc cttaggtatt cactctcaga 15180
 ctgagtgggtg tcctgttttc ctgggtggctc atttgccagg gtctgcaaag gacaagggtgc 15240
 ttagcagggg cagagagggg ctccaggtgg ggagcagtgg tcaggccagc tgaccttagt 15300
 gcttaccctt ggggtcccct cctagaattc ctgcatgcc tccaccatgt ctaccctctc 15360
 cccacagggg ctgtttccct tttagttctc ttgccttact gtgggcaaaa gcttactttt 15420
 cttggttcct cttatcccag ctaagaaaag ctgatgttca caacaaaatg aagagacttg 15480
 actaccaggg aaggtagaca taagaggaaa cagccacaaa cagcacaggc atacaagtgt 15540
 cttgcttccc aagcacaggc aggtttttgt gccacactt cagggtctgt ctgactgcac 15600
 tggctctgcc aggcgctgcc cagcttcctg cttcccctcc cagtcctgct tcttctgtca 15660
 gtaaaacaaa attttagagt gatgctttgc ttgatgcttt gggggctgca tttttatcac 15720
 aagtgtggaa attgtacttt gctctgggtt ctacagatgc ctaaacaatgt ctaggttgct 15780
 gttgccttgg caactggctg agctgttggt tgtgtgtgtg gtgaagggtt tggcaggggc 15840
 ttcattggaag cctctgggtc tccttagagc ttttcagcta agctggcggt tactcccaag 15900
 cccagaacca tatgctctta gctccccttt gtcgtttgga gcttcccatt ttatttccct 15960
 gagaaatccc tgctcagctg ccctagactt ggggtgggtgg tatggcatgg catggcatgc 16020
 actgataggg ctgatttatt tcgagactgg atttttggta acacaagacg tgattccata 16080
 gcagcccttg gcctaagttt tatttgaggt cggacggggc ccaagggtg cttgtgctct 16140
 cccttcaagt gccttctgct ttttaggcct ggggtgcttg ttggggccca gagctggggc 16200
 ttgcctgaga gtcacgcaaa agcagggcct tccagtactc cagaggggta ctctgacca 16260
 cctgtaagcc atgggaagaa agtcaatgtg attggatcac ggcagtgcc cacagtatgt 16320
 ctctttttat gaagactgag tgaaaggata taatcaattc cagttttatc ctttgttatt 16380
 ttctttaatg cttcttgat ttcttttgg atttatttct tttcttctta gtcttgcaaa 16440
 tgtgagtctg aagttgctga tgtgggggtt ggggctgact catcacctcc ccttatagct 16500
 cagacaagtc accttctta aagttttttg ggtgtacgct catttccaga attgctattt 16560
 tgggcccatt cttttctctc tgctgcctct actaggaact tgcaggttca cgaaatttag 16620

WO 2006/119775

PCT/DK2006/000254

aaatagttta	acataacatg	agctcttact	agtaagtaaa	gaggtgagta	gagagagaaa	16680
cctcctccaa	tcctccaaac	cctccaaaag	cctcatgtta	gcagtgtgat	tggtggaggt	16740
ggagacaaac	tcagtccatt	gggaactcgt	ggggttacca	ggtgggagag	gggagttctc	16800
tccttgggggt	agagttggca	aattaccagt	ccccacaaag	gctgcaaagc	cctgctaccc	16860
actccctgtg	catcctgctg	tccaggatgg	ctggctttttg	gggtcctgac	ctggaacccg	16920
ggttgctttt	ccccatagga	gggcctctgt	ctggagcaga	ggttcaggga	gaagcctctg	16980
gcatgttttg	gtcttgggcc	ggccccgtcc	attggggcta	caagtgcctg	cagtgagctg	17040
ttgttagtaa	ggagctctgc	aaccagattg	cctgggcttg	aatcccagct	ctgagattgg	17100
ttgtgatacc	tacctcacct	ttgaaatggg	gatcatgata	ggattgttgt	gcagatggga	17160
tgcatgaatg	tgtgccctgc	acttgtgatg	gcattgccct	tagagaactc	cattgtagca	17220
cagccatcgg	tgcttctgag	cctgccaacc	atcttgcctt	gggactcttt	gattcccact	17280
ttacagttga	ggaaataggc	tcagagaagc	ttaatatatc	ttcagatctg	ggtctgtctc	17340
agggttctct	cttcaaagca	ctggtgggtcc	cccacagtgg	ccctgtctcc	ccagcaagtt	17400
agggagtgac	tgccagggga	ctgctccatc	agtcctgtgc	ccaccaggca	tagtctgtag	17460
acagcatttt	tttttttgag	atggagtttc	attcttgttg	cccaggctgg	agtgcaatgg	17520
catgatcttg	tctcacggca	acctccgcgt	cccgggttca	agcgattctc	ttcttcagcc	17580
tccaagtag	ctgggattac	aggcatgtac	caccacgctg	ggctaatttt	gtgttttttag	17640
tagagacagg	gtttctccat	cttggtcagg	ctggtcttga	attcccaacc	tcaggtgatc	17700
cactggcctt	ggcctcccaa	agtgtctggga	ttagaggcat	gagccaccac	gcctggcctg	17760
tagacagatt	ttttaatgcc	tccgtacaca	cacatacaca	cacacattag	acattttatat	17820
gttactgaaa	tgggagagtt	ccctggaccc	cttcgaggta	cttgggacag	gggtgtggct	17880
ggtttgctta	aaccacttac	aggaggggga	gcatgcaggt	gagcaagtac	aggggccgag	17940
gcgagcactt	ttgggctctg	gccccacagc	agtgtctagg	gttggttaca	tgctctttta	18000
gccctcctgt	ctgaggacgg	cttaagtgtt	aaacggctta	agagtcagtg	tgacagcctt	18060
tttgggttca	gcatcccga	ttatctggcg	tccaggaaga	atcaggtcac	atggacttga	18120
aggatggtga	atgcgggatt	ttattgagtg	atggaggtgg	ctctcagcag	gatggggagc	18180
tggagatggg	atggagtggg	aaggatgatc	tttcctggag	tttggccatt	ctgcagccga	18240
tctcctcttc	aactgcccc	agctcaactc	ctctcgactt	cctctcttct	ctcctctgcc	18300
actcttctgc	tcctggattc	tggggtttgg	ggttttatat	gggcacagga	tgatggggga	18360
catggtgggc	tgaaaggcaa	catttgggtg	cgaaaacagg	aatgcttggt	ctcatttagg	18420
gccttgggtc	caggcttgag	ggtggggctc	tcaccaggga	ccctgccctt	ctgcctcgtg	18480
tttgtatcat	tactatatat	aatgatgct	gtgattatga	ttaacaaaat	accagtttac	18540

WO 2006/119775

PCT/DK2006/000254

tttaaaactt	gaatttacag	tagccttttg	tcattctggag	tttctaaatc	attgaatgtg	18600
aaaaatacta	ttatttaggg	atctagcaac	cttattgatc	tttaaaaggc	ttaagttcag	18660
aaaaagttta	ttgttgccct	tttcttgctc	tgtgtttttg	tgtctgccct	tctgtctgtc	18720
tggcatctgt	gcgagtcatg	cgtcagggca	tgttggggag	atggaggcca	gtgtggctgg	18780
gttgagttat	aaactggaag	caaggttggg	agcccagaat	ccaaaccagg	ttcagactgg	18840
gagaatatct	gaaatccctg	cagcagcaaa	gaaggttcct	atattcacgg	tccctgaagg	18900
aattttgctg	ggccccagg	ttctctggga	gcattgatgt	gtctcactct	catggtacca	18960
aagactttct	gtttttattc	atcaacttcc	tgtatttcct	tctgctccct	ccttaacctg	19020
gggttggtc	tagtttcaaa	gttaattctc	aaatccaaac	ctcagactta	aaggctggta	19080
actgcccttg	gatatgtttg	gggctggctc	ctgagaaacc	tcttttccag	gaaagctctg	19140
gcctgttaga	ttttgctctg	taattttact	gccttgctca	cctaacgggtg	gtcactgggtg	19200
ttctgagaca	cgtgggtgga	actcagcaat	gtttcttcac	ttgtattttt	gtttgtttgt	19260
tttttgagac	agagtcttgc	cctgtcgccc	aggctggggg	gcagtgggtg	gatctcagct	19320
cactgcatcc	tccacctcac	gggttcagat	gattctcctg	cctcactctg	cctcccagat	19380
gaagtaatag	ctgggaccac	aggcatgcgc	caccatgcct	ggctaatttt	tgttgacctg	19440
tatttgatc	ttatatattac	atgagaattt	catgcaaata	gattctttcc	ccctagaaat	19500
ggactcagac	atttttgcat	gcaattcaat	gtattcataa	atccccattc	atttttatatt	19560
atttatattat	tatttattgt	tttttttttt	gagacggagt	ctcgctgtcg	cccaggctgg	19620
agtgcagtgg	cgccatctca	gctcactgca	acctctgcct	cccgggttca	agcgattctc	19680
ctgcctcagc	ctcccaggta	gctgggacta	caggcgccctg	ccaccacgcc	tggctaattt	19740
tttctatttt	tagtagagac	agggttttac	catgttagcc	aggatgggtct	caatctcctg	19800
accttgat	ctgcccgcct	ctgcctccca	aagtcctggg	ataacaggcg	tgagccactg	19860
tgcctggcca	aatccccatt	tatgtattca	gcaaattttt	gaacgcctgc	catgtgctag	19920
gcacggttct	aggcactgga	gttacaacag	gaacaacgta	gtccatgatc	ttacattctg	19980
gtcagaggag	gcagagcgtg	attaaacacc	cccagcctgg	tgtggaggga	aggccagcct	20040
cagtgggcgg	gaggggtccaa	gcctcatccc	ctcctgccag	ggctctgctt	gttgccaaac	20100
ctcccagccc	ccttctagcc	ttcacttccc	tgaactgggt	tctgacttca	caatcgttat	20160
cagaattatt	tttaaaaaat	ttaatgttta	ttgaggctca	ttttgtgcca	ggtgtcctat	20220
ttttccttgc	gtgattacac	ttattctcat	gacaactttg	ttttgcaagt	gagaaagctg	20280
ggcctaggcc	tcatacctag	gaaatgtgat	gtgtgggttt	gatcagcatg	cagcttgagg	20340
ctgtagaact	cagtaagcca	ggcctgagga	caccatcaga	aataccaaca	aaaaattcag	20400
tgtgtccatt	ttggcctgga	ccgtctctat	tctctgcttt	ttttcccct	gattgagaca	20460

WO 2006/119775

PCT/DK2006/000254

gggcctcact	ctgtcaccca	ggctggagtg	tagtggtgtg	attgtggccc	actgcagcct	20520
ccaactccag	ggcttaatga	tcctctcacc	tcaccctccc	aaattacaag	tgtgtaccat	20580
catacctggc	caattttttt	tttttttttt	ttttgcggca	gagtccttgct	ctgtcaccca	20640
ggctggagtg	cagtggctcg	atcttggtct	actgcaacct	ttgctgcctg	ggttcaggcg	20700
attctcctgc	ctcagccttc	ccagtagctg	ggattacggg	cgtgtgccac	catgcccagc	20760
taatttttgt	atttttagta	gtgatggggg	ttcagcatct	tggccaggct	ggtcttgaat	20820
tcctgacctt	gtgatccacc	catcttggtc	tcccaaagtg	ctaggattac	aggcgtgagc	20880
caccgcgtcc	agtgtcatac	ctggctaatt	ttattatttt	ttttttgaga	cggagtctcg	20940
ctctgtcgcc	cagcctggag	tgcagtgggt	caatctcggc	tcattgcaag	ctccgcttcc	21000
caggttcaag	ccattctcct	acctcagcct	ccaagtagc	tgggactaca	ggcgcccacc	21060
accacgcca	gctaattttt	tgtattttta	gtagaaatga	ggtttcaccg	tgtagccag	21120
gatgggtctt	atctcctgac	cttgtgatcc	gccacttcg	gcctcccaa	gtgctgggat	21180
tacagacgtg	agccaccacg	cccagccaat	acctggctaa	tttttaaatt	tttctggaga	21240
ctggggaagg	gggcaggctt	cactgtgttg	cccaggctgg	tcttgaactc	ctggcctcaa	21300
gtgatcctcc	tgtcccggct	tcccagagcg	ctgggattac	aggtgtgggt	cactgtgcct	21360
ggccactggt	tcttttaaca	atcaaattgc	ctttctcaag	gaggaaaaca	tgggtggaat	21420
gtgctttaag	gcctctgtct	ggtggacatt	ataggcactt	gaagttacca	aatctgtata	21480
tttactgttg	gcaaaattgc	atgatcccat	ccttatttcc	aaataattgc	cgcgatgaat	21540
acattggaaa	tgcacattta	gttcttgact	gagtagttgt	atacacactc	tggtttttcc	21600
acaccagaca	cttgctctct	tgctgtgact	tttaagggtc	tgcccaatgg	acagagtaac	21660
ttcctgaaaa	gttgctttta	aaaccttgct	tcctcagtgg	acggagggtg	caatgaaaca	21720
acatttgta	tcctgactca	gtcccaacca	agcagaacct	ttattgtgac	tggagggttg	21780
caagctgccg	tcctcctggc	accaaagca	gcctgttctt	taaatatttg	cattgaataa	21840
ctttgtagct	ctgcttcctt	ctaaggcttg	ccacacttca	tgtcagaagc	agagctcctc	21900
ctctggagtt	attgtgaaac	atgactgaat	gccagggtgc	agtttctctg	cagtcagctc	21960
actttataga	ggaagaactc	ggctgcttca	gctcaatgtg	agttaaactg	tgcagtgtcc	22020
agtcttgctt	cctggcagtt	ggggcccggg	gcagccccga	ggcggggagg	ggggacattg	22080
tcccctttga	ctatgtgatt	gttggaagca	gggtcctgag	agggagtcca	ggaggcagcc	22140
taaatggaat	ttccagtcta	aacgggcaag	gcacagtgtc	gggagtcctg	gtgtaaagtc	22200
cagtgaagca	gagactgggg	cttgggggtg	gatggggatt	gctgacaaga	gagaaggatc	22260
tctgggcccag	ggaagaggca	gctgcattcg	ctggctgcga	ggctgtggct	ctaggtgtgg	22320
atatgctagg	gtgctccccg	cagccccag	acctttcata	tgtcagcatt	cccgtagggtg	22380

gggcatgtgg gacccctgcc cggggccctt tttccagtgt gactgtccct atagggtgag 22440
 gaatctgagg tatcaagaag tggatatctc ctaaaacttg tgcctctccc tttaggccat 22500
 gcacagggac agagaccctg gaattccctg cccagtaac cccagaccc cttacatggc 22560
 ctgcatgggc cgcctccctg gcgtgtcctg ctgagggtag gctgtgcact gtgtgcacgt 22620
 gcctccagcc aaagggtggct gaatggtggc catgggcagg gtatgtgtgg acagagcttg 22680
 ggtgtgcagg ctgggggtgc ccataagca cgcccgaggg tctttgtggc gcaggacacg 22740
 gctctggatt catgtgtgtg cctggtttct tgaatggtaa agttcatgct tattttctgg 22800
 gtgtgttttg ggttgtccct ccctatctga cacatggggc ttgggaccat gcctggagcc 22860
 ccagttgcat ggggaggttg ggttaagagc aggcctgtcc ccctgctcag cccttcttac 22920
 caccaaacca gccacacttg agcaggagcg cactgcactg tcattttacc caaagtttcc 22980
 ctctagggct ttgctgatga catctcttgg cttttgttag cattctggcg caatatcaga 23040
 aggattccag tctacatcgt ggtcttcttg agagctcata tttgatttcg gctttagctc 23100
 actggacttt agcctggagg aatcccactg aagagttttt aatgttgact tatggaactt 23160
 gtcccagcag caaaccacag atactgtggc ttgggggtga gcagctcttg tggaaacaag 23220
 gcacttccct tcttcagctg cataggagtg cttcaggaac cttttgggcc acattcttct 23280
 tgtgatttct tgacttgtgg gagggcaggg aagtgatgtc tgggtccac agcgcacatg 23340
 gctgtccctg gtgccgccct tccaggagga actgccaggt cggatgtgga gatgaacccg 23400
 gagagtgtcc ttcggctctg gtcggccttc aggggcattc ctgccagtct cagaagatga 23460
 cagtctgcaa tcgtgttatc atcaacgtgg aaaaaataag ggcaatgcag ttttcatgtg 23520
 tgtaaggtta gcagctgccc tttcttggga agggctcttt ctgtataaaa ttttgaccat 23580
 tgtatctttt acattttaaa atgtcctatc tttccatggt gatgaaaaga gctggtagtt 23640
 tctaaaggat attcttactt ggcaaaataa gaagtaagca gtcttatggt attctgcctt 23700
 caaatttgaa acagttttat tgagttataa tgaacaggcg ataagttttg gcatatgaat 23760
 aactgttaa gccatcattg caaagaagat agtgagaagt ccaccctctg agggttcctc 23820
 ttgccacttg gggtcctct gcccccactt cttccaggcc accactgctc tactctttgc 23880
 cactgtagaa tcatttgcat tttctaaaat tttatgtaaa tggaatcata cagtatgtat 23940
 tctatgtatt cttcttcttc cttttttttt tttgagatgg agtcttggtc tgttgctagg 24000
 ctggagtga gtggtgcaat ctcgactcac tgcaacctcc gcctcccggg ttgaagtgat 24060
 tctcctgctt cccagctccc agctcccag tagctgggac tataagtga tgctgccatg 24120
 cccggccaat ttttttgat ttttagtaga gacggggctt cactgtgttg cccaggctgg 24180
 tctcaaactc ctgagctcgg gcaatctgcc cgcctcagcc tcccagagt ctaggattac 24240
 aggtgtgagc caccacttcc ggccagtata tattcttttt acagtatgtg ttatttttag 24300

WO 2006/119775

PCT/DK2006/000254

tctggcttct	ttcactcagt	aaaattactt	tgagattggt	gtagcattta	tcagttagtt	24360
tcatttttta	aaaaatccaa	tgcatgggca	tgctacaatt	tgttcattca	cctattgaca	24420
gacgtttggg	ctgtttccac	ttggtgtaat	cacaaatgaa	gatgctatgt	cgtgaaaaga	24480
tgttacagaa	aaataagggtg	ctattaacat	ttgtgtgcca	gtctctgtgg	ggacatctgc	24540
tttcttttta	tcaaaaattt	aatggagaca	gagtctccct	atgttgccaa	gactgatctt	24600
gaactcctgg	gctcaaata	tcctcctgcc	tcagtcttct	aaagtgttgg	gattacaggt	24660
gtgaaccact	gcacttggcc	aacatctgct	tccacttccc	tggggtaa	acctagagtg	24720
gaatggctag	gtgctttggt	ggtttaggtg	gttaggctag	gtttggtggc	tagggtagct	24780
ttaacatgtc	cagacactgc	caaacttttc	caaagtgtt	tgcgcgcac	accagcactg	24840
ggttctagtt	ccccacatc	ctcaccaaca	cttggtacat	gtagtctttt	taatttttgc	24900
cattctggta	ggtttaattt	gcatttccct	agtgattatc	tctttaaatc	acttgcctgt	24960
tgaaaaagct	gagttgtttt	attaacttag	ttttgaaatt	tttaaaatat	attttggata	25020
caattccttg	atcaattaca	taatttgaaa	atattgactc	cctgtctgag	gctgtctttt	25080
cattctttta	acagtatctt	ttgaagaaca	gaagttttta	acatttgctg	aagtcaattt	25140
gttcttttac	ggcttggtg	tttttggttg	gtgtcacagc	taaaatcttt	gcctaataccc	25200
aggtcacaaa	tgttttctca	tgtctcgtct	agaaatgtta	tagttttatg	ttttacattt	25260
aggcctgtga	tccattttga	acttttaaaa	tgtgttgagg	tgtgtgaatt	gaagggtttt	25320
tttttttttt	tttgtctgtg	gatagccagt	tgtttcagta	ccatttggtg	aaaagtcctc	25380
aaaatttttg	gggaaaattt	ctaattgggtg	aaaccagag	gtctgggaat	tgtgggtctg	25440
agcccttgga	tctcttcacg	actgtttact	gttctcagct	tcgtatggca	tggctttctc	25500
tttatcgcac	ctgggagttc	ccagctcctg	acagcaggag	ctctatcatt	ggatagggct	25560
gggaaaggga	ctgggagatg	cccgtggat	tccgcacag	tctctctcct	tcttctctgg	25620
gccatacaga	ggcttctggg	tggtgggttc	tgggcagtg	ctcatttcca	gaggtgacta	25680
gaggcttgcc	aggacagaga	ggtggaggga	gaagaagcct	tctgtgcctg	gatgggtggg	25740
gtggggctca	ggctgagaaa	ggtgggagg	gctgttcccc	attctgagag	ggctgacagc	25800
cttgagacct	gcagccagct	gagaccag	gaactgggct	ctgctctgag	ccactgcttg	25860
tccttgacca	gctttatgac	tgggggtgat	gggggtggga	gggggttaaca	caagttcatt	25920
cttttccctt	cagtcacaaa	gtggctctga	gcagtctcca	agtctatacc	attgtacagc	25980
agcttcctag	gaggtggtaa	agtctccatc	tgctttttga	tctgggggca	gtgtgttctg	26040
actcataaac	tgctggtact	actagtccta	gcactaggag	ttccaccagc	acctaagaat	26100
gatggagctt	tcttctgctg	taagcacttt	gtgtatagaa	cccgttgagt	cctcacacaa	26160
gcgttggtgag	gtggctgctg	tccattttac	agatgaggaa	actgaggcac	agagaagggtg	26220

WO 2006/119775

PCT/DK2006/000254

accatctgga	agaagtcttt	gggctgtgct	ttttgttatt	gaacattgaa	ggacagatat	26280
ggtttagcat	ttttcattag	ctcttgagcc	aaatagagtg	gtttccatgg	ccctcttcca	26340
catcttcatg	ttcttattgg	caacggggag	gggtcggact	gtctgagatt	aaaattcttg	26400
agtggttctc	tgttgttatc	acaataaaat	ccagacttgt	ggctcgaccc	ctgcttacct	26460
gttgagtctc	atctcttgcc	acactcctga	acttcttgca	tttctcccat	gcaaattctg	26520
tctactgcaa	atgccattct	tcccgtcttt	gcctgctgaa	ctcatccttc	aaaacttggc	26580
tctaatttca	ccttcccctc	cctccagggtg	ttttgcccct	gggcctgcag	agtggatggt	26640
ccacccccgt	ccccatactc	ctcatgtatg	gtatgaccga	cctgtttgtc	ttctgtgggg	26700
ctaggagctg	ggtaggcgga	gggcagggga	ctgcattttg	tattgtagtg	tgcagttctt	26760
agctctgaaa	tggcacatca	taggcaagca	ggctctgtgg	gatgaacaat	gcaggcatga	26820
gtgacatgag	gatctgcagt	cctaagactg	attcattctt	tgggaggagt	atcccttctt	26880
acccaaagcc	agttggtttg	tggccgtctg	tgattgcaag	tgcccagggc	tgccctcatg	26940
tggaaagtca	cgttgactga	ctgcagaaat	tagctgtgtt	taaaaatcat	tctttttaaa	27000
aaaatcattc	atttttttaa	cctgttttaa	tgcttgaact	ggcttacatc	ttacatatat	27060
acacacacat	atatatacac	acatatgtgt	gtgtatatat	atatatatac	atatatatac	27120
atatgtgtgt	atatatatgt	gtgtgtatat	atgtatatat	aatatatatc	tcttttttta	27180
aacctgttta	aatgcttgaa	ctggcttaca	tcttacctat	atatacacac	atatatatac	27240
atatatgtgt	gtgtgtgtgt	atatatatat	atatatatatt	ttttttcccc	ctccaagaga	27300
atgagttttc	ttttcttttt	tttttctttc	ttgttttggt	tttttgagac	agagtctcac	27360
tgtgtcgccc	aggctggaat	gcggtggcgc	gatctcagct	cactgctgca	atctctgcct	27420
cccaggttca	aatgattctc	ctacctcagc	ctcccaagta	gctgggacta	caggcaccca	27480
ccaccatgcc	cagctaattt	tttttgtatt	tttagtagag	acgggggtttc	actgtgttgg	27540
ccaggctggg	cttgaactcg	tgacctcgtg	atccacccgc	ctcggcctcc	caagggtgctg	27600
ggattacagg	tgtgaaccac	cacacctggc	cttttttttt	tttgaggggg	atggagtttc	27660
actcttattg	cccaggctgg	agtgcaatcg	tgtgatctcg	gctcactgca	acctccacct	27720
cccaggttca	agcgattctc	ctgcctcagc	ctccttagta	gctgggatta	caggcatcca	27780
ccaccacacc	cagctaattt	tttgtatttt	tagtagagat	ggggtttcat	catgttggcc	27840
aggctgggtc	tgacctcctg	acctcagggtg	atccacccac	ctcgcctccc	aaagtgctgg	27900
gttaataggc	gtgagccact	gcgcccggcc	gagaattgag	ttttctattg	tattaagaaa	27960
cgaactagac	cagggtgcagt	ggctcatacc	tgtaatccca	gcactttggg	aggtcaaggc	28020
aggaggattt	cgtgagccta	ggagtttgag	accagcctgg	gcaacatagt	gagaccctgt	28080
gtctacagaa	aattttaaac	tagccagggtg	tcatgggtatg	tgccctgtgg	cccagctatt	28140

WO 2006/119775

PCT/DK2006/000254

caggaggctg aggtgggaga gttgcttgag cctaggaagt caaggctgca gtgagccatg 28200
 attgcgccac tgcactccag tctggatgac agactcactc ccaaaaacaa aaaaaaaaaat 28260
 ttcttttgacc acctttcaga tcatggattt aagtttctat attttttccc aaatgcatgt 28320
 gatgtcattg taaattttat ggcaagctaa aaaaataatt tgagtttttg aaaatgtttc 28380
 tttggaatgt gtgggctggg gccaacatga attagattgg atcaagtttg gtatcccttc 28440
 ttgaaagaga aggaagtttg gagagcaatg ttaactgaac tgcatttggt acagagagct 28500
 atttagaaat gttttgctac atttcactga aatcttcata cctgtggcta cattattttc 28560
 ttcccagagct gctgtagtta ggagcagtta tacttttggc cagtggaccc taagcctgga 28620
 cttctacagg ttgctgaatt tgaaccagat tttcagttca ttgcataatg tatttttttc 28680
 ttttaattact aggacgcagg attaactaat cttgggaagt agaaggttga gttgtacctt 28740
 gatgaagatg agttcaagct ctatttggtt gtggctctct cctgctctgg ccatgacgtg 28800
 cccccaccc ctcactcaca catcccacag ctgtctgtgt cctcgtgggc tttgtaactg 28860
 tgagggtcct agcagtgaat gtgaatgaaa agtgttcaga gactatcttc tccttgtctt 28920
 tgggtgcctgg gacaaaccta cattttctaa atggttttgc ttttgtgttc tctggaacag 28980
 gcatttatgt atttatattat tttgtttttg agccacagtc tcgttatggt gccagggctg 29040
 ggggtgcagtg gtgcgatctc ggctcacagc aacctccacc tcccaggttc aagtgattca 29100
 cctcagcctc ccaagtagct gggactacag ggatgtgcca tcacacctgg ctaatttttg 29160
 tatcttttagt agagacgggg tttcaccatg ttggccaggc tggctttgaa ctctgacct 29220
 caagtgatct gccgcctca gcctcccaaa gtgctgggat tacagacata aaccactacg 29280
 tccagcctgg aacaggcatt ttaaaatgtg ttctggaggt tttatggctc atgcctgtaa 29340
 tccgagcact ttgggaggcg gaggtgggta gatgatgagg tcaggagatt gagaccatcc 29400
 tggctaacac agtgattaaa ccctgtctct actaaaaata caaaaaatta gccgggtgtg 29460
 gtgggtgcgca cctgtagtcc cagctacttg ggacgctgag gcaggagaat cgcttgaacc 29520
 cgggaggcgg aggttgcagt gagccgagat tgcaccactg cactccagcc tgggtgacag 29580
 agtgagactc ctgtctcaaa aaaaaaaaaa aaaaagaatg aatgtgcccg tcttctgaac 29640
 atgggtgccat ggggtgtgtg gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgttgct 29700
 cctggactca ggggtggggtg tctcaaacca catgctgggt ggccagtgag gtccctcggg 29760
 cagagccacc tgcgtgtagg ggtgtcatct tatctagggg ctggagatga aatttttatc 29820
 tgtgggaagc taagcagccc atataataac agtgatgatc agtgttgggg taccgaaggt 29880
 tagtgctgtt atcacttttt tacacataaa tgaaatggag gctcagagag atagttaaaa 29940
 tatttgttca gggacacatg ggtggtaagt gacagagtca gggtttgaac ccaggtcagt 30000
 cttgggtgtt tgtgtataga ttcagacaca cccccagca agtcatcaaa aatgtcttct 30060

WO 2006/119775

PCT/DK2006/000254

cattgactac	acatcaaggg	tgggtggagt	ggactctaca	ggctgtctga	ggggtgttag	30120
ttagggaagg	cgtcagtggg	aatggagtat	ctatgtatgt	gtgcgtgtag	gtatgtgtga	30180
tttacatcca	ataaacttta	ctcaagaagg	ggatccttagg	ggctttatta	tataggcagt	30240
ggctatcaaa	ggcttttaag	cagaagagaa	aggtgagacc	acatctaggt	ttcgggaagg	30300
tttttttaga	atgggtagaa	aaggatagag	attagatcta	gagagacaag	tgttgatgca	30360
gtagctgctg	ccccttctta	gggtctttgc	acgtgggggt	cccacttcct	ggaatgttct	30420
ttgtcttcag	gtcatagcac	aaaagcccc	tctcagtggg	cactgatgtt	gctctctgcc	30480
ctgaagtagc	cctctgctgt	gtctgaagtc	attgtgttcc	catgtggatt	ttagtgtgta	30540
tcttcacccc	ccattagagt	gtcactgtca	tgcattgtgca	ctcagcagcg	tgcagggccc	30600
caaaaggcct	ccgtgaactg	tggaggaatg	agcccctgaa	gtggcgggtg	gtgagatcga	30660
gcctgtctgc	tggttgccat	gctgggaagg	ggttgtggag	ggagagggga	accaagggtca	30720
tttcttttag	gcgtttgatt	tcgtttggcc	ggatgaatgc	tgcccttgat	tgatactcac	30780
gagtgttttc	tttgggggcc	gctgcgcttt	tccaaacagc	cagggtgctt	accagactcg	30840
ctgaccacac	tgcagcccag	cgtcgtggct	gacgcagtgg	cacggttttg	cttagcttga	30900
attgcttttc	cttttctgtt	ggccaagtcc	tcttttggtt	gtgatcattt	acctctttta	30960
aatctcctag	catagttcct	ggctttcctt	ttcttgtttt	tcccttgcct	aatgcttact	31020
gtttctgtgt	taattctggg	gaaagggaaa	tcaccttgag	ggtgcttggc	taaaactaca	31080
gagcaagggg	atcgttggtt	ggaaagtcaa	gcgtgagcct	cctgacctgg	ggcattggcc	31140
caatcaagaa	gacctcgatg	tttctttgtg	ttgggggtgt	ggggtgaagc	caagtggccg	31200
gtggctctat	agtggccctg	taatggacat	cagacctgca	gtcagcaagg	ggtagggcagg	31260
gtgggcatgg	gaggggtggag	aatcacttca	gttgagagatg	agactgtgga	gaacctgggt	31320
tttgctcagt	gattcctgca	tggggagggg	ttaggcgtct	gggagactta	tcagatttcc	31380
aggacatttt	cagaccttag	ccctccctct	gaatataatg	tgggtgaagag	gcagacagaa	31440
ggttaaaaac	cagctggagc	tggatgacat	gtgaatgtgt	ctgtgtgtgt	gtctgttgga	31500
ggggctgttc	tttgtcctgt	tatctttatg	ttttgcaa	gtttcaaaac	cagggtgggtc	31560
cttctcaggg	gttgggccgag	cgagggctgg	aggcagtgtc	ttcagttgca	cctaggtgct	31620
gtaaagggtga	cactcctgga	cctagacctg	ggcctgcatg	tggagagaag	tgaggccagt	31680
ggggggccac	aacggagcct	tggagtgtgc	ccatctgctg	cccttggtggg	cagaacgtgg	31740
agcccatcag	gaatggcttg	tgggagacag	ccagtctctg	gggtgagcac	cccatccaga	31800
gaggtgcaca	gctgatggag	tctcgctctg	tcgccaagc	tggagtgtctg	gagtgcgggtg	31860
atgagatctc	ggctcactgc	aacctctgcc	tcctgggttt	aagcagttct	tctgcctcag	31920
cccccaagta	gctgggatta	caggcatgtg	ccgccactcc	tggctgagtt	ttgtattttt	31980

WO 2006/119775

PCT/DK2006/000254

agcagagatg gggtttcgca tgttggccag gctggtcttg aaccctgac ctcaagtgat 32040
 cctctcacct tggcctccca aagtgctgag attacaggcg tgagccacca caccagccc 32100
 atacacactt actgatttca ttttttatag aaagcagcca gctttcctga cctttggttg 32160
 ccaaacagaa caagcaaadc cacatatatc catctctagg gtttagttta gaataagaaa 32220
 aactaggggt ggggcctact tttgttttcc aacctctccc cagctgcttt tttcaggtaa 32280
 gtagttgcct cttttggggt ctcatgttta ctaagctggg tgttgatgtg ctgagaagga 32340
 agacctgtgc tcttagtctt ggggatgttt tctgccctcc ttgtgttaag ataactctgt 32400
 aagtgtcaaa ggccacagct ctgcagcctt ttggcacagg tgaggtaaga gaaatttcat 32460
 tttcctttcc ttctggttta ccctgctcct tattcccaga tgccctgaag caccgattaa 32520
 agggagtagt ttttgcacca ggaacctttg cagcaggtgt tgtcaacccg ctggcctggg 32580
 tcagggcagt gcagatcaca gtcctgcttt tctcctggcc atgtggggtt tcacctcatc 32640
 tccagatcca gtctggggag atgcaagatg ggaccagtgt ttactaggt tcacatatca 32700
 acatcaggga agcttcaggg aggccgaggg cctagagcat ttcttccggg tgggaaattg 32760
 agaattctga attctgtact cattattaac ttttcctggg ctttaaacia attatatagt 32820
 tgtttctttt cattgagggc aggaagaatt ttatttaaaa agtgctaagt gctaggtagt 32880
 gggtagtagc tcacacctgt aatcccagca ctgtgggagg ctgggacagg cagattactt 32940
 gagtttggga gtttgagacc agcctgagca acatggcgaa actccgtctc tacaaaaaaa 33000
 cataaaaatt agccagatgt ggtggtgcac acctatgatc ccagctgctc aggaggctga 33060
 agtgggagga ttgcttgagc ccaggagggt gaggctgcag tgagccaaga tttcatcatt 33120
 gcactccagc cttggtgaca gagccagacc ctgtctcaaa aaaaaaaaaa aaaaaaaaaa 33180
 agtgctaagt gctgtgcca tagcctgtgc gtagcttctt cagcccagggt ttacactagt 33240
 ttatttcctg acctgacatg acatctgatc agtggtcatt aagtatctgt ctcagtgat 33300
 ggagactgac caccagtcag ttaactgaag aacaaagaca aagattctta aaaaacaata 33360
 catagtttag ggatgtaaaa aaaatgatag gcatgatggt gtgctcctat agtcccagct 33420
 actcaggagg ctgagggtggg aggatggctt aagcccagga gttcaagacc agcctgggca 33480
 acatagcaag accctatctc ttaaaaacia aaaattgata tattctcata tgtcaatctt 33540
 ttattgtcct tttcttttac tttcagttta ctaattagaa ctgctgggta ttttttttgc 33600
 atatgccaca ccaataattt attgccttca atatctgttt tttttttttt ttttttgaga 33660
 tggagtctca ctctgttgct caggctggag tgcagtggca caatctcggc tctactgcaac 33720
 ctccgccttc tgggttcaag tgatcctcct gcctcagcct cccaagtagc tgggattaga 33780
 agtgtgtgcc accacacca gccagtttct tttttttttt tttttttttg agatggagtc 33840
 tcgctctgtt gcctgggctg gagtgcagtg gcacattgtc gactcactgc agcctccacc 33900

WO 2006/119775

PCT/DK2006/000254

tcctggattt	gagcagttct	cttgcctcag	cctcctgagt	agctgagact	acaggtgtgc	33960
gtcaccacag	ccggctaatt	tttgtatttt	cagtagagac	ggggttttac	catgttggcc	34020
aggctggtct	cgaactcgtg	acttcaggta	atccacccgc	ctcagcttcc	caaagtgctg	34080
gaattacagt	tgtaagccac	tgtgcccggc	ctaatttttg	tatttttagt	agagatgggg	34140
tttcaccatg	ttggccacgc	tggtcttgaa	ctcgtgacct	caagtgatct	gcctgcctaa	34200
gtctcccaag	tgctgggatt	acaggtgtga	gccactgctc	ctagccttgt	tgttgttttt	34260
tttttttttt	ttttttattg	atcattcttt	ggtgtttctc	gcagaggggg	atttggcagg	34320
gtcataggac	aatagtggag	ggaaggtcag	cagataaaca	agtgaacaaa	ggtctctggt	34380
tttcctaggc	agaggaccct	gcagccttcc	gcagtgtttg	tgtccctggg	tacttaagat	34440
tagggagtgg	tgatgactct	taacgagcat	gctgccttca	agcatctggt	taacaaagca	34500
catcttgcac	cgcccttaat	ccatttaacc	ctgagtgaca	cagcacatgt	ttcagagagc	34560
acagggttgg	ggataaggtc	acagatcaac	agatcccaag	gcagaagaat	ttttcttagt	34620
acagaacaaa	atgaaaagtc	tcccatgtct	acttctatcc	acacagacc	ggcaaccatc	34680
cgattttctca	attttttccc	cacccttccc	gccttttctat	tccacaaaac	cgccattgtc	34740
atcatggccc	atccccaatg	agccgctggg	cacacctccc	agacgggggtc	gtggccgggc	34800
agaggggctc	ctcacttccc	agtaggggcg	gccgggcaga	agcgcccctc	acctcccgga	34860
tggggcggct	ggccgggicgg	ggggctgacc	ccccaccac	cctcccggac	ggggcggctg	34920
gccaggcaga	ggggctcctc	acttcccagt	aggggcggcc	gggcagaggc	gcccctcacc	34980
tcctggatag	ggcggctggc	cgggcggggg	gctgtcccc	cacctccctc	ccggacgggg	35040
cggctggccg	ggcagagggg	tcctcacttc	ccagtagggg	cggccgggca	gaggcgcccc	35100
tcacctcccg	gacggggcgg	ccggccggaa	ggggggctga	cccccccacc	tccctcccgg	35160
acggggcggc	tggccgggca	gaggggctcc	tcacttccca	gtaggggcgg	ccgggcagag	35220
gcgcccctca	cctcccggac	ggggcggctg	gccaggcggg	gggctgatcc	ccccacctcc	35280
ctcccggacg	gggcggctgg	ccgggcgggg	ggctgacccc	ccacctccct	cccggactgg	35340
gcggctggcc	gggcgggggg	ctgacccacc	cacctccctc	ctggacgggg	cgactggccg	35400
ggcagagggg	ctcctcactt	cccagtaggg	gcggccgggc	agaggagccc	ctcacctccc	35460
ggacggggcg	gctggccggg	cggggggctg	acccccccac	ctccctcccg	gacggggcgg	35520
ctggccgggc	agaggggtcc	tcacttccca	gtaggggcag	ccgggcagag	gcgcccctca	35580
cctcccggac	ggggcggccg	gccgggcggg	gggctgaccc	ccccacctcc	ctcccggacg	35640
gggcggctgg	ccgaccccc	ccggcgctc	cctcccggac	ggggcggctg	gccgggcaga	35700
ggggctcctc	acttcccagt	aggggcggcc	gggcagaggc	gcccctcacc	tcccggacgg	35760
ggcggctggt	cgggcggggg	gctgaccccc	ccacctccc	tcccggacgg	ggtggctgcc	35820

WO 2006/119775

PCT/DK2006/000254

gggcggagac gctcctcact tcccagacgg ggtggttgcc ggacggaggg gctcctcact	35880
tctcagacgg ggcgggttgcc aggcagaggg tttcctcact tctcagacgg agcggccggg	35940
cagagacgct cctcacctcc cagacagggg tgccggcccag cagaggcgct cctcacatcc	36000
cagacagggc ggtggggcag aggtgctccc cacatctcag acgatgggcg gccgggcaga	36060
gacgttcctc acttcctaga tgggatggcg gcggggaaga ggcgctcctc gcttcctaga	36120
tgggatggcg gccgggcaga gacgctcctc actttccaga ctgggcagcc aggcagaggg	36180
gctcctcata tcccagacga tgggtggcca agcagagacg ctctcactt cccagacggg	36240
gtggcggccg ggcagaggct gcaatctcgg ctctttggga ggccaaggca ggcggctggg	36300
aggtggttgt agcgagccga gatcacgcca ctgcactcca gcctgggcac cattgagcac	36360
tgagtgaacg agactccatc tgcaatcccg gcacctcggg aggccgaggc tggcggatca	36420
ctcgcggtta ggagctggag accagcccgg ccaacacagc gaaaccccgt ctccaccaa	36480
aaaaaacgaa aaccagtcag gcgtggcggc gcgcgcctgc aatcgcaggc actcggcagg	36540
ctgaggcagg agaatcaggc agggaggttg cagtgagccg agatggcagc agtaccgtcc	36600
agctttggct cggcatcaga gggagaccgt ggagggagag ggagaggggg agggggaggg	36660
ggagagggag agggagaggg ttgtttgttt ttttgaggga gggctctctt ctgtcgcca	36720
ggctggagtg tcgtggtgtg atcatgcagc cttgacctcc taatgaccat gacaagctag	36780
tgtttttgat ttttagtaga gacaaggctc tgccatgttg ttcaggctgt tcttgaactc	36840
ttgggctcaa gcagtcctca tgccttggcc tctcaaagtg ctgggattat aggcattgagc	36900
tactgtgtgt ggccattaat taatttaatt aattttctta atgacaattg agaacttaga	36960
agcaacatga atgtggaatt atagatttta caaattttca cccaatactc tatacttggt	37020
ttgccatacc tataatttag ttagcagatt gtttctttct ttcttttttt ttttaattaac	37080
aacttgtttt aatagaaact tgttttaata gttttaatgg aaacaatgat gcgcttattt	37140
cataggttta tatagttaag catgtatgtg tatttatact atatgtgtat tgtatgtatg	37200
tgcgatatgt taataaatgt gtgtatgtaa atttacaatg acaagtacag ttggcataag	37260
cagggggttg cagacaaacg tgtttcagca cactgggcaa catatcagtc tgtagcatgt	37320
ggtcaggatg caagggggat ccgttaatcc attaataact tagggatttt gtcctcagga	37380
atctcccagc ctagaggagg caatgtctgt gtttaaataa ctataacca agttagaaag	37440
cggcagggtg cttagacag agcatatact cccaggagga aggagccaga gaattcatag	37500
agaaggaggt atctgacctg ggccaattac ttctttcttt gttttttgac tgtttttata	37560
tttagttaaa tatctgtata tttctttttt cttttctttt cttttttttt tttttttttt	37620
ttgagagggg gtcttgcttt gttgcctagg ctggagtgca gtggcatgat ctcggtcac	37680
tgcaacctcc acctcccatg ttcaagtgat tcccatgcct cagcctttgg agtagctggg	37740

WO 2006/119775

PCT/DK2006/000254

attacagacg	cctgctacca	cgcttggcta	atTTTTgtat	ttttagtaga	gatgggggttt	37800
caccatgttg	accaggctgg	tctcaaactc	ccgatccttag	gtgatccacc	tgccctctgcc	37860
tcccaaagtg	ctgagattgc	aggcatgagc	caccgcacct	ggccaaatat	ctctatatattt	37920
ctgtaaatct	ataaataaaa	acacaaacag	atgaaccatc	attgggcctt	gggcagtgggt	37980
atggacattt	gaattgtttt	tgaagtttct	atgtctcagc	cagtacaggt	gctggctgtt	38040
tctcctaggc	aatggagctt	tgtagtaccc	ctgttgtctg	tgctatactg	caggggggatt	38100
ggccaacagg	atgcattcct	ttctgtcaca	gttcattata	ctggtcagag	ttctgaggtc	38160
accagctaac	ctgagcataa	tcagttcctt	ggaagacatt	gggtggagtg	acctcataga	38220
ctcataggaa	acctgaaaag	aagccagggg	agaaacaatt	gttttgaagg	agtagggaga	38280
tttcccaaag	cccagctgga	atgctgttgc	caggaggagg	aaccagttaa	ggggcagggg	38340
ccagcctttt	tggaaatgct	ttgtagaatg	ccctagggaa	gaggaaatac	ataaaccaag	38400
ggccaggatg	caaacatact	tttgaatatg	tctcagatct	gagtcacact	gaacgtctgc	38460
aggccaaaaa	gcatagcaca	tcattcatag	gaactccaga	aatgttgccc	gactgcctcc	38520
cacaacttac	cgtcagtgcg	ggcatggaaa	ccttcctgct	ttctgctctg	tgtgggtggg	38580
tgaatgtgag	tgaggagaaa	cactgtcatc	tgggagtgaa	gagcccagct	ttgcagtcac	38640
tcagcaggat	gaggggtggac	ttttggctct	gccaccaacg	cagtgtgggc	cctgtgctct	38700
tcttgtgttg	cgtgtaaacg	aagactgata	ttttctgtta	gatttctcag	cgttataggt	38760
tattggattt	gttactgtgt	gctcatgcca	agagaaacag	aagatgtata	aacagctagg	38820
tggtaatcaa	ctgttgactg	cttgttgtag	cagagccttc	cacaactttc	tcctgaccct	38880
cacaaatgca	tgtttttaaaa	aaagagaggg	acctattgta	ggcattattc	agcaacttgt	38940
tatttttact	tctcttgagc	ttacacaccc	ctgaatttaa	gctcttttga	taacaaataa	39000
ttgaaaaaag	tgcatggcag	gttgtaggtg	ctcattgtgg	gatgaatgaa	tgaatgagcc	39060
agtgagtgag	tgttgtactc	tggatctctc	tccaggtcag	tagtttttag	tagttatgtc	39120
acactgaatt	gctctagtct	ttttcagact	tgtactagca	atgtatgagt	accctgttat	39180
tagttcgttc	tcacattgct	ataaagaacc	acctgagact	gggtgattta	tgaagaaaag	39240
aggtttaatt	gactcagttt	cacaggctta	gcaggaagca	tggctaggag	gcctcaggaa	39300
actacaatca	tggtggaagg	ggaagcaagc	gtgtcttacc	atggcagagc	aggagagaga	39360
cagagagcgt	gaggggggaa	gtgccacatg	cttcagatct	ccgagaacac	acttactgtc	39420
atgagaatag	caagggggaa	attcgcctcc	atgatctaata	cacctcccat	caggtccctc	39480
ccctgccaca	cagggattac	aatttgccgt	gagatttggg	tgaggacaca	gagcctaacc	39540
atatcaaccc	cttttactgg	gcatttccac	cagtgctaga	tgttgtattg	tttttgctaa	39600
tctcatggct	gaaaatggta	ttatagtctt	aattatTTTT	aactaaaaac	tgtttttcac	39660

WO 2006/119775

PCT/DK2006/000254

cattctgttt ctatggagca cttttatttt aatttgtgtt ttcctgagta ctggtgaggt 39720
tgagcaccat ttcacctttt ttctttattt tttcttctgt gaattccata cctatatcct 39780
ttctccgttt ttccgtctac tggattgttt tgaattttta ggaactcttt gtatattagg 39840
gctataaatt atttgttgtg tgcattgcaa atgtgtcccc agctattgtt tctatagttt 39900
gtttttacat tgtttaggtg tgttttgcca taaaatttta agattttttg caaatttcta 39960
tttctatttc tttgtaaaaa tgtttgtttc ttcttccatc catgggtctaa atgaaattac 40020
agatcagtct catatgaatg taatatgtgt ctattaatgc atatatatgt ttatatattt 40080
gcaacatttt aaagtaggtt tctaaatcaa ataataatth gaatacagta gtgtattaaa 40140
agattaatth ttcattgacca aacaaggtht actctagcaa tgcaaggata gcataacatg 40200
aggtattgtg tcggctggag aaacttctag ctcaaggatt agagcatggt ttgccccagt 40260
cctgccatth actaattctc tgacgttagc caaattagth aatctccctg agcttcagtc 40320
tcctcatcta taaaattggc aggacaatta caatactgct gtatagggtt gatggaagga 40380
cgaaatgggg tatctgtgaa atgattagac ctgggttttg tgcataataa gtgctcagth 40440
aatgttatac atcacaattc actgtatcac attaaaaaga aaacccatgt ggthtatgth 40500
gcatattthaa aaggcatgta atacaattct gccgctaatc ttaatgaaaa tctaaaacag 40560
tcatagaatc tacataaata cagaaatagg tgccaacacc cagcagcaag cctgtthaagt 40620
ggccattctg tgacactcca ggcatgacag ggaggcctgc atttagcact gttgcctcaa 40680
atcgatgatc ctgtggactt gtcaggctth tggggagagg agccagtctg tcgcctgggc 40740
tggccagtgt tgcctctgct ctgtctggag caagtgtgca ttaaggcctg cgcttgccag 40800
tggcctthca ccagggtgta gcttctggcc ttgctctata ggaaccagga attcctgggg 40860
ttgggctggg aagaggatcg ttaaagtgat aagcctggag gcattttgag aactacctgt 40920
gacattthaga gtgaggaggg tcagcagtgg gtgggttagc ctggcaggth ccaaggthct 40980
gtactccctc taggtgaaga attagcacag gcatagthtc tagccccctt ttctctctct 41040
acatgcaaaa taagatgatg aatgttctcc aaatgacaga ggaggggaag gcctcaatgc 41100
tgtgaactct gtagggthtc agctcagatt ggacctthca gactggcctg gaaagccatg 41160
ggggccaggg tgtctcagca ttggcccgtt tttggcctgg tatttgthga acatctatca 41220
gtgggtctth ttcttgtgtc tgtcttacag catttctctc atctaccctt agththagaaa 41280
tgththctth gataactthaa aattcatgta aaataagagg tgatggatat ataagththg 41340
gaacatacta ataatgcct gggagccaac ctcttagctg gagaactaaa ccactgtgca 41400
ccacaggga actcctagtg ggtgctccct gccagggggg aaaggcattc tagaatagth 41460
catcattcct gcacctthth tttttththta aagaatattg acaaggctct ggththagaaa 41520
agggaaaaaa aaaagaatac ttatacacat gtatctgthaa ataatatatt ctattththg 41580

WO 2006/119775

PCT/DK2006/000254

ttatgaattt	tataaaaaga	gactcataag	taatctgcaa	atttcatttt	tcattcattt	41640
ttacagctgg	ataatttttc	attgtgtgca	tgctatatat	atatatat	gtatatattt	41700
tttatttttg	tagagacagg	gtcttgcccg	aactcctgag	ctcgagcaat	tctcctgcct	41760
cagcctccca	aagtgctggg	attataggca	tgagccacca	tgcccagccc	ccttttctacc	41820
ttttaataga	tatatatttag	ttttatttgt	cttatgagta	aaccacaaga	atggggactt	41880
tctctaattg	atttctgtat	cctgtgaaag	tactaaaact	tgtggtgaat	atggtgggag	41940
cgtcctccat	gctaataact	gagtattggc	attggattta	agggagtcag	taggttaa	42000
tgatacagat	tggaccttag	agaggggtgat	gtgagtgtca	gaacaagacc	aaaatgccac	42060
ccttgttaca	atgtagtaga	agtgggtggac	tagagctcat	ctgcagtgca	ggcagggtggg	42120
cttgctgact	ctccctgacc	cacattctag	gccctggcac	agtctactca	aaggctcagct	42180
gggtctgcag	aggagatatc	cttgctccag	ctggctgcct	tgcatcaatc	tgggggtggag	42240
aggaaaggaa	aagaggacta	attcccaaag	cttatgttgt	tgttgttggt	ccttctcctc	42300
gtcctcctcc	tcctcttctt	cttcttcttt	ctcttttctt	cctctcctcc	tcctcctcct	42360
ccttcttctc	ttttcttttc	tttttttttt	tttttttgag	acaagatctc	actgtgttgc	42420
ccaggctgga	gtgcagtggc	acaatctcaa	ctcactgcaa	cctccacctc	ccaggtttaa	42480
gcagttacgg	tgccctcagcc	acccgagtag	ctgggattac	aggcacacac	caccatgccc	42540
agctaatttt	tgtattttta	atagaggcag	ggtttcacca	tgttggccat	ggctgggtgct	42600
tatgatcttt	taatggatgg	cctctaccca	gacaagaagc	ctgtatgtgg	tacttctgcc	42660
cctactcacc	aggaagagca	ctctcctcta	ctgttctcca	gtggggagga	gttaggcgtg	42720
gcttagacca	ggagcattga	gggaatggaa	atcattggca	aaggggagct	ggcatccttc	42780
cctactgttt	agaagtagat	catatcta	taacagggtga	atgctctggg	tatggcacct	42840
gctatttggt	gctcatggga	aagggccttg	ggtagacggt	tcaggcgcca	ggcccttggt	42900
ggaactggct	tttgtttttc	attgtaccta	gtcatgggtac	actggatggc	ggataaccgg	42960
aagcagagtc	gacataggga	tacggaggga	aggagactta	catagatttc	aggacaacag	43020
tgagctgcca	ggcaccaggc	cagggtgcttt	atctcatgta	atccacttgg	taacatgtga	43080
ggtgggagat	gctggtcctt	tttatacagg	tgaggaaaca	gaggctctga	ggtaaattga	43140
attgcgattc	cacagctcgt	gggctgtgag	ctggcatttg	tatcaggcta	caagttaagg	43200
ctacgcttct	gttcttggtg	cttggtcatcc	acacctgagc	tatgtgtccc	caaggatgag	43260
gctgccattt	ctctcgtggg	cttttctctt	ttcctctttt	agcatatgct	gtgcccctag	43320
accctaagcc	ctgtgctagg	tgccaaggaa	acaagggcaa	tgagggatga	tggagacggg	43380
agtggaggat	gggcccaggc	ctaagactgg	agtgtagcag	gcctgtttct	gtgtcagtta	43440
caggggagag	cagccaaccc	agtctggggt	tatggggcag	ggggccatgc	gtggcacaag	43500

WO 2006/119775

PCT/DK2006/000254

tgtggggcctt gccacacctg ggcggagaca ggcaggcatg ttggctggag agcttgcttg 43560
 cgagctttgc agctatcaaa tgtgtaaatt ccaaagtgtat tatgtagagt ttctgattga 43620
 atgaaacatt tgaaaaatga atgggggaaag catctggagt ctttttctca atgagaagtt 43680
 aaaggaggaa gcttcagtgg agccttggtg acaaaaaaca gtgtggtaca gattcccagt 43740
 tgcataattc agagcactgg ggaagttttg acctgtgttt tgctcgtcaa ttcctgagca 43800
 agtggaata atgaagttgc tcaggaggga ttgccttggt ggcgagacct gtgacagagt 43860
 ctgtcatttt gaaccttctg ttctggaaat actttttaca gtgtcttctg ctttttttct 43920
 ttttaaagat tgttttaaaa acaccacttc tcgctttctc tccttatctt tcagtaattt 43980
 gttgcagggc attttggata ggaatatctc actgatattc ttgtttgagc tgacgggtcat 44040
 gtcccagatg tgtttgctgc agtatgtggg cagatatatg actgccctgt caaagttctt 44100
 acttccgtat tgaccccagc cataaaagtc atgataaatc tggttcacaa tgaacaatcc 44160
 tcacatgctg ttcgtaatta gaaaatccag gtgtgtgctt ttatttctag aattttgtcc 44220
 tgtctgaaat gtaggaagat gctgttataa agaagcacat tcaaagccat ttaaaagcac 44280
 gggtgagctg ggtgcagtgg catgcatctg taatccaggt acacaggagg ctgaggcggg 44340
 aacatggctt cagcccagga gttgaagtcc agcctgggca acataagaaa aaaaaaagta 44400
 aaagcatggc tgacttttct tagagtgagg cagtagactg tgtgcttcct ggcgtgggtg 44460
 ccaggtcgga aagcccagag gtggaattct cttttggcca ctttggggcc atcctgggtg 44520
 agagctacaa cttcagggtt cctagtgcac agttctgatc atatcttttt ttcctagatg 44580
 tataaggaac tatttttcta aaatgggtgg ggttttaaat atatatatat atatttagta 44640
 ttatttgatg attcttgggt gtttctcgga gagggggatt tggcagggtc ataggacaat 44700
 agtggaggga aggtcagcag ataaacatgt gaacaagggt ctctgggttt cctaggcaga 44760
 ggaccctgcg gccttccgca atgtttatgt ccctgggtac tggagcttag ggagtgggtga 44820
 tgactcttaa agagcatgct gccttcaagc atctgtttta caaagcacat cgtgcaccgc 44880
 ccttaatcca ttttaacctg agtggacaca gcacatgttt cagagagcac gggattgggg 44940
 gcaaggccat agatcaacag cgtcccaagg cagaagaatc tttcctagca cagaacaaaa 45000
 tggagtctcc tatgtctact tctttccaca cagacacagt aacaatccga cttctctttc 45060
 ttttccccac atttcccccc ccattggac aaaaccgcca ttgtcatcat ggcccgttct 45120
 caatgagctg ctgggtacac ctcccagacg ggggtggcggc cgggcagagg ggctcctcac 45180
 ctcccagacg gggcggccgg gcagaggcac cccccacctt ccagacgggg cggcggccgg 45240
 gtggaggctg cccccacct cccgaacggg gctgctggct gggcgggggg tgccccccac 45300
 ctccctcccg gacggggctg ctggccgggc gggggctgcc cccacctcc ctcccggatg 45360
 gggcggctgg ccgggcgggg gctaccccc acctcccgga cggggcggct gccggacaga 45420

WO 2006/119775

PCT/DK2006/000254

gatgctcctc	acttcccaga	cagggcggct	gcccggcgga	ggggctcctc	acttctcaga	45480
cgatgggagg	ccgggcagag	acgctcctca	cttcccagac	atggtggcgg	ttgggcagag	45540
acactcctca	cttcccagac	tgggtggccg	ggcagagggg	ctcctcacat	cccagacgat	45600
gggctgccag	gcagagacgc	tcctcacttc	ccagatgggg	tggcggccgg	gcagaggctg	45660
cactcttggc	actttgggag	gccaaggcag	gcggctggga	ggtggggggt	gtagcgagcc	45720
gagatcacgc	cactgcactc	cagcctgggc	aacatagagc	actgagttag	cgagactccg	45780
tctgcagtcc	tggcacctcg	ggaggctgag	gcgggcagat	cacttgcggt	cgggagctgg	45840
agaccagccc	ggccaacacg	gcgaaacccc	gtctccacca	aaaaatacaa	aaaccagtca	45900
ggcgtggcgg	cgcgcgctg	caatcccagg	cgcttggcag	gctgaggcag	gagaatcagg	45960
cagggaggct	gcagttagcc	gagatggcgg	cagtacagtc	cagctttggc	tcggcatcag	46020
agggagaccg	tggagggaga	gggagaggga	gaatattttt	atcatggtgg	tttgtaatgc	46080
cacctccttt	cattgttttt	agcttagcgt	tagcctgcag	accctcagaa	actgctgagt	46140
ggtgagaggt	ggtgcagggg	taggctgtag	cagagagggc	tgtggagagg	ctgggggcag	46200
ggcaggggtg	tgcagggagc	ctgggtcgct	gtctgaccgc	acagctgtgg	gtggcgtggg	46260
gaggcgctgc	aggccagcct	cccattctcat	ggtcagcact	gcttggttacc	acaagcactc	46320
catatctggg	ggtgcactgc	catggagcac	ctcgggtgac	atcagcttac	acggggccct	46380
gccttccttt	ccagggtgctt	gcttacttaa	agattttaac	aaatcatttg	atttcttttag	46440
gcttcatctt	ccttggctaa	attttaaaaa	gaagaccaca	ttgtccctga	ctcctttttg	46500
caaaggacgt	aatgaatggt	cgttaaactt	ttgtattatt	atttctataa	taaacttttt	46560
attttacaag	aattctatat	gtgcagaaaa	gttgcaaaac	taatagagtt	cctgtctacc	46620
ttttgcctga	ctttctctaa	tgtaaacatt	ttatgtaacc	atagcttatt	ttttgaaatt	46680
aagaaattaa	cattgatgca	gtattactga	ataaactatg	gactttcttt	ggatttcacc	46740
agttttccca	ctgatgtttt	cttgcagttc	taggatccca	tccaggctcct	tgatgatgtt	46800
tcctagtatt	agattgggtg	atgaatttgg	ggagagtccc	acagagggtga	agtgcccttt	46860
tcattacatc	atctcaggag	gcacctgata	ctgccatgac	tcaccaccgt	gatgttaacc	46920
ttgcgggttt	acttttagtcg	ttctcctttt	taaatgtgat	tcagggtggca	ctggttcccc	46980
aaaagattgc	tgatcccttc	aggccaggac	cctgctctgc	ttgctttgtg	cctcttccca	47040
gcgtgaggac	ttgggctgtg	gtgccagtgt	cgaagggggc	atgccaggtc	aggtggagat	47100
gcattttgggt	tctgtgattt	tgtgagcatt	ctacagtttt	cacatcacag	gtgaaacagg	47160
attggccatg	tggtgatact	tggtgaaacg	gggggtcatt	ctagttttctt	ttctttttttt	47220
tttcttttttc	tttttttttt	taagggccag	catctcacc	tgctactcag	gctagagtgc	47280
agtgggtggga	tcatgattca	ttgcagcctc	accctcctga	gtagctggga	ccacagctgt	47340

WO 2006/119775

PCT/DK2006/000254

gcactatcaa gccagctta aaaatttttc gtagagatgg gttctttcca tgttgcccag 47400
 gctggtctcg aacacgtgga ctcgagcgat catcatgcct cggcctctca gagtgctgga 47460
 attacaggca tgaaccactg tgcctgacct agtttctact ttctacgtag aaactacgta 47520
 ggagaaggga aaactatttc caatctacag cgataatcaa tcccatattg cttggataat 47580
 gccggtgtta caatggtcat cgttgccaga gttattgagg tgggaggatt actcatctta 47640
 aatgtatttg gttcttcccc aaggcctcca taagaagaaa aacttgcata taaaccagt 47700
 ttaagtggta ctgctaataa cagggttaata ataatttcc ttaacatctt tcagggagaa 47760
 aaatcactgg tgtgccatca gaaaagaatt ccaattgaaa tgagggtaca gattagttgg 47820
 aatacattta tagaatacaa agaaagttta ccatgtaatg tcatgatcat ttgttgaaaa 47880
 gaaatttgta aatgagata cacttttaag ttatagttg gtattctatt tataattatc 47940
 tccatatttc aaactgagaa tggaagcttt tagaagttag tgaccatgtc ttttaagttg 48000
 tatttatgta acacttgcct agaggcttta gatgtatact tactacgtgg ctgtattagt 48060
 cagagccaat aggttgatg gatggatgga tggatggatg gagatagata gataagaagc 48120
 agtttattag gggaattggc tcacatggtt gtagagtctg agaaattcta tgacaggctg 48180
 cctgcaagct ggagactctg ggatgccggt agtggggctc agtccacatt tgaggggccc 48240
 agaactagag aagccagtca ttagctctc agtccagtgc tgaaggccta agaaccagg 48300
 ggccgctggt ataagtccag gaggccaaag gctggcaagc ttgggggttc gtggcccaag 48360
 gcaggaggtg agtgcacccc agtccaggga cattgataga cacatttgcc cttttcctgt 48420
 ttttgttctt tttaggcccc cggcagattg gatggtgcct gcccttattg atggtggctc 48480
 ttccccacct catccactca cactaatgtg ctagtctctt ctggaaacag ccccatggac 48540
 acaccccaaa atcaggcttt actaggtttc taggttttcc taggtgttcc agtcaagttg 48600
 acacctaata ctctaaaact aacctgaca catgccaagt attgaaaatc tcgaataagg 48660
 gttcaaatgc tacctatttc ttttcccatc ttgccattt tctttttccc cagcactgat 48720
 ccaaaatata gaaggagaag ggagaagata atggtagcgt cctggaatgc ttggcaccac 48780
 gatttcattt tcttacatat ttataccctt cctggtttca ctaagggttt gagttggtta 48840
 ttagggacta tgaaaaaaaa ggtggcagac ttgaactctt cttttgcctt ccccaaaaca 48900
 acaacagtga aacagatggt tagttctggg aggagccgtt gcaatcacc acagctgtct 48960
 tgcagtgtca cgctgtttgt ggaggtccag gacacactcg gccgtcagag ggagccaggg 49020
 ttgcagcttc ccgcgtacc attgctgtgc aactgtgggc aagccacttc cccaaacctg 49080
 tagttttctc atctgaaaac tggcgcagta ctccacctt gcaagggttat catgagtgt 49140
 tatcctgagt atatcattgt atataaagat ctgaggagcg tgtataaaat gtcataaaat 49200
 atgcactcag taaagtacct aatttgatca tccatttgca ggatttcttt ttatttttga 49260

WO 2006/119775

PCT/DK2006/000254

gatggagtct	cgctctgcgg	cccaggctgg	agtgcagtgg	cgtgatctcg	gctcactgca	49320
aactctgcct	cccgggttca	tgccattctc	ctgcctcagc	ctcccaagta	gctgggacta	49380
caggcgccca	ccccacgcc	cggctaattt	tttgtatttt	tagtagagat	ggggtttcac	49440
catgttagcc	aggatggtct	tgatctcctg	acctcgtgat	ccgcccgcct	cggcctccca	49500
aagtgctggg	attacaggcg	tgagccaccg	cgcccggccc	atttgcaggg	tttctaaca	49560
ccaactcaca	atgccagttt	cactatcatg	ctagacgatg	aaacaccaac	agtataatta	49620
ttctgttaca	aagttaaaga	tcttttagttt	aaggcttttc	attgattctg	gggttggtatt	49680
acctgtcttt	tggtgttaaa	atgtctgaaa	tgaaataatg	ggagtgatag	gtctttttct	49740
tttttaaagt	tccttagggg	gcaaaataaa	gttggcgggc	ttgtgttggc	tgcttttcct	49800
gccctcagga	aaaagccctg	atcagtaaca	ttgggatctt	ctatctggtc	cctgctctcc	49860
tttgaggag	gaggaagccc	cagccaaaga	ggggaggcca	cagccagagg	tgagcagga	49920
cacggccctt	gtttacgtga	ttcagctcag	agtccagcct	tcctgggcat	cttctcacat	49980
ctcttcctgg	ccccctgtcc	cagccccacc	ttccacattc	cctgcctagt	ttctgttacg	50040
gagcacctgt	caccctgcat	ggtcctctgg	gtactcgtcc	gtctacccta	ccagacttga	50100
agctgcttac	ggacaagttc	gtttaagtta	attatttagt	ttgtaacttc	cagagcctgg	50160
acagtgctga	aaaatatttg	tcagatggct	gaattaaact	ataaataaat	gaggggaaaa	50220
aatcacacc	gataaaacat	taacataaca	tgcccctgac	tttattcctg	aaaataccca	50280
cagactcacc	aggaaccttg	ctcacattcc	ttttgaattt	cagatctttc	ttggatcatc	50340
ttccattttt	aatgtgttgc	ccagagaagg	ctcaagctgt	tgccaggacc	atgaggctgc	50400
gtgacatctt	aactgagatg	tgctagaaat	ctagatgttg	cttcttctgt	ggcatttggg	50460
gattctataa	aatgcataag	aaaaaaatta	acatgttccc	tctatcacat	tggtgctttc	50520
aagtttttcc	tcaattgggc	tgtctccatt	aattagtctc	acttcactgt	taaaaaaatc	50580
tttggacaag	tgccagagaa	gcaaagaagc	ttatgtaaag	ggttttaaca	ttttctcctt	50640
tgccaaagat	agaactgata	cttagctttg	ctttgtcctg	ggaccttctt	ctttgataga	50700
gaatgttaga	catttggtct	ggttttattcc	tttttatgta	cacgtgttgg	atgagtaaag	50760
acaagtggct	tttttttttt	tttagatgca	gtctcactct	gtcgccaggc	cggagtgcag	50820
gggcacgatc	tcattgcaac	ctccgcctcc	caggttcaag	caattctcct	gcctcagcct	50880
ccaagtagc	tgggactaca	ggtgcacacc	accacaccct	gctaattttt	gtattttttca	50940
aagagacggg	gtttcaccat	tttggccaag	atggtctcga	tctcttgacc	tcgtgatccg	51000
cctgcctcag	cctcccaaag	tgctgggatt	acaggcgtga	gccaccgcgc	ccagcccaag	51060
atggctctat	ttacatcttg	gttcatgaaa	catatgaaag	gctaacaagt	aggatttgca	51120
aatagccact	accctgttct	cagccaaaag	aaaaatatga	aaaagattga	cctatcacat	51180

WO 2006/119775

PCT/DK2006/000254

ttccacataa tttatattatg attatattatt tatttatgat tatttatatta tttatgatta 51240
 tagtagggttt ctagaattta agatttactt atttgaagag agaaactaat gccttaaatt 51300
 atcatattaa attttactgg tttactgtct agcagccatt gaaacctatt gctgacacct 51360
 gccacttcca ggctactgaa ctgtcattgt catgcctttc atcgcggtgc tagtggaag 51420
 aaccaagct agcctggta gaaagtggac catttggatc ctctaccaa gtcattggta 51480
 gaggatcagt tattccaaga attgaggccc aagaaataca cttgggctaa aatgacacag 51540
 aacttacttt aatcttttga atctgaggca ctttgtggtt cttgggttaat gtcagtcacc 51600
 actgtattgc agatgcagac gctgggtagg aaggaagctc tggtactaga cttctgctgg 51660
 gggcttttct tccttttctc ctgttttctc accttctttt cccctccttc tcccgccttc 51720
 ttcttcttta ttgaggtatt agtttacata caattgaatg catcctttta ggagctttga 51780
 caagtggaaa cacctgtatg actcagtcaa gacagaatat ttccattgct ccaaaatttc 51840
 cttgtgtccc tttgaagtca gtccccacc ctgccccctg gccctggcaa acattaatct 51900
 gtcactgtag attggttttg cttggtttcg aatttcattc acatagacac acacagtggg 51960
 tgctattttg catctggctg ctttgctcag catgatgttt ttgtgattga tccatgtcac 52020
 atctgtcagt agctcattcc tttttattgc cgagtgggtt tcagctgtgc gccgatgcca 52080
 cacacagttg ggttttgatg tcctctatag agatgtgtga caaacctagg tgctacagga 52140
 gtgtgtgcat gccctcccca tctttcagca gcatatgagg aaggcagagg gaactagcaa 52200
 ataaatgaac cagatttggg ggaatcttct ttttatccaa gccatgggtc ttgctagtcc 52260
 ttgactagtt tggcttctca gaacaacact attctagtta cccctggctt gggccagcct 52320
 gtaggatctc ggtgtttgcc tttctggaga tggggagtta caaattccta atggaaaagg 52380
 cttgtggtca gagcagctcc ttttggtctt ttgtctaact ttgtcgtccc tgctgccgat 52440
 ggatcacctg gctctgtgtt tggggcatgc aaagttcagg gaccaggtta cataggaact 52500
 gtagagcctg ccctgtgcca ggctgccagg taccacacct cctatgccag ccacctgcca 52560
 ctcaggtctc ctgccttgcc cttgccctgt cctggctcag ctgttagttt tagcctgcta 52620
 gggaaagact gaggcgagcg cagaactgca atctgtggta ttctattaac aaccaatta 52680
 ttgtgtgccc atgatcactt taagaatcac cttttttttt tttttttcca gacggagtct 52740
 cgctctgtcg cccaggctgg agtgtaatgg cacgatcttg gctcactgca acctccacct 52800
 cctgggttca agcaattctc ctgcctcagc ctctcgagta gctgagatta caggcatgtg 52860
 ccaccatgcc cagctaattt ttgtactttt agtggagaca gggtttcact atgttggcca 52920
 ggctggtctc caactcctga cctcaagtga tctgcgcacc ttggccccac aaagtactgg 52980
 gattacaggc atgggccacc gtgcccagcc aagaatcacc ttgtaaaaca caatgactgc 53040
 cgcgggctgg ggaaagggtt gggaaatagag ggatgatagc taaagggtac agggtttctc 53100

WO 2006/119775

PCT/DK2006/000254

tttgaggtaa	tgaaaatgtc	ctacccttga	cgggtggtgat	ggttgcacac	acctgtgaat	53160
atactaaata	acattgaatt	gtgcccttta	aatgggtgaa	ctgtgtgacg	tgaattatat	53220
ctcaataaag	ctgttttttt	tttaaggcag	tatagtatta	gagcgatgac	agaccagtag	53280
agaagtgggtg	aaaaaaaaat	tgtagaaatt	ttattgtttc	caaatgacac	gacctttttgc	53340
caaatatctt	gatttttcata	gcagggtccc	ttccatggga	ggcacgggtg	tggtagagag	53400
aggagggact	ttgtatcaag	tggacctgac	tgaatcccat	tctgccctta	ttagctgggt	53460
gaccttgggc	agattgctta	tgttctgtgg	tggtttcctt	ccctgtaaaa	cagggttagg	53520
agtacctatc	ttatagtgtt	tcaaagtgagc	gtggtaaacg	tggtacaagc	agtcattttt	53580
accatgcgac	aatcatataa	caaaaggcct	cagatagaag	cagcagcctt	gctggatcct	53640
gtcctgtctt	tgagtcctgc	taggctgatg	aactggcctg	ctggagctgc	actgaccgct	53700
gggcgaaaga	atccccctgg	gaagtgctgc	cctgcgtgtt	gtttgccgcc	tctgcagggg	53760
gcctgagctg	atgggtggcg	ctgctggtct	ttagttaact	tggtggactc	tctgtttctg	53820
aattttcttc	cgccttaatc	cacatggttt	gagtaggctg	gttgcacact	cttgagagac	53880
cagcacgaga	tgtccctcat	tttattttatt	tgcacttctg	ctcagcctgt	acattgtggg	53940
ccggggcagg	gaccttggct	gctttgactg	tgggacccta	agcactgtgc	ctggcattaa	54000
tgtttgttga	atgaatgtcc	aatggtttcc	agtctttgat	gacagaagcc	agatcagaac	54060
tggaattcca	gcttaggctc	tgcatgccag	aggccagcag	gtgagcttcc	ctgagcgggt	54120
tctgtcccag	catcacctaa	agagtggctg	ggcagctgta	tcactgccag	ggacaggaga	54180
gctgcctgag	tcagtcttgt	cttctgccat	agtccacggt	tcctgcctgg	aatccactag	54240
tgagatgtgg	ctgacatggg	caggggagagg	gtgtcttttag	ccacactgga	gatctgggtct	54300
gatgacttcc	tagtcccact	actattcagt	tcaaagtcag	ttttcttttt	ctgcgccgcc	54360
cccccccccc	gccacacact	tgctcttgtg	actgggtact	tggggcagga	ctcacaatgc	54420
ctccacccac	cccaggcatc	caggctgcac	tttgcctgag	ggcatgtgag	ggtggccctg	54480
tgacttggac	tgtgaatggc	tgctgggctg	gccttccctt	gccctgtgct	ggatgtggcc	54540
ctcaccccct	acccagacc	tcaccctcca	ccccccaccc	ccagttctgc	cttgggcctc	54600
ttctcctttt	ccaagttcag	ctttttatta	agcagaaact	atccaagacc	cagttctgac	54660
cgttgaggga	agtcaagttg	gaggggcagt	gttgtgtagc	ctggcttaca	ggctggctct	54720
ccttgcgcca	gctgtgggac	acttagcaag	ttgcataacc	tctcatctgg	gaaatgaaga	54780
taacaatgct	tacccccaaa	gagctatttt	catgattagc	cagaaataaa	gtaagtgctg	54840
tgtaaattgtt	tattgagcag	cctgtgtttg	ttggatttaa	aagaagagag	tgagtctttt	54900
aaacaaaaag	ccctcccata	accagcagcc	ctaagtgggg	ctctgggaca	aaaaaaaaaa	54960
aagttttctg	gtatcctggc	ttgggggagg	gaacaggtag	acaatttttg	tatttttttag	55020

WO 2006/119775

PCT/DK2006/000254

tacagacggg	gtttcaccat	gttggccagg	ctggtctcaa	acttgcttcc	ccagaactca	55080
ggcgctgctt	ctgtcttagt	ggagtgagt	cttgccagct	gtggcgggcg	tggcactctg	55140
gtgttcatct	tctgtgatgg	tgccgtggat	ttaactaagc	caaccacagg	atatagtga	55200
atatgtgttg	acttctctgc	ttttttggtg	acgggctgtg	tggggagggg	tggtaaataa	55260
tttctcaagt	agtgtaaatt	attatcatat	gtcaagtaat	agagtctcat	gttatgtggt	55320
ctttgctttt	attggtattt	ttccattaag	gaatttcttt	ttctcttttt	tttgagacag	55380
gatctcgctc	tgtcacccag	gctagagtgc	agtgacgtga	tctcagctca	ctgcaacctc	55440
tgctccctgg	ggtcaagcga	ttctcgtgcc	tcagcctcct	gagtagctgg	gattacaggc	55500
atgtgccact	acacctggcc	aatttttgta	tttttttagta	gagatgaggt	ttcaccatat	55560
tggctaggct	ggtctcaaac	tcctggcttc	aagcaatcca	tttgcctcag	cctaccaaaag	55620
tgctgggact	acaagtgtga	gccaccggcc	cagcctaagg	aatttctaaa	cacaagattt	55680
tttttttttt	ttttttttaa	ttgatcattc	ttgggtgttt	ctcacagagg	gggatttggc	55740
agggtcatag	gacaatagtg	gaggggaaggt	caacagataa	acaagtgaac	aaaggtctct	55800
ggttttccta	ggcagaggac	cctgcggcct	tccgcagtgt	ttgtctaaac	acaagatttt	55860
caaaagagat	ttttaaccta	cttcgttttg	tgttataagg	taagactaga	tggctgtttt	55920
taactgaaca	tttttgaatg	aatatgagta	gcaagaacac	tggtaaatag	tgtatatcca	55980
cttagcacag	tttccccttg	cagtatgttc	ttttgccaat	tacgtatttt	cccatcccta	56040
catacacccg	cttgggcat	gttgaaagga	ggaaccgccc	tgaactggcc	ctgccctggc	56100
cattaccact	tagatcctgg	cttgccaagg	ggaatggggg	agtgtgcatg	tgtatcttct	56160
aggagggcac	ctagcagaaa	ttgatccctt	gaaatatattg	tttcctgtct	gtttttgcag	56220
gttggtttcc	atatctactc	tttcctgtag	tgtctgtatt	ttttgttact	gtatgagatt	56280
ttaaaaagtg	ttattgaata	atatgcatac	ctaaagtgca	cagatctcaa	gtgtacatat	56340
cagtgaattt	tcacaagccc	agcacaccca	tagccagttt	aagagccaca	gcacagctcc	56400
aaaaccctcc	tggcaccctt	tcctgcctgc	ctcctctttc	tcttccctgt	cccggcttcg	56460
aagggtgtag	attagttacg	ccttttcttg	tgctggatcc	aaagggaatc	gtacagcatg	56520
tattcttgta	ggcacaagca	ttttaaaatc	tttctctcct	cacctaaagc	tttattgttt	56580
ttttcttcac	tttaaattct	ttgtaaaaat	tgaagaagcg	aggtgttttt	atttccatgt	56640
aatcacagtt	gcgtatttga	gtgtcctagt	tcttaggaga	cagtcgtgct	gctgagagac	56700
catgggggtca	cctgcctgtc	tccaggggac	tctgtgtccc	cagtcctgtc	tcctgcatga	56760
ggcttcctgt	tcatctgctg	acaggcattg	ccttgttgtg	tgtagtcaca	gaaggaagtt	56820
gacttccaca	tgtagggaca	gtgtaagatt	atcctgtgca	agcaagtccc	tgtcccactg	56880
actgctgtga	cgcccctcag	atgcagagag	agcatgctgt	cttttttctc	ttgtaagcag	56940

WO 2006/119775

PCT/DK2006/000254

gaggatccca	aatacaagaa	gatattagaa	gactcatgta	tggtgtaagt	tagtttgtat	57000
ataatttgat	tttttgtgtg	ttggggaggg	ggtagtatga	attttgttct	cagacctctt	57060
ctccctgatt	ttccgtggaa	atactgaaat	ttctcttttg	ccactttctg	aaaagagagc	57120
tcttctgtga	tggccacagc	aggcaggagc	gtgcagtctg	ggaacctggg	actgtacagc	57180
gaatgcagca	gtgtatggga	atgaccttat	ctcggcagcc	aacaggaaat	gatgaaatgt	57240
gctgggagaa	ttcagctgag	agacagaggt	cagacatgcc	agcagggttg	gaaagaggca	57300
ggtacggaag	ctgagctgca	gagttagccg	gcattctgaca	gaaaaatgtg	agtgtgctgt	57360
ggaaggcttg	catttccttt	aaggtagacc	tctttgaaag	tagaattttc	acggaattgt	57420
tttctttctc	actgagtgac	tgatttaaca	gaactctcac	cttctatgca	agcgcaccta	57480
gttttagagc	tccgtgacgg	caatgttcaa	ttagttcatt	attaactgag	caatgcagaa	57540
tgctttatcc	agggtgcttg	gagaagatgt	cagagtaagt	gtttgatttt	cttgagaaat	57600
agatttcggc	ttaaacaaat	gaaggattgt	gttttgtaat	ttttgactct	ggtttttagaa	57660
catacgtgcg	tgtttccttt	caaaatagat	gcaataactt	tgcacctttg	ctttggcatg	57720
tttttagagta	gctgacatac	tttagctgta	tttttttctt	accgagtagc	taaatgagtt	57780
acttgagtta	gctgttaata	gagggcctttt	ccccctactt	taaggcgcta	atgtttatta	57840
catgctgcca	ttttctttct	ttacagccaa	atctaatact	caaacacata	agtgggactt	57900
gagagagaaa	gcaaattatc	ttttccagga	gactgtaatt	gctgtttcat	gactgtttac	57960
tgcagagtct	ctgatcacct	gccaccgctc	caggaccctt	cccgaccacc	ttcccatcct	58020
ccccctcagt	tcttctcatt	tctcatccca	gctggaccct	gtgccctcgt	ccccctctga	58080
gggctgtcac	tagacactgt	ttatatTTTT	gtcgggtaac	aaaatgtaaa	agtcttattg	58140
ttttaataac	tgagcagcct	tatgtacgta	ttcaattgtt	tcttagaagt	catattctgg	58200
aatttttgaa	aaacagtggc	tcccctttat	tgctgtgggtg	ctgtaattct	aattagattt	58260
ttaaaatcct	gatacttaag	tgacccatca	ctcagtgaat	tatacttact	ggattactga	58320
ttgccattaa	aaaggcagca	tcctctctcc	gtcctccaga	aaatactgtg	ctggtgcagc	58380
ctgtgcctca	cagaaaaaca	tgcttttcttg	ttaggatgtt	acctactgtg	aagtgtgtgt	58440
ttcaaatact	tttaagattg	cataggaaag	gtttctagaa	acccccaagg	tatatagtat	58500
gtctgtccga	ccttaatcgg	gcccatcagt	tttgtctttg	ctcttgtagc	tctgaatatc	58560
tctttcttca	gatcccttcc	ttcttgtcag	taaaacatcc	tactgaattc	tgcatgaatt	58620
ttaaatgctg	tatttatgaa	aacagcaaag	aggtttcata	acttcaaacc	aaagcatttt	58680
tacaggatatg	agatttatgc	caataagtaa	aaatacgcct	tccttttttc	tttttaaaaa	58740
gtgttagtcc	tttgagaccc	gtttgcccac	ccacacaatt	ctgttattca	tttggtggca	58800
taatgcagaa	ggtgggaacg	caatatttag	agctgtaatg	gctcttggac	ttaatattcc	58860

WO 2006/119775

PCT/DK2006/000254

aggatcctca ctttaccaaa ggaattggga tccagaggtg acgacttgct caaggctact 58920
 cgtctcgttg gtagccgagt gagtctgaaa ccctgggtcat gagactcatt actctttcca 58980
 ttacagtaat ctgctttctt cgcattatta tcttaagatg aatatcactt agctaagagt 59040
 gttttttttt tctactgatg agtccgttca tatttgtttc actttcaaac actcagcttt 59100
 tattcaataa cttttttgct gtaagtacca ttgcataga tgtatagtga attacctcaa 59160
 gtacaaagag cattgacatt tcttaaaaat caacaattct ctatattata aagttttgct 59220
 tgcctctccc ctagctcagg ggagagacag tcaattgtta ctgatagaag cgggcctcac 59280
 gatgggattg cttgccagat gagaatgcca aacttaatac aatcacaact aatacaagaa 59340
 acagacaagc actcaattct tctttgaaac aggattcttc aggggtgtgaa agggaaaccc 59400
 caaaccttta gcatttctgc ctttcagccc ctataaggga atactgactc tgtcttgagc 59460
 tgaactgaag aggtattgga atatgaaaga ctttaaaata atacaacata atataaataa 59520
 aggagaaaac gatttaattt ataaaaatgg aatttactca caattttcca gaatcattaa 59580
 aaatgggtctc tgctaaaaat tcagtttgaa tcctctgaaa ttcaggtgtg ggagttagtt 59640
 ataaacaggt atgctgagaa cagtgggtga ctgttcctag acctcggtgt ggccatcact 59700
 gtctgtccaa agtcacctta aaataccttg gggagggttca caggatagcg cttatttgta 59760
 ttttctaaag tatttagaat atatgccttg aagaatttct gaaacaaaag cacttcagga 59820
 aaactggcac tttcataaat atgttaattg tatttgaagt tgaatttaat aactttacaa 59880
 ctatttttaat ttatgaaact gaagaatgaa gtggaatata taagcaatat ttttatgcag 59940
 tgtgtcagaa ctgtcttccc tcctgggtgtt ctagtgggca ctttctgtca gtctccttaa 60000
 agaggtgctg ctttcatttg gtgagccagg cgacattatc ctagagaaga agtggggctg 60060
 caggctcagg agccgcagtg gcttagttcg cttctcccat aatttagtta cttttgttcc 60120
 ccaattagca ttttatactt actattatat tttcactata tattaagtat agaattattcc 60180
 aatcggcttc cttttcttag tgtataattg accttcttat cttctgcagt ctataagtaa 60240
 catttgatat gcttccctcc ctggcctcta gcttctgccc acattagtat aaatgttgct 60300
 gtgcttgtga gatttaaaga tgagaaagag ataccagact gttgcagtgc tggcattgct 60360
 ccaaactgag aaggggttaa acggcagccc cctggggctg ctggttggtt agaaatctca 60420
 gagtttgtgt tgttggtgtt tgggcctttt tcatgtccac ctctgtccag ggcagctctc 60480
 cagaccagat tctagaagtg gcctcagaac cttctctaag actatgaaga aacatcccc 60540
 agtatcctag gcccttagag tttggtaaag aggagcatgt atgtgtgtgt cagtatggtg 60600
 agatagtaat tagacaaggg ttattgattt gaatgaatga atgaatgaat gacattgttt 60660
 gcttttccat agtgattact tttgaaggca ctatagtcac atttcactag gtggtctaaa 60720
 tgcatttgca tgtcacagtt ctgccttgac accaagtaac aggttgtgtc tacctgacgt 60780

WO 2006/119775

PCT/DK2006/000254

ctcctccagc cctggggctc tggcctaagt tttttcttat tagactctgt catatatgaa 60840
 tcataagaca taaatcttta tgtgggctca tgggtacatat tttattttga gaaacttaaa 60900
 gaatctaata ggaagttaga attaatttga aatgtgtgcc agtgggtcatc acaatagtat 60960
 ggcattttaga ttatcttttt gttaatgtga aggggtgctat gaatttcata ctgagtcagt 61020
 tttcccatca ggatttggtc tacagtggag tccttgaagg agaacaagga agaatgattt 61080
 tgctggaaaa tgcttacttg acgaaatggt gatgtaatta acaggacaaa atgttgaact 61140
 tccaatttg gaactctctc ttgactctt aaattcctaa catagattgg tagtgctagt 61200
 ttaacaaaat aatgaagca catgacagtt ttggggataa tgaaatgact tctctatctt 61260
 ctctgctatg aactgacaca tgaagaaaaa atagcacagg gcgtagtaga ctgatcacca 61320
 gtatctattg atttggaaga cagaaacatt tgttactact ttcccatcta tagggacggt 61380
 cacctcaggt tcacttagtc agatcctacc attactgtgc aggaagaatg ttagaaagca 61440
 agactgtaaa gccatgaagg aaagtggcaa gaccactttt ggtggataaa tctttcatac 61500
 ttcacattga actcgaagac attaaattga atcttaggtg taccagtgga acttctgagg 61560
 ttgaagagtt tgcatttggg atggtggaag ttgtcatgat gatgaaaagt catgatgtga 61620
 ggctatgtaa actgtacttg tcccagcacc tgacagcctt tttttttttc cccctgatgc 61680
 tatctccagg tatgacagct ttatttggtg gacatgcagg cacctgagtg agctctgggc 61740
 ttatttccag ccatccgtta gatcactgcc tggagggaca gttcacctgg atgctggctg 61800
 gctggatcag ttgatttggt ttttcgttcg ttcgttcggt tgtttgttca ttcattcatt 61860
 cattcattca gtttaccttc ttggtctctt gtccacgcag ttcacatcag gcactgtgct 61920
 agactctagg gatccataga cggttaggat tgagttctgg tctctgagga cctcacagtc 61980
 cagtaaggta aagagacttc taagcagcag ggtggtgagt cttaaagagc tattgggacc 62040
 ttattctgcc atgaagaatg gaggagattg gggggaaact ccagtttgat gatgagtttt 62100
 agagggagtg gtttgtcagg tacccaagag aagcagcggg agcaatatat gcaaaggcaa 62160
 ggaggtcaga atctgccaga cttaggagct tgttgtggct ggagtgccgt ttatatgtgg 62220
 acatgcttca ggaacatcat atttatcaga tccaaaccag tttgttctta attctatttt 62280
 tctctttgtg gcattatcac ctaccatgcc tccgaaattt gagactttga atccctgccc 62340
 ccatcccca ttccctggca tcatccatca ctatgtcctg tctgtttttc ctcttgata 62400
 tattcctctt tcctggtctt ctaaccattt ccctttctcc agacttgaac ctccagaact 62460
 gcgatgtcca gtgtggtagc cactaaccac gtgactgttt acattaaaat tgactaaaat 62520
 taaaaagtca ttgagaccag tctggacaac atagtgagtc ctctgtctct ataaaaaata 62580
 aaaaattagc caggtgcagt ggcacatact tgtagtccaa gcttcagaag gctcaagtgg 62640
 gaggatcgct tacacccagg aggttgaggc tgcaatgagc tgtgatagtg ccactgcact 62700

WO 2006/119775

PCT/DK2006/000254

cagcctgaat	gacagagggga	caccctgtct	caaaaaaaaa	agtcagtttc	tcacttggac	62760
taactacttt	ttaagtgtta	atagctgggtg	gctgccatac	tggacagccc	aagactagag	62820
gctcaatggg	ctgttctcca	ctctctgtcc	aagggaacct	tccttttatgt	gcttttttgct	62880
ttcaagatag	ggctcttgccc	agtcgcctag	gctggagtgc	agtggcatga	tctcggctca	62940
ctgcaacctc	cgcctcccag	gttcaagtga	ttctcttgcc	tcagcctccc	aagtagctgg	63000
gactataggc	acatgtcacc	atgcctggct	aatttttttga	gtttttttggt	agagacaggg	63060
tttcactatg	ttggctaggc	tggtttctgaa	ctcctgacct	caagtgatct	gcctgcctca	63120
gcctcccaaa	gtgctaggat	tacaggcatg	agccaccgca	cccagcctca	tgcctttttaa	63180
cctttgtact	tgctgttctt	attctctcca	gaataccatt	cttgctctaa	gcattcctct	63240
cttccctggc	tttccttagg	tgaatttcgt	ctctttttttc	ctccagtgtc	cttgtaatca	63300
cacaatcacg	cttgcctctg	agagggcaca	tatatgttta	tcgtgtgtcg	agacatcata	63360
gtatagtggg	taagccttgg	ggacagattg	tcttaatttc	aaaactagcc	ggttacttgc	63420
caggtggctt	tgaacacatt	acttaacctc	tcggagtttt	agtttcttca	gctataaaat	63480
gaagacaatg	gttttttctt	cgtgttaaata	gaaaccctta	gcactgcatt	taaatatata	63540
gctctcaaata	tgtagctgag	agctctcatt	ctagataaat	tcccttcctc	tgttcatcag	63600
ccatgctaaa	accattttaat	ggtttggttaa	tgctaataat	gctgttacgt	ttcttctttt	63660
ctgattgtta	caccttttag	ttatttttct	tattgcagtg	gctaggccct	tcaatgcaat	63720
ttgcatagac	atgtaatttt	ttattcctgt	aatttctatt	ggctgtagtt	cttctgggtga	63780
gattttttatt	ctctcttttg	ttttgattgc	atttcatcat	gcccattttc	cattgcattt	63840
ttttttcact	cttagtttag	aaattaaatt	ctcagcctgc	tccttttagtg	gttatcctag	63900
atatttttgt	gtacttaaca	aaatcttaag	tgatatcttt	acccttctct	tgaacaatgg	63960
aaggatagaa	aacatttttaa	ctccatatgc	cctctcaact	tgtccagttt	tttttttctg	64020
attttttttaa	atcctgcaaa	tttgccattg	tctttttatac	agctaatagct	atttagattt	64080
atccaatata	tttatctgca	tatttgctta	ccattccttc	ttgtgcttca	gggctttcca	64140
tctgaaataa	tttccttctt	ttaaaagtat	atcctttaag	ggctgggcat	ggtgggttcac	64200
gcctgtaatc	ccaacactgt	gggaggccga	ggccagtgga	tcacttgagg	tcaggagttc	64260
aagaccagcc	tgtccaatat	ggagaaaccc	cgtctctact	aaaaaacaca	caaaaattag	64320
ctgggcctgg	tggcacacac	ctgtaatccc	agctactcgg	gaggctgagg	caggagaatt	64380
gcttgaacct	gggaggcaga	gattgtagtg	agccgagatc	gctccactgc	actccagccg	64440
gggcgacaga	gcgagactcc	atctcaaaaa	aaaaaaaaata	aataaatagg	ccgggtgtgg	64500
tggctcaacg	tctataatcc	cagcactttg	ggaggccaag	gtgggcggat	cacgaggtca	64560
ggagactgag	accatcctgg	ccaatgtgaa	aaccctgtct	tactaaaaat	acaaaacaaa	64620

WO 2006/119775

PCT/DK2006/000254

ttagccaggc	gtggtggtgc	gcgccctaaag	tcccagctaa	tcaggaggct	gaggcaggag	64680
aatcgcttga	accaaggagg	cagaggctgc	tgcagtgagc	caagatcatg	ccactgcact	64740
ccagcctggg	tgacagaatg	agactctgtc	tcaaaaaaaaa	gataaactta	aaaataaaaa	64800
ataaataaat	aaaataaaat	aataaataaa	taaataaatt	tcttcaacag	gagagtcatt	64860
aagataacctg	gtttgctgtg	ctgctgaaaa	cagaattatg	ctctgtcatt	ccttgactca	64920
ttcattctag	ttgtgtttat	tatgtgcctg	ttacgtatca	gtggtaggaa	cacgggcaaa	64980
caaggcaaaa	tcctggtgca	taaggaacct	acattctcgt	gagtggagag	agaaaagata	65040
at ttggcagg	tgaggttgag	tg ttaagaaa	ataaaacaga	gttagggctg	gagtatcagg	65100
gctgggggaa	ggcttttcta	agaagagggg	aatgtttgag	caaatatctg	atgctgggag	65160
gcagcagatc	ctacagatgt	gggaggcagg	tattcccggc	agaggggagg	cagggtgcaa	65220
ggtaagccga	gtctggtgca	tctggagAAC	cacatggagg	tcactgtagc	cagagtccag	65280
ggtgtgggga	ttggaaggag	cgagactggc	gctggaagca	ggggctagaa	tcacatggga	65340
ccttgtaatc	agggtgggga	gtatcaataa	ctgtgacagt	tgacaatagc	tg ttggcaag	65400
at ttgagttt	agtagattga	accagggtaa	ctgagacaga	ggtgggtgtga	agcggccaca	65460
ttctggctct	gttttgaggt	agaggtgatg	ttatgtggac	atgagaggag	gagaggagtc	65520
agggatgaca	ctggggcatt	tgtcctgagc	agctgtgtcc	tggaggagcc	caaccggggg	65580
atgcaggagc	aagaggtggg	ttttggccag	tgtgtcggag	aggcccgtga	gcctcccagg	65640
gggagatgct	gagccagcag	ctcggggccat	gccggatctg	gagctcaggg	agaggctgag	65700
tgaggaggca	tcagcacgtg	gatggaattt	taggtcacag	gagggagagg	gtcaccagg	65760
aaactagtgt	gcacagagaa	gaggaggctg	agccctgggc	accacgggtg	tgcacagggt	65820
gataagggca	ccaattgggt	gctttgcagg	actgtccgga	attacttgtc	tagagagaat	65880
gcagttaata	ctgcaggccg	gttatcgtga	ctagaagata	aggctcgtgg	attcctttac	65940
cagaaaagct	tggatttcat	gcttacctaa	agggaagggt	cccacgtgct	gccttatctt	66000
cctttttgaat	tttcaaaagc	aactcaatca	gtgtattctt	atggcagAAC	aattaagtga	66060
tacagatatg	caaaagccca	gactttcatg	ttccaccacc	cagagatgtt	cattcaagac	66120
tcaggacaac	cttccctgac	caccatatgt	ccaggtagac	cctgtcactc	tctgtcccc	66180
acttggttc	atgctttccc	ccagccctta	atgctgtctg	ccgctacatt	acctactgg	66240
gtggctttga	aatcccctga	cccctactgg	aatatataca	ggggccttgt	cacttctacc	66300
catcccggga	ccccacaggc	ccagcgagg	gcctggcata	cggtaagtaa	ggctgcatta	66360
catttcattg	aatggaagaa	ttcacatttt	tggtacatag	ccatctggac	tttttttcta	66420
ggtaattacc	cactcttaaa	catatatatt	tagataaagc	ccaagtgttc	agcttttttag	66480
acgtaagaat	gctacttatg	gcgttgaatg	ttactgtaag	ctaactttgc	actagttggc	66540

WO 2006/119775

PCT/DK2006/000254

tggggctgct	gttttcagtc	agtttcagca	atggtggctg	cacaggcctt	cagagccctt	66600
cctgagcctg	tgcttgaca	gcagcgtgaa	tcgtcagtcc	acctcatgct	gtctcctagg	66660
aggctgtgca	gctgcaggcc	ttctttcctc	aaccaaagt	ctgtgcctgc	ccacctgccc	66720
actcccttcc	ctccctccct	ctcagcctaa	tctggtatga	ctttgtctta	tctttcctag	66780
tatttagaca	atttgttcct	ttgtatcttg	gaggattggt	ccaggaccct	ccaaaggata	66840
ccaaaatcca	tggatgctcg	tctctgatat	aaaatagtgt	tgtatttgca	tataacttat	66900
gtatactctt	cagtattctt	taaatcatct	ctagattact	tataatacct	aatagaatat	66960
aaatacaatg	taaatagtta	taccgtattg	tttagggaa	aatgacaagc	aaaaaaagtc	67020
tgtacatggt	tagtacagat	gcagctaccc	acttctttcc	cccgtatatt	ttcgattcct	67080
gactggttga	atccacagat	gtggaaccct	tgggtatgaa	gggctgagtg	tgctgtgagc	67140
tcctggaccg	cagggcctgc	atggtgttat	tcacttgttt	tcatttccat	cccctagctc	67200
ttagcaaggt	ttctggcact	cagaaagtaa	atgttttttag	aatattggga	gagatagatg	67260
taaaataatt	tgtaaaatca	atttccttgg	cttgatattt	actcagttta	ttagaataaa	67320
ttaaacattt	tattgaaaat	atttaagaat	ttgggacaca	ggctatgtct	tgagacagcc	67380
aaactgacaa	atttttatga	aggacagggt	ctttctgcct	gtgtgtgtgt	gtgtgagaga	67440
gtgtgtgtga	gtgtgtgtgt	gtgagtgtat	gtgtgtgtgt	agtgggggta	ttgggtattg	67500
taatgaagtc	aagagagtaa	ccaagacatt	ttcttcaggg	cttctcagat	aatattattg	67560
tatattgtaa	ttcggcattt	tatcataaat	aagaattgtc	attaattttt	cccctttata	67620
agtggtcact	ggaaatgttt	attaggcagc	tagcatcaca	aatgtgtgt	gtgtgtgcgt	67680
gcgcgcgcgt	gtgtgtatga	ggcttaacgt	tgaaggctct	tgttttgtgt	ttctagagtt	67740
ctcagtggca	cagctgagaa	tgggtgtgta	agtcctgctt	gccaaagcca	tatcctccct	67800
gtgtgggggtg	ggagttgcca	gagaggccag	gaggagtcac	actaggagca	agaaggctgg	67860
gaaggctgga	gactgacatt	ttcaggcact	cacatacaat	gtgtcattta	aattcctact	67920
ctgtgtaggt	cagccacctt	tattcccatt	ttgctggtaa	gaaggtagct	tagaagttaa	67980
aatgtgcacc	tattatttaa	gaaatgtttt	gtccttatag	agacaggatc	ccaccttggt	68040
gctcaggctg	gtttgtaact	cctgggctca	agcaatcctt	ccacctccac	ctcccaaagt	68100
gctgggatta	caggcgtgag	ccactatgcc	tgaccactc	cccttgttta	agggaaactt	68160
ttgctattaa	ttttatgctt	gatttaagcc	tattagtttc	agattgttgc	ctgacaattg	68220
agttccctgt	ttgtttttcc	aaattactgg	ctagttcatt	ttatgtgaac	tttttttaga	68280
aataaaagca	ctgtcaagtt	agaattcatt	tatcagtttg	tcacacgaaa	ctggctccat	68340
attcttctct	acagtagagt	gctttggctc	tagctgatac	ctttgaatga	aatagaatat	68400
tttcttgaga	aagcggcacc	atagttgagt	gcatagatgc	ccagaaaata	tttgaaaacc	68460

WO 2006/119775

PCT/DK2006/000254

aaatcctctc	atggtttgta	cttaaaattt	tttttttagag	atgggatctc	actgtgttgc	68520
ccaggctgga	gtacagtggc	tactcacagg	cactatcata	gctcactgca	gcctcaaact	68580
cttgggctca	agtgatcttc	ctgcctcagc	ttcccaagta	tctgggacta	tagatagtca	68640
tgtgccacca	cacccagctc	gtgtggcttt	ttaaacttca	aaggcacata	tttatagtac	68700
tggttgagaa	ggtatttagg	tagttcactt	ctagatagcg	cttctgatgg	gtgtgacctc	68760
ttcttgtcca	tgcttaccta	gcattcttat	ggtgggcttt	tatcaggaca	ttctactttt	68820
tattgctggg	aggtaaattg	agcagaagtc	actcttgtgg	tggagtaact	agtgacttga	68880
gctgggtgtct	gatcttcagc	gccctgccat	ttgtcagtcc	cctctcccct	ccttcacagc	68940
cacagcagct	tcccagggtt	tgccctccag	gaccagcggg	accctgcact	gggctggttt	69000
ctaactcccc	tttgtcacac	aggttcctct	acctgggatg	cctttccttg	ctgctcctgc	69060
tgaagaattc	ccactgaaac	atcagccctg	ggcacccttt	ctggaccctc	tatgtgtccc	69120
cgtagcatgt	tgtgcagaac	tgtttcctag	ctcttggccg	ctgtgttgga	ggagggttgtt	69180
caagtgtcca	tcgaccctgt	gagaatggag	tgccccagtg	tgggggcacc	tggcttacgg	69240
gctcaggagc	tgggctgccg	tggagcctgg	ccccagtcac	cttctccctg	gccaatgctt	69300
gtggaaggtc	cccatactcc	aggcactgtg	ccagggtcca	gccggatgcc	atttgatccg	69360
tgcaacctgt	caggcagatt	ggattctatt	ttacccatga	gcagtgtcag	gcgcagagct	69420
cttaggtatt	ttgcctgagg	cagctggatc	attaagaggt	ccccccccc	acccccaggg	69480
gtcttgtgaa	gactcataga	tgcagatttc	agagaccttc	tagatttgtg	catccaggct	69540
agcgtgttct	ttgggtgttg	cccctcttgc	tgtctctcat	attcagcttg	ttcagaccag	69600
ggcttcagct	cctcttgagg	attccactca	tcagactgtt	gggtagggaa	gagcctctga	69660
aacggctgcc	tccaagcatg	ggaagaggcc	gctcctcttt	tgtggacaaa	gggccacaga	69720
ggccctttgt	gtgacaggaa	gtcctgcttc	tagaaggggc	tgcttgaaag	tggagggggtt	69780
gaaagaagct	gaagcgttga	cctggggcca	ctctgactca	tgcattgttg	caataatact	69840
ctttgtctgt	cttgggaacc	tggcacagca	gacctcgaag	gtggcctgga	gcattcctct	69900
tctccctcct	tcctgctcac	tctgccgcct	gtgaatgccc	ctatcttctg	acctgaacaa	69960
agggcagcct	cctgacctat	tctatatattt	ttccttctat	ttatgccagg	agaatctttc	70020
tgtctcatct	gggcctgtgt	gagttttctt	ttttagggtt	taagtgcgtc	tcttgagttg	70080
gaaaacttgc	agcttaaggg	agagctgtcc	cactgggcag	gtgcctctga	cggagaggag	70140
gcgtcgtcgg	gaagtatgtg	gccgtggctg	ctctggcatg	atcctgaaga	ctggcacgaa	70200
ggggaatgga	gccgcacaga	gaatttcagg	catcagaggg	gaagcgcagg	taccgggacc	70260
cacggcccag	ctccctgcaa	catggtccac	atgtctcctg	tatttgtggc	catctgtcag	70320
ttgttaaata	ctgtgggttag	gatgtgagga	caccagaagc	gagggaagtg	gtggcagaga	70380

tgagctgacg	cgctcctcatg	tgctctgaag	tcctgatccc	cagctggacg	agagaggatc	70440
atgtgtgtgc	gcatgcctgt	gtgaagggat	cctaagatta	tgttggaata	acactgcatc	70500
agtcctgggc	aggcagcttt	gtgtgtttgtg	taactggggc	ttcttgctgt	ggggacatgg	70560
tggcaaatca	gtgagcactt	gttcttttttg	ttttttgaga	cagagttttg	ctctttcccc	70620
caggctggag	tgcatgtggtg	cgatctcact	gcaacctcca	tcacctgggt	tcaagtgatt	70680
cttgtgcctc	agcctcccaa	gtagtgggtg	agcccgtggt	acattgctgc	tctcgccagg	70740
ctcagaggcc	aatgtgtgcc	tccttcctgg	atgcagtggc	agtcagtgag	tgcttccact	70800
cctgccccca	gggacaaatc	ctttggcttt	attagctttg	tccttacttc	ccccttagac	70860
tgtggcaggt	ttgggggtacc	catacttctg	gagaccctac	cctggattat	attaggtatg	70920
ttatgctcac	atgcagttta	aacacttctt	tatgttataa	aagtaattgg	aaggattctt	70980
attattttgc	ttttgaaact	atttggctat	gagaagttga	tttctcctcc	acctcctcta	71040
attcatttgt	gcatttgttc	agagggaaca	gtgtgtgtga	aggggtggag	ctgataggag	71100
tgtggtgctt	gtgttggtag	acacaggtga	ttaacggtgg	ccagagcgtg	tatagtgtgg	71160
ccaggagtgt	ggtgtgagtg	gtgaggctgg	aggagcagta	gttacagcct	ggaggccctg	71220
ctggcttcgg	gaggaagtag	ggggaggagt	ttgagtagag	gcaaaacctg	atctttccag	71280
aagacattcc	aaccaccatg	gtagagtcac	agtaatgtca	tggagaatga	agggagtcgg	71340
ggactaaaca	agtggctgca	ggttagttag	gacgttgтта	aggattaggc	aggaggcaat	71400
ggtggcctta	gcgaaggtag	ctatcatcat	ggggacagag	ccaagtggga	ggattcaaga	71460
gatattgaag	agctgatgta	gtccggtctt	ggaagctgat	tgaatgtggt	agtggtgagg	71520
ggggttgggt	gtgggaagaa	gaggagtttc	tggcttgaga	agctgggtga	atgaatgggtg	71580
ggaccatttc	tgacgtagaa	aaccaggaa	caggagtggg	ctcataggga	aggggatagt	71640
ttgcaaggac	tgtgtgatgg	ccaagtggag	gcatccagta	ggcagagatg	cgggtctggg	71700
ccagacagca	gatcaggggg	cttgaggtta	aaggtggttt	ggtgtgtggg	cacagaagac	71760
gtaagagaga	gggcctggga	cagaacctgg	ggagcgccag	cagccaggaa	acaggcagag	71820
gaggttgaga	gaaggcacac	tgagccact	acagaaatga	gctgggtgagt	gtcagctccc	71880
tccccaaga	ctcctcttgc	cctcgtcccc	cactgatttt	gcttctgtca	gcttggtaat	71940
tgccggtgcg	tacatctagt	cactgtgcag	actcatttga	gcatctacct	catgctaagc	72000
actggggata	cggagatgac	acagcccttg	cctctttttt	tttttttttg	agacgagtct	72060
tgccctgtcg	tccaggctag	agtacagtgg	tatgatcttg	gctcaccgca	acctctgcct	72120
cccaggttca	agtgattctc	ctgcctcagc	ctcctgagta	gctgaaacta	caggggctga	72180
ccaccacacc	tggctaattt	ttgtattttt	agtagacata	gggtttcact	gtgttgggcc	72240
ggctggtctt	gaactcctga	cctcagggtga	tccacccatc	ttggcctccc	aaagtgctga	72300

WO 2006/119775

PCT/DK2006/000254

gattataggc gtgagccacc gcgcccagct acagtcctca ccttataagt tgatgtgtag 72360
 tggatgcgct gtgccactct gccttttttg cttataccca atttcggatg gaaaatagct 72420
 cacctgctag ttctctgtaa cggatgtggt gtgtcctttc tgtcaactcg tcattgtttt 72480
 cttagctttg agcagctgac tcgtttccaa tttttgctgt tataaatatt gcagtgatga 72540
 acattttgtc caagctattt tttctttgag tttatttgct tggggcatgt tccccagagt 72600
 gggacttaag ggtcagaggt tggtaactgt tatctgcctt gttctgtatt ttcagatttt 72660
 tccatgaagc ctattagcaa tgttagagta catttctggt ttcctgtagc cccaccaaaa 72720
 ctggcattgt catattaatt ttccttattt tacaagttat ttccatattt taaaggaagt 72780
 tgtgcttttt tttcctgcat gggaatttgt tgtattatat ttagctaggt tacattgcag 72840
 taacaagcaa accccaagtc tcagagactt agcctaacac aaatgtcttt ctcgcttgat 72900
 ctgtgtgtcc cttgtaggat ggctgggggt gcggttccac agagtccctc aggccaggct 72960
 gattaaggag acagcactct ctgagagctg ccactcgagg cctgctcagt ggacaacaca 73020
 gtagacagga gaggctgggg gctttgtgtg gacttttcac tactgatgtg gccaggacta 73080
 ccaacacagt ctggcttagc aggtagaact tagtgagcat tcctgtctct gccagccat 73140
 ctctcttctt gtttgttcct ctccagcaac ttctttgtcc agtggaattg gatcctctgt 73200
 cttcatatca gctcttaatg tatttttaggt aataagattt ttatcaattg tgtctttttc 73260
 tttctttgct ttctatatgt tattgtgcaa aagttaacg taagttttgg acacattcat 73320
 gttttcctct gataccttat tatcacaat ttcttttctc aaagccagag aataagctgc 73380
 ttttctttga gttgcttaaa atggatttta tttggctcct tatgacactt gaaattcatc 73440
 aaaatatata atgtcaggtg gttttaatct aagttaatta ggtccttggg gatgcctgtt 73500
 ctgcaacatt aaaacaaaac aaaagaaaga aaaaacaccc aacccaagc aacacctctt 73560
 catgacagaa acactcagca aaccaggaac agaagggaac ttgactggat aaagggtatt 73620
 aatcaaaaact tccagggctg gcagtgaag actaaaagct ttcttcctac aatccagaag 73680
 aataggacct tgatgtctgt tctcactcct tctgctcaat attgtactgg agagtctaac 73740
 caggaacttc gaagttagac aagaaaacaa aaggcataca gattggaaag gaaaaggtaa 73800
 aacactctct gttcaaagat ggcatgatct tatataaaga aatcataag gaatcaccaa 73860
 aaaaactctt agaattaata tacaggttca gcaaggttgt aggatataag atcaatatac 73920
 aaaaataacg atatttctat aatctagcaa ggaataattc aaaaaatgga attaagaaaa 73980
 cagtttcatt tattatagta tataaagaat aaaataggag taaatttatg aagagaagtg 74040
 taagacttat attctgaaaa ttacaaaaca ttgttgaaat aaatttaaaa ggacctgaat 74100
 aaatggaaag agatcctaca tccatggact ggaagacttc acatagttca gatggcagta 74160
 ctctcaaat tgagctacag agtcacagca gtccctttca aaatcccaa gcaacaagcc 74220

WO 2006/119775

PCT/DK2006/000254

aatttctaaaa	tccctgtgga	aatgcaagg	accaagaata	gtcagaacag	ttttgaaaag	74280
tcagagaaca	aagttggagg	gctcacactt	cccaattttac	aaaaatgtat	aaatctacag	74340
tattcaagaa	acagtgtggt	gcgggcataa	ggatagacat	ggaaaaatga	gatagagtca	74400
agaaaacaaa	actaaacccc	tacattttatg	gtcagttgac	tttggacaag	gatgctaagc	74460
cagttcaatg	ggggaaagaa	aaatcttttc	aacaagtgac	gctgggacaa	ctggatatct	74520
gcttgcagaa	gcatgaattt	ggatctctgc	ctcatgccac	atataaaact	taactccaga	74580
tggatcagag	acctaaatgt	aagagctaaa	tctgtaaagc	ttttttaatt	aaaaaacgt	74640
ttaaaaaaat	tagtcataat	aagtgatgta	aaactcttaa	aacatagcat	aagtctttgt	74700
caccttggat	taggcagtg	tttcctagat	atgacaccaa	aacaatcaac	aaaagaaaaa	74760
agtgggtaaa	ttgaacttca	taaaaattaa	gaacttttgt	gcttcagagg	atgttatcaa	74820
gaaagtaaaa	agacaactca	gaatggaaga	aaatagtttc	aatcatata	tctgataaag	74880
ataccttgaa	tacataaaga	actcttgggt	gggtgcattg	gctcacatct	gtaatcccat	74940
tactttggag	gatcacttga	ggccaagagt	ttgagaccag	cctgagcaac	atgagaccct	75000
gtctctacca	aaaaaattga	aagaattaaa	aaaaaaaaaa	aaaagaact	cttaaaattc	75060
aataataaag	gccggacaca	gtggctcact	cctgtaatcc	cagcactttg	ggaggcagag	75120
gcgggcagat	cacctgaggt	caagagttcg	agaccagcct	ggccaacatg	gtgaaaccct	75180
gtctctacta	aaaatacaaa	aattagctgg	gcgtgttggc	ccatgcctgt	aatcccagct	75240
actcaggagg	ctgaggcagg	agaatcgctt	gaacccagga	gtcagagggt	gcagtgagcc	75300
gagaatgcgc	cactgcactc	cagcctgggt	gacagagtga	cagagtgaga	ctctgtctca	75360
aacaaaacaa	aacaaaacaa	aaaccccaaa	ctcagcaata	agaagacaca	ccagttttaa	75420
aatgagcaa	aggatctaaa	ccgacgtttc	ttcaaagata	gtatacaaat	ggccaataag	75480
cacatgaaaa	gatacccaat	atcactagtc	attagggaaa	cacgaatcaa	aatcacaatg	75540
agatatcact	tcccacccac	taagatgggt	acaataaaaa	agacaacatg	tgttggtgag	75600
gatatgggaa	attggaactg	tcacacactg	ctggtgggaa	tgtcaaatga	tgcacctgct	75660
ttggaaagca	ctctggctga	atgttggaaa	ctcctgactt	tggaagtttc	tcattatgac	75720
ccagcagttc	tactcctagg	tatctactca	aaagaaatga	aaacatctgt	ccatacaaag	75780
gcttgtacac	aatgggtaat	agcagcacta	ttaatagcgg	aaaagtggca	aacaacccaa	75840
atgtccagtc	acttggtaat	gggcaaacaa	attgtggtat	attcatacag	tggaatatta	75900
cttagcaaca	aaaaggagtg	aaatacacat	gctgcaacgt	ggacgaacct	tgaaccttga	75960
agacattctg	gaaaaagtga	aggaagccaa	tcaccagagc	cacatgttat	gatcccattt	76020
atctgaaatg	tccagcatag	gcaaattccat	agagacagaa	agtagattag	tgattgccag	76080
gattgggagt	agaggagggt	ggggatgact	ggggaagggt	gtggggtttc	tttttgggggt	76140

WO 2006/119775

PCT/DK2006/000254

gatgaaaatg	ttctaaaatg	attgtcatga	tgtggcacia	ctctgtgaac	acactaaaaa	76200
tcattgtgta	catttttagat	gggtgaactg	tgtgaattat	tatcttaata	aaactgttag	76260
aaaaatagca	agaatctcaa	gtacttcacc	catcatccct	cccacagcct	ctcctggcta	76320
aataatccca	tccctttggc	tattcttgac	agacacattt	gggtcttttc	catttggtata	76380
cccagaattt	tattcttttt	gagcctccta	aattgtgact	tccgggatta	aacacaaagt	76440
cccaaggtag	gggaggaggt	ggtggcggtg	gcaggggctt	caccagcttg	aagagaagca	76500
gctgggtata	gggaaacca	ggacctgggg	tggcccctct	ccagcatgct	agtgagggct	76560
ctgtgccctg	tcccagggcg	gggctcatat	tctggtggca	ctgcttctgg	tgtaactggc	76620
tgaccttgga	gactgggctt	ctttccccct	ccctttccat	gaatgtaaaa	ggaggagtgg	76680
attgaggagg	aagtttgaga	gctacctggt	gtaatccgta	ttcctgttga	caggaacagt	76740
gttactcatt	cagcaaacac	atgctgacct	cctgctttgc	acagcgtctg	agattcagag	76800
atgaagacag	tataagggtca	cgatctaatt	actcactctg	cactacattt	ctttccttct	76860
attgtataaa	atagaaaaat	ctatttttgg	tgtaaaatga	catcctaatt	actgagtagt	76920
taaggagttt	gagtttttaa	gcccgtgcat	gctttccact	tggttttact	gtttgttctg	76980
taaggcgctg	tgtaggaata	gtactaggta	acatctgtgg	agctgatagc	aagcattgtc	77040
ccgctcctgc	ctacagcctt	cggaggagct	tccctgcagc	aggctgctgg	atgcaggggg	77100
ccgggggaat	aattccattt	tagaaacca	gtaattgtgt	ctggtgctag	gtagaatttt	77160
ccttcattca	gtagtttatt	gaaggccggg	tagacgccga	ttgccactaa	catgcagtct	77220
ctagaagtca	gggactttat	cttgcacgct	gccatctcct	gttcctggaa	catcgcaggt	77280
gctcaagatg	tgттаattag	cgaagaagca	aatgactaga	ggatgcagtg	ataatgacct	77340
ttcgttcacc	ctgagaacta	ggctggccag	ggagtgtctc	tccctcttgt	aggtaaggaa	77400
acaggctggc	aggaagtgac	tgcataagat	taggcagact	gtaaaagcca	gagcggggcac	77460
aagaatccag	gtctcctgac	tccatccttc	cattgtgtgt	tctagtccca	cccaactttg	77520
aaagcaaaaag	ttgttatata	tttctgttaa	aaacgttgcc	ctgtgggagt	gcttcccatt	77580
ccctctgtgg	ggttggctcc	actggaatgg	ggaggggagg	agctgtggag	ttgcatcggt	77640
ctggttcagt	taggcctgct	ctaactgcgt	cacttccaat	cattcttctt	tcaaactctt	77700
agactcttgt	acatcttctt	acactactgt	ttgaattctc	atgttaatat	caaatttgaa	77760
aaagtgtcat	ttttgtcctt	tttaaagtaa	gtttttctgt	tttaaacctc	tggaacatca	77820
actgagtatt	taggttgggg	aagtgttcag	cctccatctc	cacattccca	gggtgcttgc	77880
caaccagagg	ctaaaaaaat	attgctgtaa	aatcagttcc	aaggcaagtc	tggaatctgtc	77940
agctgatgtg	ttcccattgc	ataagataaa	cgagttttgg	ggacacatct	aattacacgc	78000
tgttagctcc	agctcttcat	taagttgcat	aggtgagcta	agagctgatg	ggctgctgctg	78060

WO 2006/119775

PCT/DK2006/000254

ggacgggtga ccttggggcc tgtcctacct ttgctgctga ttctgatagc ccaggtggtt 78120
 gtcagctgga gcaagcggtc tgaatgccac caggatggct ggtcatcagc ttgtttcctt 78180
 tgacctataa agatgttggg gaaaacatgg ttattttaag ggttactata tgaatctgca 78240
 tataggattt ttaaaaaata caaatgtgtc acagacagca tttttacctt ctccacctcg 78300
 ggatccagca caaacagcac ctctctctt cccttccctt gtcgtcttct cttgttactt 78360
 gagcacttgc tttgggttag ataaaaagg tgaagccctt ccctcgtgga gctccagggtg 78420
 gagtgggaga aaggctcttc accagccagg cacgggtagc acacaccgga atgccactgt 78480
 gccaggctgt cgagagggag atggtgcaac ctaaattggc ctgcggaaca tgcagttggc 78540
 ctgtgcagag gcattcttca tgtttggaca gcctctgctt gctgtggtcc ttggcacaca 78600
 gcacacacct ggtgagctct tcccatagaa aaagattccc tacctgcttc tgagccactg 78660
 ctgctgtgac tgaccttcct ttcctctaag gtcattggaa agggagaagg tcaggcttct 78720
 ttctgacgcc tcctgtgtgt tccagagggg aggattcctt atgcagtttc tgcattcaggg 78780
 gcttcccttt cttcttcagg ccaattttaa gaataagatt aattggcata tctttctgtg 78840
 ggatcttcag aaccgtgtgt gttttcatcc tcgggtttta ggaattcact gcctagcatt 78900
 ctctattcat accttctgct gcttcgagc cccaggaaac aacttggaac ctggatctcc 78960
 agtggagaga gtgattgatt tcaaaagacc ttccccaac ttacctcaag tgattgcttg 79020
 gtttgtttca aaaagatttt aaattttgct cgacttaaat ggcagtcgtt tttagggctg 79080
 tcatctagaa tctagtattt aacatgatga ttttgtttta ctctgtacaa ctgtgggtag 79140
 agggcaggga gcaagttagc tgggaaaaat caaccaggca ggcgcagggtg agacttctac 79200
 cgctgtcatt ccagagttct ccagtgggag attgaagcct tctggagatg ggggtccaga 79260
 ctgctcgagc caggagaagg gtccccctgc ggggctccat gtctgtgtga ctggtcttca 79320
 gatcgtgtga ccccctaggg atgctggact ttgctgttct cggagctgca tgggtgccat 79380
 ggaaagggtc gggagccagt gtccctgctc ctcaccaca gtggggcctt gggtgagtct 79440
 gagggagagc cgagttaggg gaagtgcagg atttggtggc aggcaagggg tgtggttaagt 79500
 gaaagcacct cgtatgctga ctcacatgcc tgactgcttg cacagggtc cttgagaatg 79560
 aggccgtgtg gagattcagt gtgtggtagt cactgcacac aaccttggtg atcgagccgg 79620
 gacatctagg caggggctca gaggtggact ttggtgggtg ggctagaagt tgaacaaaga 79680
 ggtcattttg ttttgtcttc ggtgcccgtt taagctcatg catcgttatt tctactgccg 79740
 cctgcctggg ctcaccaca gccatctcac caggaactgc tgccacacct gtctttctgt 79800
 gcctcagaca tgccagggtg gcctctcttg ctttttgctg gggctgcctg tgctctgggg 79860
 gcccatatgg cttacttctt cccatgttgt ctcagggtgc accttctttg gggaccttct 79920
 ctggtcactc tgtttagaag cagtctcctt tctgtctcct ccctatctca tcctcgctta 79980

WO 2006/119775

PCT/DK2006/000254

tttctctgca tagcacctgc tgggtcagtc acctgtgcca gaacacagac ccatgagggc	80040
aggagttttg gtctgttctc tccgctgtta gattcccagc attcagaaca ccgtgtggct	80100
gcacaggcac ccaatataca ggtgttgaaa gaatggaaag taggcattgg tgggtagatt	80160
actaatgcat tcgcaatggg tgtaaataagg tatttgctaa tttaattttac tgcccctgcc	80220
cccccccctt tttttgaggc agagtcttgc tgtgttgccc aggctggagt gctatggcat	80280
gatctcggct cactgcaacc tccgtctcgc gggttcaggt gattctcttg cctcagcctc	80340
ctgagtagct ggaattgcag atgtgtgctg ccacacctgg ctaatttttg tatttttagt	80400
ggagatgggg tttcaccatg ttggccaggc tggctcttgaa cccctcatct taggtgatcc	80460
acccagttgg accttccaaa gtgctgggat tacaggcatg agccaccatg cccagccata	80520
accccatctt tatttctgct ttgatatatg taaaagtctt ttttcttttt ttgctttgat	80580
acaaaatgca tactacatat tttcacttaa accagtccat cctttgcctt tgcctatgct	80640
tcatgtcccc ttagaatgtt tgagtcttca taggacttaa agatcatctg ctccatgaac	80700
aagggaactg aaggtcagag gggaggagtc ttgtgctaag tccgggggtca ttctgctcca	80760
tctccttcca ctgatactgg agccaggga ggaaccaac atggattgtt tgccagatga	80820
gagacaggta tttccccgca tgatctaagt gaatcctcag aagtcacagg gtgtgcattg	80880
tcatccctgt ttcattgagt gaattctgta ggattttttt tttttttttt gagacggagt	80940
ctcgctctgc cgcccaggct ggagtgcagt ggcgcaatct tggctcactg caagctccgc	81000
ctcccggggt cacgccattc tcctgcctct gcctcctgag tagctgggac tacaggcacc	81060
cgccaccatg ttcggctaata tttttgtatt ttttttagta gagacgggggt ttcaccgtgt	81120
tagccaggat ggtctcgatc tcctgacctt gtgattcacc cacctcggcc tcccaaagtg	81180
ctgggattac aggcgtgagc caccgtgccc tgctggaatt ctgtagtttt caacaggaag	81240
aatgaagtga tagaatgtac ctgggtaacg tggccagtaa gacttgatgg ctaggtcatc	81300
tctcaggggt gccatggagc ccctgggatg cctgggactc tctgagctct ccaggctctca	81360
tgctgtctga ggggtgctgct tgtggggcca cctgtgttcc ctgaggacca gtgtgtttca	81420
gtgtctggcc cgtgaacctg ttctcaagtt gcccttgtct ctgtggaaac actttttcct	81480
cgaagccagt atttgggatt ttggtaacaa gccaaattga aatctcagtg attataggac	81540
ttgatgaata atttgtgttg caaaagattg aaaaatcatc cctcacgaag tccattttta	81600
gtgttgccta agtacagtta cttcattgaa actgctcacc agctgggtggc aaagcttatg	81660
tggaacaag ttttttggtg ctctgaccct agacaagacc atgagtctgg agccatgtcc	81720
acagtaggtc tccgcacatg tggaagaggg gcagtctgtg caggagagcc ttcttcccc	81780
caccccgccc catgggaaaa ggagagaaat ggtgagttag gtcaggggaa agggaaacag	81840
ggaggaggca ggtcctcgat aatcaggaat atggcgcaga cactgctgga ctctgaggtt	81900

WO 2006/119775

PCT/DK2006/000254

tcctaaaaaa	ttcatgccct	gggctgggca	cggtggctca	cgcctataat	cccagcactt	81960
tgggaggccg	agacgggcag	atcatgaggt	caggagatcg	agaccatcct	ggctaaaaca	82020
gtgaaacccc	gactctactg	aaaataaaaa	aaaaaaaaaa	aaattagccg	ggcgcggtgg	82080
cgggcacctg	tagtcctagc	tactcaggag	gctgaggcag	gagaatggcg	tgaacctggg	82140
aggtggagct	tgcagtgagc	cgagattgcg	ccactgcact	ccagcctggg	cgataaagcg	82200
aaactctgtc	tcaaaaaaaaa	agaaaaagaa	aaaggaagaa	agaaaagaaa	agaaagagaa	82260
aagaaagaaa	agaaatccat	gccctggctc	cattttttct	ttcagtgacc	ctcactgagg	82320
agctgtgggt	accgtgatag	gcactaggga	aggcaggtga	ctcgagtatg	gtctgatcta	82380
ccagggttaa	caacttggtt	ggggaaagag	gaatgggggc	caaataatgg	caccatagag	82440
gaactataac	gggacagtgg	tgggcaaagg	gtcactagaa	cgggtggggg	cttcttgagg	82500
gagatgacac	tcaaactggg	ccttgcaagg	taggtgggag	ttgaccagca	aggcggcagt	82560
ctcaggtggc	agtgtgctgc	acgtgtggct	atcagggccca	gcccattgagg	ggcaaggctg	82620
ctgggcaggc	tgggaagcca	gagtggggag	gccccaggag	gtgttctgtg	tgcagctgcc	82680
acattgaaag	gtgtgcaagt	gtagttggag	gatgtgcagc	tgaggcaggg	cccaaaggag	82740
ggggagtaca	cagaagccaa	gaagggaaaa	gaagggaaaa	ggcgatctgc	acacacacag	82800
tgccttgaag	gtcttgcggt	tcccgtagaa	attgtgctgt	agctctttac	cagattgact	82860
ggtattttggg	cacacagttg	agcatactct	gggcctcggg	tatatgtttt	gaatttagta	82920
tggtttttat	tcaaagcctc	agtgtctctc	gcacgtttcc	atactcccat	ctcaactcta	82980
aaacatataa	gttgttattc	agatcaactt	gctggtaaaa	ttagagaaaa	tttcaaagct	83040
cctttttact	ctctttgtct	cacatttctc	caatggggaa	aaagcctgga	agacaagtgt	83100
gttaaagaac	tggctgttct	tgtaagtatt	tgaggagAAC	agtgtgagtc	attgtagctt	83160
ttcactggta	tttgagtatc	cggttttatt	ttaaacagtt	ctgtgaactg	gagttttctt	83220
catagtttct	gttaccatgg	cgccataata	tttcctcgtc	tgcttgttca	tactttcaca	83280
cacttcagca	tcacattttac	aattcctttac	aatcagaata	tcttattcat	aacttccaag	83340
aatgttttct	cagctgcaat	agcatttttt	tccaaatggg	aaaagtccgt	tgttctgcca	83400
cttgccagatg	ctgggaagaa	agaggggtctg	gttctgtgaa	ggacgggtgca	gggtgcgggg	83460
gctgtgaggg	ccccgttggg	gtttgagtgg	gtggacctat	gtcatctgtg	tggagccact	83520
tccccttttag	gcgccacggg	tgagagtata	ggacggacag	gctgctgctt	ttgcttgggg	83580
ggagttggtg	ttatacattt	tggtctagac	atctaacaga	ccacccttgg	ttctttgggtc	83640
ccctcaaaca	tgaacatggg	tgtgttctgt	tgctcttatt	tattttgtca	tttatttttg	83700
gacttgaatt	gaggcttcca	ctttgggcag	gtctcttggt	ttgagaatct	gcagtccaga	83760
gaaaaacaat	gctgggcttg	gcagtcactg	gtgaagtcgg	ctttgcccgc	cctccaggca	83820

WO 2006/119775

PCT/DK2006/000254

tgccaggatta	ttcagagagg	taccttttcta	catctcattg	ctttgagaaa	tatatatttc	83880
ttaaagagtt	gttcaggatt	taaggaaagt	tggggaggta	tatggaacca	tagagatagt	83940
ttgtttccaa	tttcttaaag	tcttagacct	caacttttca	gggataacaa	ggttttcctc	84000
tgacacataa	agataataat	agactagggt	ctgaatatgt	gcttttgact	taattatggg	84060
aacttggctt	tgccagggtg	tttcaaattc	tactgtaa	ataaatgctg	attcaatgat	84120
acagcaactc	ttgatacttg	agaattttaa	gtgctgcgta	actgattgga	catttcgtct	84180
cctaacacag	ttggttcctc	tgattcagag	aggaccctt	tttgaagatc	aaaagcaaaa	84240
taaaagtgtg	gaaagtgtgg	acttctagt	gcatgtggac	aggggggtgt	cctgggggct	84300
gcccaggagg	ctccagggtg	gaagtgcctc	tgacttagtg	gctctacaag	ctggccattg	84360
tggaagaatt	ggctttttcc	tttttttttt	ttttagaatt	gactttgtaa	gaatcatcaa	84420
aaaattgttt	tgctctcagc	tcaataaaca	tttgtagtta	ataatccgag	taactaatct	84480
tttttctctc	ctgtgtaatc	caagatatgt	cagtatatgg	tcttcgtgag	taaaattgat	84540
accaatagta	cattatccta	aaggctttat	ttgatatcct	tttatcaagt	caacacttta	84600
ctgtcccacc	atttcacttg	taaagatccc	caagtgtgtc	gctgtcagtt	tcaacaaatg	84660
ttagaaagga	cctgttgtca	tatgtgagat	acgcttctgg	aattcactgt	ctagagggga	84720
aagcagactg	tacacaggta	aaaactgagg	tagtggtatg	tgctttgagg	aaatataaaa	84780
agagttagaa	ggtaacaggc	tagagttagt	ggtggtggga	ggtgtggcta	gattctctaa	84840
gggatgcccc	aatggtgtg	agggggagcc	ggccatacgc	ttgtctcagg	gaagaacatt	84900
cctggcagag	ggaacagcat	gtgcaaaggc	tctgaggcag	ggatgagctt	attatggtga	84960
agaaacagag	agaagatcct	gtgctgggac	agagaagggc	cagcaggtaa	gggagagtag	85020
ttggggaatg	ggtggaagg	caggcagggg	ccgggtcatg	acgggatttg	cagtacagcg	85080
gaggagttag	ggttgaagt	atgtgatggt	ttagagaagg	gaagtgacat	gatttgagta	85140
tttaciaaata	ttgccctgaa	gaggctagaa	actggaagt	aggatagggt	ttctgttctc	85200
tttaacggtg	agtttccatg	ggaaaacatt	aaacgttttc	atttttagctt	tgaaaattta	85260
gggttattaa	taccacctgt	gttccttaca	attggggggt	ttgtgacatt	ttctgggctg	85320
ttttggaaga	gagggattgt	gtaagtgaat	ataccctggg	tgtaaccat	aataataaca	85380
acgtgatctc	tgcatggctc	tgttaaatta	aaaaatgtaa	ctattaatag	tgtgtgaaat	85440
agttctcagc	ctagccagca	gctggaagat	tacacattat	gagaaactca	tcaagtttaa	85500
accacatggg	gatttctgtg	tagaaatata	ccagttcctc	cttccttcct	cgctgttcct	85560
gtttcagggc	agccatcggc	tttgtggagc	cctctgaacc	ctggctgggt	ccattgacgc	85620
tgactgtgtt	cctccactac	cttcccatca	cctaagggtc	cttctgtctc	ggacacctcc	85680
tgcttctccc	actctgcacc	tgccccattc	ccatccaagg	ttcaccttgg	tgctcgtggc	85740

WO 2006/119775

PCT/DK2006/000254

ttggaagcca	tttcctcact	tccctgggtg	gaaagcagct	acttcagtcc	tccaacaatc	85800
ccaaggagtg	gtaggcagtg	ttagttttgt	tttctggatg	aggaaactaa	ggttcaggaa	85860
ggttaagtaa	cgtgcccctg	gttaccacagc	tggttttgcag	caaggctggg	attcattcat	85920
tgctgcttct	cacatatcat	tgccctcgtgt	aagccagggtg	ctgtgcccag	ccgtgggata	85980
caactgtgaa	caaggcagggt	gtgagttagg	tgtctgcctc	cctcaacacc	ttgagctggt	86040
tgagcaggca	ggcgtctgga	ataggcatgc	agtgaacact	gagcgaatga	atgagtgaat	86100
gaaagaattc	ctgtcaacat	tctactcttg	ttggactcgg	gttggcatgc	tgccctccctt	86160
agcatccagc	ctgtcactct	tcagagtcac	gcacacgtga	ctcacctctg	ccttttccttg	86220
gtgtcttttg	ttgcctgtgt	ctcatgcagg	aggcagcaca	gcctgggtggc	tcaggaggagc	86280
tctggagcac	ctgggtctggg	tttgaatcca	agctctgccc	cttactagtt	gtgtgacttg	86340
ggcacgttat	ttaaacactc	ttatgcctct	gtttcttcat	ctataaattt	acttttgaac	86400
ttgaactgtc	aaaaactttc	attttttaat	tcttaaactt	tcccactgac	ccaaaccaga	86460
ttgagaaatt	ggataaccaa	aaccatggag	gtattttttgc	tttttatgag	cgtgcctttt	86520
gttattttcaa	aatgacctac	tttcatactt	ccaaatgcat	tttagaaaca	ttttgattaa	86580
tttaaaagac	aatcattttc	aaaattagtt	tattgaataa	ctttttccct	tgtttatgtg	86640
gtgtgaacgt	gggtatccat	gtacttcttg	ggtcaggatt	gagtgggtgc	ctcatctgtg	86700
ctaggctctg	cgtgagacac	gaggacagga	gggtccagag	aaggctctgag	agctgagaca	86760
aagggtgctga	gtgtgtccag	ggcccaaacc	cagaggctct	gggtgctgctg	catgttgtgg	86820
ctgtggaatt	ggggcagaga	gcacaggtag	caggaaagtg	tgtgagttag	gctgcatgtc	86880
acagagaccc	tcaaaataat	agtggcttaa	agagaagttt	attcctctca	tctgcagagt	86940
ccagagggtg	catggctggg	gggcacacca	cactgtcaga	gacctagact	tcctctgtct	87000
cgtttttgtgc	atggctctgt	tcccaagtca	cctcatgggt	cgcgatgaga	aaagcatggc	87060
attcatattt	aacacggcga	actttgtgct	aggattggga	cacaaaaata	aactccatgt	87120
ggttcctgcc	ctcaaggatc	ttaggtctct	tgggagagac	agatgcctgg	atgaccgccc	87180
tccatctgag	gttgattgga	aggagccctg	tgtgagagga	acacaggcag	gcaggccagg	87240
cagcagagct	gacatttggg	aataaaggaa	gtcttttatga	aagagggtggc	atctgatctt	87300
gatcttaaag	aataattagt	atcttgacac	tcatctggga	ggctgtttga	gggagaagga	87360
tactgtgaag	ggaggtacga	agacctgaaa	gggcagggct	ctgttttgag	ccccgggtgg	87420
cgggtcactc	tcctgccttg	tccgagttga	agtagaccat	cattctggag	cagacgtcac	87480
acaattggaa	agatgtcccc	agctgtgtga	ttagtatgac	atccattcta	aaagggtgaac	87540
aaaagaacct	ggaaagaaac	cagacaggat	cgtttttgag	ttctgagata	ctcattcttt	87600
tgccctgtact	ttggaagcag	gatcaagcca	ttcgtaggca	cgccactccc	actgaacatt	87660

WO 2006/119775

PCT/DK2006/000254

tcacaatcca	ctgagcccag	tgctgtcact	gtgttttcagt	tctcagaagc	tattgacagc	87720
atcttcaatg	agaagtcctg	ggacccagag	tcagagcact	ttttactggg	tttgcaagat	87780
acaaatatta	attgctgttg	actttccctt	taaattttag	cctctgtatg	tgggcaggac	87840
ttgtagcata	gatgagacct	tgtgttacca	acattgtaag	taattttatgt	gatgtggacg	87900
tgcttgggaa	gcatggactt	tcttttatgct	gactgtggct	tgtttcacag	tgagaccctg	87960
ataatgagct	caagcccgga	cattgggggtc	ctgggtgagta	tagttgcaca	gggcggccgt	88020
caggcctgga	tctgatgaag	tagggaggat	ctttgcagcc	atgctcactg	cacatacaca	88080
ctcttataat	tactcggcctt	taattcacgt	ttattcttcc	ataatctcct	gcttatagga	88140
ggctgtctta	aaacatttta	aactcagaga	actctaagag	cttgttgaat	cccactctct	88200
tctttttcatg	tgaaagtagt	tggatgctga	gaagataagg	gacttgccca	atgtcccatt	88260
gctgtttaat	ggcaaatttg	ggatgaggct	tctaaatttt	ggcactgaga	gctctttgta	88320
agtgtaaaaa	atgacttcag	aattatcaaa	aatgaaatta	tgttcagtgt	agctttattg	88380
attttgtttg	tttagaggta	tatgtatttg	tgtgtatttt	tccttaaatt	tggtgttcca	88440
ggtaggggat	accagtgat	tcccttcacc	agcctgatta	cctacaagtg	cctccaccac	88500
tatttttggg	tgtacttcat	tgtcttcaga	cattaacacc	atctactgag	tttaatgagc	88560
atcagtgcat	caggcccagt	gaaacactct	gacaacttac	atattaaagg	agtttcacat	88620
ctaaaatttc	acaaatgagg	cttatgtatt	tgaagtttct	taaaaataaa	aacaaaacca	88680
acctgatttc	tgtactctga	tgcatgaaat	tagaagctat	gtgaggctca	cctgcctcca	88740
cagcaactga	agtcttgttt	tatgtgaaca	ccttcttttc	tttcctcctt	ctgtccattt	88800
ctgatgcatg	tgtaagatgg	cagggtgat	ttcatcgctt	acagcactct	cagattcagc	88860
acgtgaagag	acagtgaatc	actttaagga	cttcagaggc	agcgacattt	tttagccaag	88920
gtgttttcttt	gggccatttt	cctcctatgc	ctctccgggt	gtgttcccgt	ccttttctcaa	88980
gggttttatgg	tttgcccctt	aagtgtagat	tgaggatgtc	agattgaaat	cttgctcttt	89040
aaactgtatc	cagaggttct	tcaactctct	ttggaatgtg	tgagccagag	agcctcaggg	89100
aaacagtggg	acagaagtgg	aagctgcatg	tgccatgaca	tccccaacct	gggctccgcc	89160
tgaggactac	ctggaaaccc	aaacatctgg	acatgagggc	tcggctcagt	ggagtatggg	89220
accatgacat	gacgtttttac	agccattaaa	aatgactgag	gccggggcacg	gtggctcacg	89280
cctgtaattg	tagcactttg	ggaggctgag	gcaggcggat	cacctgaggt	ctggaattcg	89340
agaccagcct	gaccaacatg	gcgaaacccc	atctctacta	aaaatacaaa	attagccaga	89400
cgtgggtggg	catgcttcta	atcccagcta	ctcgggaggc	tgaggcagga	gaatcgcttg	89460
aacctgggag	gcagagggtg	cagtgagcca	agatcacgcc	attgcactcc	aacttgggca	89520
agaagagcaa	aactcagtct	caaaaataac	taactaaata	actaaataac	taaataaata	89580

WO 2006/119775

PCT/DK2006/000254

aataaataaa	agactggggc	caggcgtggt	ggctcacgcc	tgtaatccca	gcacttttggg	89640
aggccaagat	gggtggatca	cttgaggcca	gaagtttgag	accagcctgg	ccaacatggc	89700
gaaaccctgt	ctctactaaa	aatatTTTTa	aaaattagct	gggcatgggtg	gcggggcacct	89760
gtcgtcccag	ctactcggga	agctgagaca	tgaaagttgc	ttgaactcgg	gggcagaggt	89820
tgcagtttagc	taagatcgtg	ccactgtgct	ccagtctggg	caacagagcg	agactctgtc	89880
tcaaaaaaaaa	aaaaagtaac	tggaagatac	ccagacaaca	gggagaaatg	ctgcttatgc	89940
agctgaggaa	agcagaacac	aggatggctt	gttggtctgca	aggacagtga	agtcagtcca	90000
tagatgatgg	gagaagcaag	ggctaaatga	ggagagcgga	ggctgggtggg	acagaactgg	90060
ccatgaggag	cccctgttgg	ccttgacagc	tctgtgctag	ccaacatttg	gacttaaaat	90120
catcatgggc	ttggcagtgt	ggggaggctg	gaaagaggag	agccagaggt	tctcctgaac	90180
aggaatcgtg	agcttcagag	ctgccgagga	gcccagccag	gtgtggtaag	cagaagaact	90240
gcactgaggt	tgctctgttg	tttgTTTTc	cttaacctcc	tccaagcatt	TTTTTTTTT	90300
taagagatag	ggctctctta	tgttgccgag	gcaggagtgc	agtggctatt	cacaggcgtg	90360
atcatagctc	actgcagcct	tgaactccta	ggctcaagca	gtcctcccac	ctcagcctcc	90420
tgaggagctg	ggaccacagg	tgactcagc	tccagaatga	TTTTTTTTT	tttcctgtgg	90480
gaagaagagg	gcttggtctg	tcatgttgag	tgaaacaaga	tgtgttttcc	cccctctctt	90540
tttctctttt	cttcttaaca	gagatattgg	agagtctgtg	ctgtgtttct	ttgaagtgat	90600
tctgtgtcct	cgtggagttc	ctaagattcc	ccttcctccc	actgtttttc	ttcagtttgc	90660
agtggatagt	tctgtgccat	gagggtggtt	gaaatcctct	ggtttttcgg	gatggtttct	90720
atagcaacaa	tgagttgtgt	ctgcctggag	atgctggctg	TTTTTTTaat	ctctggaagc	90780
acctttgagt	gtgttctgcc	ccaccctgcc	caggccgatg	ggtgcttctg	gagggcctag	90840
gacctgctga	gctctaggcc	tcagtgtgat	tcccttgagg	gaggggcccag	cctttaagcc	90900
caagcatctt	gcaatgcgca	agctgtttcc	cgttgttgaa	tcaaattgtac	tacaccatgt	90960
tccagagaga	tataagtcac	tttatTTtgct	gatctcttat	ctcttccttt	gaggaatggt	91020
cacttttcca	gatgaaatca	ctggaaacaa	aatgagttgt	gccaggtagt	tgagctctta	91080
aaaaacatcc	cagcttttgct	aatggtgcaa	agcactatct	taatgtgggc	cttacttcct	91140
gagcgcccac	tggcttatag	ggggtgacct	gggtggtaga	gacaagcaag	agacccatt	91200
cttccaggag	actgactgga	catacatgaa	tagcccagagc	tgagccaagg	ctggtgacat	91260
agagggccca	gggacctatg	ggccgcagct	gggttgctcag	gcagcaggag	ggcctcccgt	91320
gaggccggtg	gtagacaggc	agggtagaac	cacacagccc	ccaaccgtgg	ttgctaaacc	91380
catgtgtggt	ttagattaac	atgtgcagta	gtgtttacga	aagaattcta	accattagag	91440
ttatagcacc	ttcacataaa	tgctctcatt	attttccatc	tttgaatcct	cagtctccat	91500

WO 2006/119775

PCT/DK2006/000254

aaatcaaaat tacccaggct cttcaacttt tcttgtcctg aattcatcca gcaattttaca 91560
 aagcaactgc tttgtgtttg acactggatg ctggggctgt agcagtgatc acagctgaca 91620
 tttgtccctg ccctcctaga gctgatactt tttttttttt tttttaaaga cggagtcttg 91680
 ctctgtcgcc caggctggag tgctatggcg tgatcttggc tcaactgcaac ctctgcctgc 91740
 caggttcaag tgattctcct gcctcagctt cccaagtagc tggaactaca ggcatgcgcc 91800
 accacacaca gctaattttt gtatttttag tagaaatggg gtttcaccat gttggccagg 91860
 ctggtcttaa actcctgacc tcaagtgatc cgccagcctc ggccgcccac agtgctggga 91920
 ttacagacgt gagccaccgc cccggcccta gagctgactt tctagagggt cagacagaca 91980
 agaaataaaa catttaagta aattatgtag ctcgttagga ggtggtaagt gcttgggaga 92040
 gaaaaataaa gcagggaggg ggtctgaggg ggctggggca gggttagaaa agttctcact 92100
 ggggccggac gcagtggctc acgcctgtga tcccagcact ttggggccgat cacgaggtca 92160
 ggagatcgag atcatcctgg ctaacacagg tgaaaccccg tctctaccaa aaatataaaa 92220
 aattagctgg gcgtagtggc aggcacctgt agtcccagct acttgggagg ctgaggcagg 92280
 agaatggcgt gtacccggga ggcagagctt gcagtgatct gagatcgtgc cactgcgctt 92340
 cagcctgggc gacagagcta gacactgtgt cgaaaaaaaa aaaaaagaaa aaaaacctca 92400
 ctgggggtggg ggctgataca ggcctgtggg gcaggaggga gaccatttca gggaggagac 92460
 gagcacattc actgtgcctg gcagtgccca gaccagcgga ggagccaagt ggagcttggt 92520
 gtaggggtgga cagcagtggg ggcgggttgg agagcagggt ctgaggacaa agtagctcgg 92580
 tggaggtgag cagatgcggg ctgggaggac ctgggctttt actctgcatg gggggcttaa 92640
 gagcagcaat gtggtctgcc ttgcattttg agagggtcac actgttcaga gtggactgaa 92700
 gtcagattgt taaatatatc atgcattgag aaagcatgca gcagtgcata cagctacagt 92760
 tgtatgtaca agtatgaaat taatatgttc taatttttaa aaacagtagc acagatacaa 92820
 aacactgagt tactgtggga agctgggagc tctgttagag ggaagaagac tgcagccaca 92880
 caaaggagag gtgatagtgg cactgacagt ggagagatgc aggggggtctg gtcatgtcct 92940
 tggctttttg gaggcaaggc cgggtccagga ggtgctcatg gacctcccc gctccctaaa 93000
 gctggggcct ggggggtctt tgggaaacct ccttccttct agtatggggc cattcatgaa 93060
 acaaagtcgg gtattctcca ggattttttt tttttttttt taaagagacc aggtctcgct 93120
 ttgtcaccca gattgaagtg cagtggcaca atcgtaactc actgctgcct caaattcctg 93180
 ggctcaagca gttcccctgc ctcagcccc tgagtagctg ggagtacagg cgtgcactac 93240
 cacacctggc taattttatt ttttaattta ttttagagat gggatctcac catcttgcct 93300
 aggctggtct tgaattcctg gcttcaagca gacctccgc ctcagcttct caaagtgcta 93360
 ggattacagg catgaaccac catgcccagt caggataaat ttttagagga aatagttgag 93420

WO 2006/119775

PCT/DK2006/000254

gccgggcgca gtggctcacg cctgtaatgc cagcactttg ggaggccgag gcgggcggat 93480
 cacgaggtca ggagatcgag accatcctgg ctgacacggt gaaaccccgt ctctactaaa 93540
 aatacaaaaa attagccagg catggtggcg ggcgcctgta gtcccagcta ctcgggaggc 93600
 tgaggcagga gaatggcgtg aaccgaggag gcggagctta cagtgagccg agatcacgcc 93660
 actgcactcc agcctgggca acagagtga actctgtctc aaaaaataa aaaaaagaag 93720
 aagaaggaga aatagttgaa gatgaagtag cttagtggag gtgagcagat gcgggcctgc 93780
 gggccaggag gacctgggct ttactctgc acaagggact taagaccagc aatgtggtct 93840
 gccttgcat ttgagagggt caccgtgttg agaatagact gaagtcggat tgttaaatat 93900
 atcatgcatt gagaagacat gcaacagtgc atacagctac agttctgcgg ccaattatga 93960
 aatgaatatg ttctcat taaaaccagt agcatgggta caaaactctg agttaccatg 94020
 cagcttatct atgatgccag attagtcctt tctaaagtgc accctgcccc acagaagcat 94080
 gtgagcttcc gctttgttgg gtgctgggac agccacaaca ctactccct ctgcctctgc 94140
 cggaagacc cctcccctgt cttctcatg cccactgccc atcttcctgc agtcctcaga 94200
 gagctcgcat tctatctcag taagcccgcc ccaccccag gtcttttaga gtgccctact 94260
 tctctgtgtc ctggtgcttt tcacaccata tagctgtttt tactttgctg tgtgccaatc 94320
 tgccctagta ggatgtcagc cccaggaggg cagggacca gtctgccttg tccactgccc 94380
 ctccccagc ttagaagagt gcctggcaca tggtagatgt tcacatggta aatatgtccc 94440
 ttatcaaagg gccctctcac cttgggcca aacattggct tagaataaat cgttatttcc 94500
 caaccactga agaattctccc cccagatagt gaagaggaaa cgagctgcag gtcagcctgt 94560
 gctcagtga aggcctgtgg agcggccaga acccagccag ccagggtgtgc tcgggcccgtg 94620
 gcagcgagga tgtccacctc tgcctacctc tgcagttaga ggattagagc ggggtatcgg 94680
 gggctttctc ccgaggccag ctttcaggag agatgagcaa catggtcttt ttactcagaa 94740
 tgttcttgct ttatgaaaat agccttttgg cctaaaaaat gacctgttct tttgtctggt 94800
 tgtggaataa attaagggtg ttaaggggga gctggatgtc cttgtgaata catcatttgt 94860
 ttacctgata tctgttatca agcccaaagg ggaggaagga ggttatgcca gcttccagcc 94920
 tgtgccacgg tcaccttagc aacctctggt aacctctgag actctgcaca ccatctaagg 94980
 aatggctctg actcactgta cttacataca ttgctggcag aagtcactgg aatctgagat 95040
 ctggaactgt tttcaatacc tcagtcattc ccggattccc agattcctga gaagtgcccc 95100
 ttagaacaaa agggttgggt ggctcacacc tgtaatccca gcactttggg aggctgaggc 95160
 aggtggatca cctgagttca ggagttcgag accagcctgg ccaacatggc gaaaccccgt 95220
 ctctactaaa aatacaaaaa ttagcctggc atggtggtga gcgcctgtaa tcccagctac 95280
 tcgggagggt gaggcaggag aatcgcctga acccgggaga cggaggttgc agtgagctga 95340

gatcgcgcca ctgcactcca gcctgggcaa tgacaaagca agattctgtc taataaaaaa 95400
aaaaagcttc taaatgaagt ttccaaaagt tagaaaagct attttctggc tcctgattgc 95460
tcccttttca tagcagcctg tttgtgtttt atgggtacat aacccttgag tctctccaag 95520
aattctcatc agatttcatt agaatttctc tcttgttccc tcaagtattt gttttcccct 95580
gatttcagtt tgaactttat gaaagggtag ggctgatggc ataggaagct gagctgggtc 95640
cctcttgggc cttcccctaa atgggagagt ggccacatct ctgtaagagt ctgggcaatg 95700
ccccagggtg gtttccctga aagtcactct tacttgatgg cttggctcca aatgcagctt 95760
tttttttttt tttctttttt cttttttgag acggagtctc acactgtcgc ccagggtggg 95820
cgacagtgc acaatctcgg ctactgcag cctccacctc ctgggttcaa atgattctcc 95880
tggtcagcc tcccgagcag ctagaactac aggcgtgtgc caccacgcct ggctaacttt 95940
ttgtattttt agtagagaca gggtttcacc gtgttagcca agatgggtctc tatctcctga 96000
cctcatgac tgcccttctc ggctcccaa agtgctggga ttacaggcgt gagccactgc 96060
acctggccca aatacathtt taataatgac cctggagcag gcacagtga tcacttgagc 96120
ccaagagttc aagaccagct tgggcaatac gacgaaacc cgtctctact ggaaaataca 96180
aaaattagct gggcgcacgc ctgtggttct agctactggg gaggccgagg tgggaggatc 96240
acttgagctg aggaggttga ggctgcagtg agctgtgatc acaccactgc acttcagcct 96300
gggcggcaga ccgagaccct gtctcaaaaa aaaaaaaaaa aaaaaaaaaa aagtgtgata 96360
attcatgata attgaccctc acccctgcca ctcaactt tgcatttctc aaccctgagc 96420
ttggcggttc taccctgcca aggcctgtct ccagcactgg cgttttttagg tggttttctt 96480
ctcctgggtc acacctggct gctctcttcc agcacgtcct ccaccatggg cggtgcggcc 96540
ttgtttctgc gttgttacag cttccacctc ttaattcagc aggccttttg gaggaaggag 96600
aaggaggtat gtttgcctcc cttgaatttg tttctttcca aaaagtgaca ctaaaatagc 96660
catgctgtaa tgtcattatt ttggaaaaga aataaattca gaggaactac atgggaaaca 96720
tattttctta catgtttttg ccattgtgat gggaaagaat ccagaataac tcagcacaca 96780
ccacccatt ctaaaagttt atttcaaaat cacttggttg aattcggatg gcatgcttct 96840
gattcatctt gcatggcacc cagggtagcc cgtgagagt agttcagggtc tggagtgtgt 96900
gcaggaacag tctccctcac gtgtagccct tggtcagagt tcacagaact cattcagcat 96960
tgaattccgg tattggagct cagtttcttt ttttcttttt tagattcttt cttattttta 97020
atttttaag ttctttttaga gaccaggtct cactctgtca cccagcctgg agtgcagcag 97080
tatgatcgta gctcactgca gcctccagct cctgggctca agtgatcctc ccacctcagc 97140
ttcttgagta gctgggacta cagggtgtgta ccaccacccc tggctaattt ttcttttttt 97200
ttttttttaga gatgggtggc tcactgtgtt gccaggggtg gcctcaagt atccacctgc 97260

WO 2006/119775

PCT/DK2006/000254

tttggccttt	caaagtgttg	agatttccag	gcgtgaacca	ccgtgcctgg	ctgagctcag	97320
tttctttttc	tttttctttc	tttcttttctt	ttttttttttt	ttttttttttt	tttgagacag	97380
gttcttgctc	tgtcacccag	gctggagtac	agtgggtgcaa	tctcggctca	ctacaacctc	97440
cacttcccgg	gttcaagtga	ttctcctgcc	tcagcctcct	gagtagctga	gattacaggc	97500
gcatgctacc	acaccagct	aattttttgta	tttttagtag	agacacagtt	tcaccatggt	97560
ggccaggctg	gtcttaaact	cctgacctca	tgatctgcct	gcctcagcct	cccgaagtgc	97620
tggaattaca	ggcgtgagcc	accacgcca	gcctgagttc	agtttcttta	ggaagaaata	97680
atcagtcaac	cactgggtgc	cagacagtct	gaaatgtgca	aaagaagtgc	cagatgagtc	97740
cctgtccttg	gaaagaccag	ggatgactca	tttgtctgtc	actttgcact	ttctcaaat	97800
tctttcacag	ctgtcgcccc	gagcctctta	ataccctggg	aggtaatgag	agcccatggt	97860
gtactgcctt	acagacaagg	acagtggctc	tcagagggga	caaatagctg	aattacagag	97920
gtggagctga	gactcatttt	cgactcaata	tcacactctc	ttcccaggaa	ctttaccttc	97980
cagatgagaa	gcccccaatt	aattattaaa	tacacatttt	acataatttt	tgcatataa	98040
aattgatttc	caaaggctcc	agcactgttt	ggcattaaag	agcagcttgt	cccatacgtg	98100
tgcacaggta	tggtaaaggt	cgttgtagtc	ggctcaacct	ggaagaatga	atgaagcaga	98160
gccaggaaac	atgccctacc	tgacctcggt	ttgaagaaac	aatcgctctc	agaagcctct	98220
cagtctctc	agcttatgtg	tttcttcagc	tgagaggggtg	ctgcaagcag	attcattttt	98280
atggttaagg	acatgtgtgt	gtcaaaaaaa	aattatagta	gtacttcctg	acaggcccag	98340
cgcctttggg	gatgtggtag	agatcagggt	caatgtggaa	gccctgaggc	tccagcctcc	98400
ctggcagggtg	gtgggactct	ggagaatcct	cctgcccttt	cttctttttcg	ctgccagaga	98460
ttgccggatt	ctggaaatgc	ctggtagctg	ggcatgtggt	gtcttggttg	agttgggcag	98520
cgtggagtct	gatggtctgg	gttgagtctc	agcctagcac	ttaagagcta	tgggtctgtg	98580
accagtgatt	tgaattatct	gtatcttggt	tgccttttca	tcaaattggaa	acaataatac	98640
atctttttata	ggattttttga	ggtacaaaaa	taatataatct	gaagttcacc	tcacccatct	98700
gttggtgctg	agaggtactg	agtagctgggt	agctgcttgg	atggtgggtt	atcccacttc	98760
tagaaaattt	gtagcttggg	ctcccagggtg	gagcagtctt	gcttggtttg	atgcaaacad	98820
ggcagactcc	aggcaatata	gctgtctttg	gggatataa	ttcagtcata	aaatagcttt	98880
tgaacactgc	atggtgccat	gaggcagctc	ttctagctcc	ttggactccc	tatgaggttg	98940
agcatggcta	ttctaggggtg	tggcctgtct	gcagtggccc	tggtgtgcag	ctggttccta	99000
ctcaggatgg	ggtcatgccc	accatagcaa	ctgttacctg	caaaggacat	ttttacttta	99060
caaagcattt	tcacatcctt	ctcagttgac	cctcaccacg	attctgccag	gcaagtgggg	99120
ataatattgg	tgcagtaaac	attatattta	aatacagttg	ccctgggact	gagtgttccc	99180

WO 2006/119775

PCT/DK2006/000254

gtgcaacaaa tggcaccatt aaggcataaa aagatagagc atagccaggc agtggcacac 99240
 actctgtagt cctagctact tgggaggctg aggcaggagg atcccttgag cccaggagtt 99300
 aaaggccagc cggggcaaca tagtgagatc ccatctctaa aaagaaaaag aaaaagaaaa 99360
 gagtaccaat cagtgagggg aatgcagatt aaaccaccat ggcttactaa catcttctga 99420
 aaggccaaga tcgacagact gtgcctggaa ataggactat agcagtagtt catcacggca 99480
 caccgcagag cagtgcgtgt tcactttttt ttttgtttgg ttggtttttt tgagataggg 99540
 tcttgctctg ttgcccaggc aggagtgcag tggcacaata atagctcact gcagccttga 99600
 cctactgggc tcaagcaatc ctctcgctc agcctcctga gtagctggga ccacagggtgc 99660
 gtgccaccat gcttggctaa ttttttttaa tttctacttt ttgtagagat gggatctcac 99720
 tatgttgccc aggctggtct tgagctcctg ggcttaggca ttccttctgc cttgtcctct 99780
 caaagtgctg ggattacagg tgtgagccac cgtgtctgac cgatatttag tttttcatgt 99840
 tatatgttct agaagatttg aactcagcac tgagggtggtt aagatcagtg gttaagagca 99900
 tgactgctgt cagcctgcct ttattaaggt ctcaactccg ctgcttgggc aagctccaac 99960
 cctttgtatg ccccagggtc cttgctttga caatgcccac ggcgaaaccg ctaccccata 100020
 gagttgttag gactaaatgg agtgcacgcg tgtagtgcac ggggcgcagt gtctggcaca 100080
 tgaggagtcg ccataggggt gggatatgact ggcttctttg ctgttgaaga agggagggct 100140
 ggtggcaagg tcacagtggc aggtggcagc gtcaccaaaa gtgcttatcc cagtagcaac 100200
 aatgacagga aaacaaaaat caaaccagc cccagtcacc aacatttccg agtaccttcg 100260
 ggtggcttgc tgtggtgatt ctaagaacaa tggctatcac ttttaggggtg tcagatgtat 100320
 gccaggtgac ttcgtcactg ctgggccttc tccatgcctg cgaggaaatc accttattgg 100380
 gttgagaggc tgcctcttga gagctgctgg tgtgtccctg cacatggatg ccgtttgaag 100440
 tcatgggagt tccagagatc agtgctggga gggatgggag gagccataga gcggactgca 100500
 aggtgtgcaa gcagttgggg gaggaaaggg aggcttgtca gggtgagagc ggggagccag 100560
 gagagtcca tacgctgagg ctggtgcggc gtgccgtggg aggggtgagg tggttggaca 100620
 gggcccctgc tgctcatatc ccggagcagg ggcaggagag aggggacttg ggagttgtca 100680
 ttgagaactc agctcggtgc ctgccctcaa ggccccctt cccaccccc catccccacc 100740
 cccaccccag gttgttgctc tgtgaacatc agaaatgagt tcccagattg gccatggttg 100800
 tacaatcagt tgggttttct ttgtggtgca gattcttgtc cccagggtcc tcccctttac 100860
 ctgcattcct tgtgactgtg gagctgaagg aagtagcaag gacactttga agcccagggt 100920
 gccggccaag aacaatgctc ccattgaggt aggccgtgcc acggcctgtg acctgggggt 100980
 actgagacct tcattctatt taaccttcac tgaaacctgg taaagcacag ctcaggaact 101040
 ggggttcaga gaaattacat ggctttctgc ggtcaccag ctggccagt gtagagctcg 101100

gatttcccg cttccaccat cccaggcctt cccacaggc acctcgtctg cattacgtca 101160
ctgactcttt gggcaggaga ctttcatttc ttcccctcac agtctctgtc tttttgggtt 101220
ttctatatct tgcctcagtg tcgccttagg gcttcggata tctttgctag attctgactg 101280
cacagttcag tgagaccttg tataagatgt gcttagaggg ttcagatgaa agtagttgat 101340
tgtgaattat tccagctcaa gaactgaagt ggtgggtgcag agcttcctaa agtaatttct 101400
gttgggtctg aaatgccttg tgcattttat aaaaagattt gtcttcctaa ttctgtttta 101460
cttaaattaa taaagatgta tataatctcg gagaatgtct aagttggaag gtgacagatt 101520
tgtggcaggg aggagctggc tgagcattaa gatgaactct taagtaatgc acatggggat 101580
agttgagctg atgttattac agattgagct gtggccacag gggatacccg tggggtacgc 101640
ttgtccctca ttgcacctga aggcattcta caggatgtca ctgttgctga gtgcagtgct 101700
atgatgagat cctggtccca gacagaagga aagtgaataa gtatgggcct ttcccagctc 101760
tgccaccatg taaccggaga atgtcccttc cccttgctgg ttgctccttg ttccttttgg 101820
tcagctaaga gcatatgtcc tgtctgatgc ccttggcaga cccagaatct tggttatatt 101880
acctctctcc tcagtgattc aggctcttac tcctaaaggc tgcccactgt gctgagcctc 101940
tccttggctc ctggcactta gtgggctctt aaatcttagg cccggtgtgg tggctcatcc 102000
atgtaatctc agcaatttgg gaggccaagg caggaggaac gcttgagccc aggagtttga 102060
gaccagtctg ggcaacagtg agaccccagt ttccacacat aaatTTTTTT ttttaattagc 102120
tgggcctggc agcacgtgcc tgtagtccca gctactcagg aggctgaggc gagaggattg 102180
cttgagccca tgagttcgag gctgcagtga gctgtgtgat catgccactg cactccagct 102240
tggtatgacag acaagacgct ttgtcagtca gtcaaccacc agtcagtcaa tcaatcagtc 102300
ttgtgcagaa ctgtgtcaag accacagctg catccagaca gaatctgaat tcctctgccc 102360
aagtggctca ggtgtcaagc tgcctctcag tccggagcag gggcctgctc ttccctgcag 102420
ctgctcccat ggctggccag gcagcccagc tgggtggcttg attgggccac tttgagactc 102480
acctgagccc aaagccctgc tattccctga gaagcctaga gcagagacca gctgctctgg 102540
gtagaaaact ttctcctgct gaaacatggg gtgaggcgac tccaggaggg ccctggctag 102600
cagcatagga ttccatagaa aaaaggctca cacctgtaat cccagcactt cgggagggca 102660
aggtgggttg atcacttgag gtcaggagtt tgaaaccagc ctggccaaca tgggtggaaac 102720
cccatctcta ctaaaaaata caaaaattag ccgagcgtgg tggcacatgc ctgtaatccc 102780
agctactggg gaggctgagg cactagagtt gcttgaacct gggaagcaga ggtttttagtg 102840
agctgagatt gcactactgc actccagccc gggcaacaga gccagattcc acctaaaaaa 102900
aaaaattact gtgagaggtg gcatcagggc ccaggccagg tggcctgtcc agcctctgtg 102960
aatcctagtg gttcagatga cacgctgccc ctatgattat ggagcagagc cctaaccagt 103020

gtgggagaac ccctcactgg agttctctcc atgagagcct ggtaggagc gtgcatgcat 103080
cttagtcctg caacagtctc ctcacagcgg caatgtcttt ggaggacttg gctttgaagt 103140
accgttatct gttgtggact ttgtctctgt tgctgcacag cagtgtgttt acctttagga 103200
gctctgcaat ccaccagatt aagccctttt aattagcaat ccaatttttag gttgtaggaa 103260
gcctgattta cttaactttg agattgtttt tggaaaacat cttggtgtaa tggagagaaat 103320
attaccttag caagcaggca gagctgggtg caaatactag tctccaccag ccattacttg 103380
gacaaccgtg ggcaaattac tgaacatcaa ttgtcttgac tttaaaatag agacagaagt 103440
aatgatagta atcataatta ctatgcaata ataatagtgg caacaataat tcattattta 103500
ctatgttcgt agcacttact aaggggtgtat aacagtactt agtaagttcg tagcacttac 103560
taagggccaa gtattgttct aagctcttgt atatatattga cttgcagcaa ttctgtgttg 103620
taggttccat tagtgtctct cttttttttt tttttttttt tttttggaga cggagtttcg 103680
ctcttgttgc ccaggctgga gtgcaatggc acgatcttgg ctgaccgcaa cctccacctc 103740
ccaggttcaa gcgattctcc tgcctcagcc tcctgagtag ctgggattat aggcacatgc 103800
caccacgctt ggctaatttt gtattttttag tagagacggg gtttctccgt cttggtcagg 103860
ctggtcttga actcctgacc tcagggtgatc cgcccgctt ggcctcccaa agtgctggga 103920
ttacaggcat gagccactgt gcccggccta ttgtctcttt tttccagata aggcaagcga 103980
ggcacagaga gggtaggtaa cttgcccact gttgcccacc tcataagtaa tgggtctggg 104040
atttaagcct aggtagtctg actctggagt ctgtgtttct tgctccaatt ccaggctgtg 104100
tcacagagaa tcatgaagat tcatgtgatt acatatgaaa caggacctag acaaagtct 104160
tgtgccttgt tgggtgctcaa aacaagttgg ccattctgtc tcacccttc ctgttgctaa 104220
gctaaggctt tagactgttg actgtggcag tgaggaaggt cttggatctc tggctgtgtg 104280
atctgaacac agggtaaggt aaagaaaggt ggggtcacct ttttaccaaa aagtcagaca 104340
gtttcatctc ggctgcggac aggtcatcac actagcaagg tgtggtttgt gcttatggga 104400
gtccttcagt ttcttccctc caacagggga aataattgcc gcctttaccc ttgcctggat 104460
tgtagattgc tgagtctgtt attcagaagt gatccattta gctttccagc tgagccacac 104520
acagacagtg tcagaacaaa ctggtctgaa tttgtggctt cagtggcgga gtcagtaaag 104580
gtggagttga gtggggacag tgctcacaca agtggaatgt ggtagccact ctggagaggg 104640
gagtaggggtg tcaaagggtg gattacctcc actcacgttt tatcccgggt cctagcccta 104700
ccagttaggt agagtaggtc gggatgggca agcccagaga tccaaggact ctgcctttta 104760
tgcacatgca cccctcgccc ccatcaccca ccatgcacac acacagacat acacatatgt 104820
gccagagtt gattagagtc tactggcttt ctgccaccta aaggaaatat tatagaccgc 104880
cctgtactga aggcagtgat tagaaacggg aagttgtgaa atagagatga cttgtttctt 104940

taccttgga atttaaaac cactatgtta tatgtttttc atttataggt ttgtctacag 105000
at ttgtgagcg ttcgaagttg acccctgact aagtatactt tgctgctccc tcagcctttg 105060
aaaaaatgtc tgtcacatat gatgattccg ttggagtaga agtgtccagc gacagcttct 105120
gggaggtaat gcacctgctt tcttccagtg aatgaagtg ccaggccaat agtgtctggc 105180
aggccctttg gtttttttg tggaggttgg gaactgaggt gaacataccc ccaaaccagc 105240
catgcatttg tagtcagacc tagatagcag cagaaaaggt ttggacactt gtttttaggc 105300
agttcaggta tgccttagcc atcatcgaag gcagtgaact taccttgcta ggaaaggctg 105360
atacttaaag aggttagggg tgtttatgaa aggtcttggg gatcaccttc ctttttatgt 105420
ttgccatggg gtgagtggga gagacagaca ctaattgctc agcttatcct acgaggcact 105480
ttactgttga gttttatgta gaaacaagga gggctgtgat ttgacatcat gctggacaac 105540
acaggcaagt ggccagaggg ccaaatgcct tcaaacttgc tggagcatct gccgccagtc 105600
tgtacccatg agtggccctc ttcctgattg gattatTTTT gtttttttgt ggaggcggag 105660
gggtgtgttc ctgagtccca cttacttttc cttcctgcta gttaatggaa agtccatctt 105720
TTTTTTTTTT TTTTTTTTaa cagttgaggt ctcgagagcg caggctggag ttcagtggca 105780
ccatcatggc tactgaagc cttgatctcc tgggctcaag ccatcctccc acctcagcct 105840
actgagtagc tgagactata ggaagtcaat ttttggctta attttgtccc tgtggcaaaa 105900
cccctggagc tttcttggcc ttcattggaaa gaggccttgc atggccgggc gcagtggctc 105960
atgtctgtaa ccccagcact ttgggaggcc gaggcgggtg gatcttgagg tcaggagttc 106020
aagaccagcc tggccaacat ggtgaaaccc catctctact aaaaatacaa aaaattagct 106080
gggcgtgggt gtgggcacct gtaaccccag ctatttggga ggctgaggca ggagaatggc 106140
atgaaccgg gaggcggagc ttgcagtgag ccgagacagc gccactgcac tccagcctgg 106200
gtgacagagc aagactccgt ctcaaaaaaa aaaaattat gtgttaactg agtacctgta 106260
gtgtgctagg caagggatgt gacagactga gttcctgccc tccagaagct ggcttctgcc 106320
acagcagaaa taaatgtgga aaggcacaga gcatgcaggt cctctctgct gccaccctc 106380
ggctgggtggc atccttctgg aggggtgagg ggggcaccca gctcaaattt cagcagccgg 106440
agaggcacat ctgcctgaga aatatttcac ttaagacagc aagagtgaca ggagcagtga 106500
cagccaaaca ggtcgccagc cacatctcac agcatggcag gaggttggcc gtgaaactag 106560
atgaagcagg accagcagtt acacgacgtg gtttggaggt gcccctggaa tagtgctaga 106620
attcaggact gcctttgccc ccgcttgag gtacagggtc acagtaatca ggtcccgtg 106680
gctggcccc ccagagaggc tgtggcctca ttagtgagg agcgcttatg gagctggcag 106740
cagctgtgag agccttcctg tgagctctcc agcctttaga agtgcaggcc cttgttctca 106800
aactgaacc aaactatag atagaaacct ttccagttat ggggagcctc tttttgttaa 106860

WO 2006/119775

PCT/DK2006/000254

ctgtagtgac attgctggct ggcaccaatc attccctgcc ttgagcttgg tggtaggtgca 106920
 gacaggacaa gcacactcaa aaggcccttc cttgtgaggg ataggccagg agaggcaatg 106980
 gggacaggtg aagacattgc tttggagcca ggaggtctgg gtttgactgc tgacctgac 107040
 acttacttct gtgtatggct ttgggcaagt tactgatttt tctttgcctt catttcccaa 107100
 tctataagat gggggtaata acggtaccca tatcctagta ttgtgagggg taaaagaatg 107160
 aataaacgtt aagtgtttta gggcatcagg gcaaggctgg tgggggccct tccctctgac 107220
 acgggagctg agggagtcgg tgtgcacagt caatgtctgt tcctgcggtg cccttggcga 107280
 cattgtcttc attaccact tggtaggcaga aactgctggg aaaattatgt agctaaatat 107340
 agtgctatta cagaattctt cagctgcttg ttaaaaaaca ttcaggtgaa cgttgccagt 107400
 gtatatgtgc agcatatatt aaatcaacta gctttaacaa gtccagtaat taagctctta 107460
 tgctccctgt taaattctca actgcaataa tgtcagtaga caagaacatc tccttttgcc 107520
 tgaaaatgct caggcgggct ctgctcagct tgtgatgctg tccctgaaat tactctccag 107580
 ctggtatcaa ctccgcagtg ctgtggctgt ccgccttctc catgacggat gttcgttgag 107640
 tggcggagtg actgctcagt gatggcctgc tgggatctga gagactcagg tgcacactgc 107700
 caactgcaca tcttcccagg caggtagcaa gttcctcttt cccatggaag gacacttttc 107760
 caaagaaagc cttgtgggtct tgaggtggaa aagttttttg tgtttaattt tactccttcc 107820
 caccaaaaga tactaactgg aatatgtcat tccagagcca tcttctgtgt ggccctgtcc 107880
 tccaggagcc tttggctgtg acatttggag gaataaggta gacactaggc aggtcatcat 107940
 gccaggtgcc gtaggatttg ctaggtttgc ccagggggaa gggcaggaaa ggcagcatgg 108000
 ggtagtggac gttgctgcac ttggatgctt gctggacgaa tgaaggaacc catactgacc 108060
 tgtgggagga gaggttacct acagggacat gatactgaat agagctattt aatgcatatt 108120
 cggaaggagt acctcctgca tcccccaac cctgtagata taggggaaaa ataaaaagag 108180
 gaagagggga gaggggatcc agttcccaca aagaagatgc ttatggagta tgagagcagg 108240
 agggaccacc cccagctctg cctctcagta acatggtctg ctgggcagag tcagcagagg 108300
 aggcctgtgg ctgagtaggg ttgggccaag aattgggaac aagtagaagg gatggcattc 108360
 caggttggaa ccaagtctgg agcctgaatg agtggtgtct tacagagtgc tgaggacctt 108420
 gacctttttg gagctaaggg ccagtttttg agcaaaatgg gaagtcgtgt gggattacag 108480
 tggcgggggc caagtggagg cagttcccca ggtgggtcgc tccctcagca agtcagagcc 108540
 tggctctggcc cgcccacagg cttcacggca gggcatgggg cagcttggct acttctggca 108600
 agttttcctg acttgttgca ggagcccagg ggccctcctc ctgcctttcc acttctgggc 108660
 actgcttcca ggccttctct ttggttcttg gagctggtgt ccaaagggcc tgggctgtga 108720
 gccttggggc ttttcacgtg tcaccagtgc ttctggcct cccttacctt cattcagtag 108780

gctctgggaa caggatggca ggtctgtccc tagcagggct gagagccgct gcctgcatct 108840
 ctgaaggctc cttggggggcc aggctaggag ggcagacgcc acttatttta ggaactggag 108900
 cctgagtctg cctcaacaaa agctccccgg aggcgggatt tctgcacagc tgtcagctgt 108960
 ctgtgcattt agagctcaga tcttctggat gaacccaggc tgagtctgac ctctgtgcac 109020
 tatcttgtgg gtacaaagag cccctgagaa gagtatttgg ggagatgaag agaacataag 109080
 tatcttgtgt caaatattct gggaatgtta gccctgaaga gttggcacgt gactctcaac 109140
 atttaccggt ccctgccagg aaggcgctcc ctgtaggttg ttgcaacttg ctctgagtcg 109200
 ggctggtttc tctcagactc tagcactgca ggacagctca tcacgtgaga acaccttcat 109260
 attcatcatc catgtggagt gatactttcc ccatccttcc ataccaattc tataggacac 109320
 atatttagaa gctgccatga cccggggcct gggctaagtc ctgggcaccc catatgagag 109380
 cccagcctcg ggcctcctag agccagcggt gtatggagag cagttaggag gcggcagagc 109440
 agtgccctga gcgcaagcac gaagcgagcg aatgtcaata ccacctcatg ctcagcagca 109500
 gtgactggct ctgtctgctg cccagccgcc tccacgcacg ccgaccccaa cggacttttc 109560
 cgtgtgatgg cctgcctgtg tcagccttgc aggagcactg tgcccacatc catgggctcc 109620
 tcgggacagg gtgcgcatca gcttaattat caccaccg tgctcacata tggctcctag 109680
 aaggttagt gcccacatta gtttactaa agtttctaaa aattaccatt aattatttct 109740
 tttcactgaa ctagtatctg ctgaaccatg agcaaggatt aaggcatcaa ctctcagtc 109800
 tccacccac tgccagatag cccatctaga aaagcagatt ggatggctcc ctgccctgct 109860
 gggaaagcttc agttttctgc atcaccttca aaatagaagc atactcttta gcatggccaa 109920
 gacacgagca tgtgttgaca gccagctgga ctccctggct cttcctgaga agctctgccc 109980
 tgggctccag ccaagctgag ccaccccagg ctcccatcca caaagatctg tgaccctgcc 110040
 tgggtgtggct tggaggattc tttcctccat ccagacttcc atctccttcc tagggagctg 110100
 ctctgggaaa ggaagagcct tccctgactg ctccgggcgg gcttgggcag tcttcctctg 110160
 tgctctccag catttcctag gctgtaacat ggcattgctc atttcctact tccaggatgg 110220
 ttgtaagaat tacagagaac acccacaag aggctaaaat tgttgtgtgg actttttttc 110280
 tcttttggca aaaactcttt ttcctacccc gatgttacac acaaaatgta gacatgggtgt 110340
 atattgcgta atcttcgttt aaggttctta tgcttatttt ggggagagtt gtgtttctga 110400
 tggatatttg ctgttctggg atttctgaac aggctgctgc agcctgctct tcagagcact 110460
 gtatacctac ctttttcagg aatgagttat attcatgtga attccttcta agccagtga 110520
 aaaatagttt tcctaagatc ctgcccttct ttgggtgggtg accctcccc tgagctatag 110580
 tggacatttc cattgtgggt cagatagact ttgtcgtcct atggggcaga tcagggccct 110640
 cgacgtgtgg gctgcctttg tcttggccat gtttggctga ccctgaatct cgtgactctt 110700

cagaattacg tgttttttaa tggttggttt tcacagttta gaatacacaa cctaagcaat 110760
gcttggacgc ttatgaaaac agacgtcaac attggacaga ctttgacgtt gagattagag 110820
cagaaagcag cctttgctgc ccaatctcta cccagtaatg gaggtttttc tcatgatgtg 110880
atgggtttta accaagctgt cctttgtccc ccttggtcag tcagtcaggc tcatgcctgg 110940
cctgagcttg gggctgggac acagaggtga gtgtgacagg gacagtcctt tgtcctcttg 111000
gagctcacag gccagcagca gcacaagaca gttccacaag tgactgccat ttcccttgga 111060
cagttttcag cactgctgcc atccctgtcc atggccactg tcccctagct tggcagggtg 111120
tgcttggctg gtgttgccag ggcctgggtg catgcacagc acatgggagg gacttacata 111180
cttgaaggag tgggtggcgg tgggtggtca ggtagatatg actgaaggct tcacaggagc 111240
ctctcctgct gtgttccagc agctgcagaa ttggcctggt tcccacaggc ggcttcctgc 111300
ttttagggca aacagggcag ggaccagccc tggccagcac caatcccttt taccatcca 111360
cccaccagga gactgcagg gagccagccg ggcagtttat tttggctggc tgctgcaggc 111420
cgtcgaatct gtcatttttg ttttcagctc tttcatgccc tttggtgtat gacatctaca 111480
acattgtctt cggcctgcgt gctcagccat tcccgatact gggaaaaatc aaattccact 111540
ttcagtggag gaaaagaaag cagaatttgg ggtgcaccgt cgcagggtcag gctcagccag 111600
cagtgatggc ctgctgctgt tctgtgatgg gctgcccagg ccagggacag cacttgctgg 111660
tagttgggca ctgggcccagc ggcaccaatg agcctgttga caccagagt ttattttgtc 111720
tgaggtcaca ggacacctgc atggcaaggc caagtctaaa tgactcccga gcccgagtgg 111780
tctctgagcg tcctctctac acttaccgtt tctgcagcca cacgcgggag catcgcacac 111840
gcggagacag gaaaagggtg atgaatgagg tcacctttcc ttctgttgtg cagacagacc 111900
tgtcttcccc ccgtgtgaat gccatcattc tgaatgcaga gaaagagttc tttgggtttt 111960
gcgttgaggg aatttagcac ttcgggggtt tgtgctgtca agaagcttgt cttctgaaga 112020
taaaattttt taaatttttt ttgtgcattt ctgtagatga agataaaatt ttattttctt 112080
atttgatctc agtgatgtca gaataacttt ttaaaactgt ttaagagaca catttttcct 112140
ttctattcag taacaaatct cttttcattt tagagccggt ggggtccgta gtgctcatta 112200
gactgtggtg ggcttagctc catggagaca tttgtggata aacgattgtt tgagctgggc 112260
acagtgcagg ggggcggagc tggggtgccc aggcacagag atggggagag gagcccacat 112320
tccactagtg gggactggta gggtaacagg agtcataaag taaaccagga ttcgggtgca 112380
gggggctagg gccaggggtg aggtagactg gagtgcagta tgggggggct gtggaagcac 112440
caaactggga agaggaggga acggaaggag gtggggaggc ttcttggaag aattgatcct 112500
ccagcttttc ccagctggcg tttccccact agatcatcgg ctttggacct taggggtcga 112560
gccagttttg gatgtctctc acactccctg tcagtcaggg ggtgccttgg gaattctgcc 112620

aaaagcctca ggtcttcagc gagcttaggg accgggcagg tgatctcgcc atggcagctc 112680
ttccagcaag gtgatgcctg catcagtggg ggcgcctccc tgaggagtga gggagagcaa 112740
aaacggcagc cacataccaa attcctgcca cttgctgggt gctcggctga aggcaagtga 112800
gcagtgcattg tcttcccaga gacccagggc tccaggatgg gagatggcac acagggcggc 112860
atatgctgtg catgtgaacg ccatgtaggg tgctcggggg ccttctcgac atctgtgtaa 112920
cacgcaccaa gggagcagag catggggagc ctggccgacg gccttgagga agagttagtt 112980
caacacagct ttgcgtaaac acatcctttc tttctatgtt gtctatctta tgtgttctgc 113040
ctttggagtt aggatctgaa aattgcaatt tctgggtttc tgtgtcatct attttttatt 113100
ttgtcatttt gaacatttca gttaaagatt cgtaactcag tgtgtcttta tgattgttga 113160
aatctgggat tcatgctttt aaatgatcaa acggcaaggg gtgctttcag ggtagtgagc 113220
tccattttct tcaccttcgt acttgaggat acaggtttca ggaacagagg atacggccag 113280
tcagtgttgg gagcaagaac tgaaaacccc aagttgagac agaattgagg ccccgaggaa 113340
aggacaggac tagtgctggg tcagaatcga gtctggattg tagatcagac tcaaggcgat 113400
tctgaatctc taaccctctt gggccctgca aggtaaatac catcacctca tttgcagatg 113460
gcgagcccca gtgagggttag gtgttagtct gagcttgtgg ttcaaggcag caccgaagtg 113520
tggaactccc agggcctaca ttccagggct gagacaagtt gggagcccag tgggtggatga 113580
gctggtggga agtttgggga ccaggacgag tggaatccca tagaaatgag gatgtggaaa 113640
ttgcagtttg gcacctgaaa gcaaggcacg gactagatgg tttttcaagg acccctgtta 113700
ctctttggtg ttttcaaaca cacatcacag atttcacaac cctaatatca acagtcattc 113760
ccgtttggaa aggaagattc agcatagggg ttaggatagg ttggaagaaa gttgcctgaa 113820
gctgtcacca tggggacatc tggcaggtgt gtacaggata caacatctgc caggcacacg 113880
cacaccctcc ccagggaagg tgggtgagag gtagtgactc tgcagggcac agtcctgagg 113940
agaagggact cagtcctgga gccccctcta gctgcctggg acgtgggttg gcttagggat 114000
aaactaactc aggccctgga tttccctgaa agctttgtag ctgcagtga gggccggtga 114060
ggccctgcag ctggctgtgc ttatgtgact ctctcaggaa tgcggcagaa gcgccgggag 114120
ttgctcttct gccggccatg aagttactgg ctgatgacgt gctaagagtt gattcctttt 114180
agctaccatg gagagagctc gtgtgactcg caggattgag aactgataaa taggccgggc 114240
gcggtggctc acacctgtat tcccaacact ttgggaggct gaggtgggtg gatcacttga 114300
ggtcaggagt tcaagaccag cctggccaac atggcgaagc cctgtctcta ctaaaaatac 114360
aaaaattagc caggcgtggg ggtgcatgtc tgtaatccca gctacttggg aggctgaggc 114420
aggagaatcg cttgaaccca ggaggcagag gttgtggtga gccgagatca tgccactgct 114480
gagatgcacc actgtccagc ctggaagaca gagcgagact ccatctcaaa aaaaaagaa 114540

WO 2006/119775

PCT/DK2006/000254

ctgacagata aaggatcatg agttgttttg ttgaccatgc ctctgagatc tcagcgactt 114600
 ctcagtcctt tcctttccag ggggaatgga aatcccagct cccagcactg tttgggctct 114660
 gagaagatag ggtgtcctct ttctgcctgg catggatgat ggagacatat agaggcatag 114720
 tagttgccct cagaaccac tattttgatg cattttggca tgtgacagca catgtccacc 114780
 agggggaggg gtaatctgtg ggggaagcgt cctggggcag cagagcccc agctcactct 114840
 ccgtgggcct catggcattt tcctcagctg tgagaaagcg gccttgcctc agagctgtct 114900
 tcagctaccg cctcaaacaat cgcttcctcc aggcagccct gaccaccag agtgttcctt 114960
 ctcacaatgt gtcactgtct cattgcatta tttttctggc cataagttcc atgagggaca 115020
 gtcgcatggc agtttggttca ccactatatc cctagggcct cgtggccctt cacacacagt 115080
 ggatgttcct taaatacctc tcagtggcaa gagagtgttg ccagacgctc agaaatggag 115140
 tgatgaatgc ttcagtgggg agccagccag ctggtggcat cagcagactg gaaactgctc 115200
 cctctgcacc acggagagga agggccaggg gcgagccagc tgggtccgagc tgaaccacgg 115260
 aaggaagacg ccctgggtgg gcccctgcag gcaaagccca tggaaccaga gtcgaatgtc 115320
 tgaccagag ctggtcaggt atgccaggaa gcaccaacaat cgtcttcaac cagacaccct 115380
 ggaatcctga aggctcacct cccccccac cctgaccccc agacaggggtg aagaaatggc 115440
 ttagcaagaa gttaggcctt gtcaggcgat cgccgtggca ggcggagagc acagctggct 115500
 gggaagaaac ctaccgtgca gtgtaaaccc ctctatcct gggccatttg gtcctcacag 115560
 gaatagctct gactctacct gacaactgaa agggcagaaa tatacggaga gggacttgta 115620
 caaagttaca cattttgggg cagggccaga tctgtgctgc tcccctcatg agagcaggtg 115680
 aagggagga ggtcttggtt gggctgggaa atccgatgga tttgaacaag ttatcgaaag 115740
 cattttgccc gtctcgttgc caactgcaga cttggaatta aggaatccct aaagagttgc 115800
 ctcatccttg tctgccaggc tgacagagga gatgctagag cttcttggca tatttggtca 115860
 gctttggata aaacatttga ctctggttcc atgggcttgc ctggcgaggg gtacgccagg 115920
 agctcttctg taggtcagct tgggatagct ggctcacct tgacccttcc tctccagggt 115980
 gcagagggag cggttggtgt gcttgggcct tcagggccac agattaaagg ggaagatcat 116040
 cctgaggtta aaatcagggt catccttggc aaccggctcct agaagtttac aggcttgggg 116100
 ttcctgggaa ctttggttgc atatggtggg gacaggagca gggcctgtgt gctgcatgga 116160
 gtctggaatt ctgtttgtga cttgctggcg cccacagggc tggctggaag cacacgcac 116220
 atctcatgtt gacttggtaa gtctcagttt acaccagctc tgatcccggc ctgagggatg 116280
 tttgggcagg tggcctctct ttctgttgca tgaccatgag tgactgcatc tcttcaaact 116340
 tcagtttcct tttctctaaa atagggacac taggtcgggt gcggtggctc atgcctgtaa 116400
 tcccagcact ttgggagtcc aagttgggag gattgcttga gcctaggagt ttgaaatcag 116460

WO 2006/119775

PCT/DK2006/000254

ccctggcaac atagtaagaa cctgtctcta caaaaaatat aaaaaattag ccaggtgtag 116520
 tgggtacatac ctgtggtccc agctactcgg aggctgaggc aggaggatca cttcggacgg 116580
 gaaggttgcg cctctgtgat tgtgccactg cactccagcc tgggcaatac agtgagaccc 116640
 ttctcgcaaa taaaataata aaattttaatt taatttttaa aaatagacat taatagtatc 116700
 ccctggctag cattctcaca agattaagtg ggggtggcatg taaaactgtc agtaccctgc 116760
 ctggccaagt atgggatgct ctattatgac gccgtcgcta gtgaaaatca ctcagtgtgt 116820
 ggcacctttt ggcgggaggc ttacagtcta gctaggggaag tgaggctaata gtgctcgagg 116880
 caatttttgag acgagctgtg acccgaaggg catggggacg gtgttcacaa gggccatcat 116940
 ggcccccagc tggaccagat ggcaggagtc ctcagatgag tctcccctgt ctgtgggtgc 117000
 aggtggtgta gactgggttg gtgctccacg ctggtgctcc agcaaggacg ggctgcatca 117060
 gggaagcctg cgtggggaaa aggagccaag acttcacaga gagctgagat gggggcaggt 117120
 aggaggggga ctcgctcatc aataaccgtt cctgtctcta tttctccacc tgcagttgat 117180
 acctaggccc cctccccatt tctaattgta cagcttttgt tagcatctgc tctcataaac 117240
 ttcgcacccc tacttcatgt gcacgtattc ttaattgatg taaatggcat catgccccac 117300
 atcattctgt ttctgttctg ggtcagtgtg tcggcgtagt cgcgcacatc cttcctcccg 117360
 ctgcactgct gggccccgtg cttacctgtg cattccccag ggagaggccc cttacctgtc 117420
 ctgggaccaaa aggaacact tctatgggaa tctaagccca gaatgaagct gctgggtcat 117480
 cgtagggcac ttgcatgctt tatgatttaa tgactgccta gttgtgcaca gaggcctgcc 117540
 cagtctgttt ccaccagtat atccagtaaa ggggtgctgta caccacctc tgcccacact 117600
 tggcctttcc cactctctgg tggttgccag tccacaggcg tagagctgcc ttgctgttct 117660
 caggcacatc cctctgggca cccacataca gaggtgtctc tgtgtgctct gcctgctgtc 117720
 atcccagcta ctcgggaggc cgaggcagga aaatcgctta aaccgggag gcagaggttg 117780
 caagtgagcc gagatggcac cactgcgctt cagcttgggc aacagagcta gactccgtct 117840
 caaaaaaaaaa gaaaaagaaa aattttgttt aaggagccat gccgtaagtc acaaaggat 117900
 tcgcgttcat tttcttctgg tcatgttatg gttttacctc acatttgtat ggggtgcctg 117960
 ctgtgcggtg gaccctgtgg aaagcaatgg agaaagagt gtggattagg aagacacagg 118020
 tccacaccct cttagggcc agagtctgaa ggtattagg gttgggagca ccacagagct 118080
 tggcatggtc ctgtgtctgc acagcagacc ttctccagcg tccttgtgaa aggcacctct 118140
 ggactgttta caacagatgc atgcaggatg gaggaggagc tgacgtggcc tcttttgata 118200
 gtccccccag cacgcctgtg acaataagct gctgggtttc aggtagagcc cttttctctg 118260
 ctctgattg cacggatgtt tctgcacagg gatggttgag attccagctg ccacctcatc 118320
 ctagctgact ttaaaactga ggcttggtg acattgggtt ctcagaggaa cgattagaag 118380

WO 2006/119775

PCT/DK2006/000254

cctggctcac tgagagggag tgagtgcagc ctggaggagg gcaggcggca gggcgagtcc 118440
tgggctcggg gaagggcatg gggaaggag tggattccct ggaggcagag cttgattgtg 118500
agggatgtgg cccccaaggt ccatccctgg gaggcaggag agcggcctcc ggaggaccct 118560
agaaaggagg ttctcagccg ggcacagtgg ctcacgcctg taatcccagc actttgggag 118620
gctgaggtgg gcagatcacc tgagttcggg agtttgagac cagcctgacc aacatgcaga 118680
aaacccgtct ctactaaaaa tacaaaatta gctgggcgtg gtggcgcagc tctgtaatcc 118740
cagctactcg tgaggctgag gcaagagaat cgcttgaacc aaggaggcag aggttgtggt 118800
gagccgagat cgcaccattg cactccagcc tggacaacaa gagcgaaact gtctcaaaaa 118860
aaaaaaaaaga aaaggaagga ggttctcttt cccaagggtg gcacagcctg gagaaggcag 118920
gggtgctgta tgcagagaaa cctgtctggg agccagggat tcccacgacc tttgtccccg 118980
catcctccac ctggaggatg ggcagctccc ttttttggct cgcatgccct gggttccact 119040
gagcagcccc attctctgga ggctgctctc aattcttcct tccaggatgt cccaccccc 119100
gccagccttc tgagccccac cagtcccaa gcctgacacc cttcaggatg caggcctagc 119160
cttcccagtt tctcagcatc tgccctccct cattgccatt tgctccctgt gttctgtgtt 119220
ccattcccag ctgacttctc tcagggcccc agggttgcca tgtggacaca ctttcctttt 119280
gccagtgtg ttccctgtgc ctggaattat tttcctggcc tcaccttta tctggccagc 119340
ctcctacctg tccttgaaga tgccctttgg atattgcgtt caggatacct ctcaccccct 119400
cctcctagtg cccagttcc agtgccctgt gcccaagcac tgagtccctc agtaggtaca 119460
tgcttggccc tccgtctgcc aggcagtccc tcaggggagt gaagggatgt ccaggcacat 119520
gtatgctgag catttgtctt tggtagctta gtttctgagt ggagttagcc tgaatttaga 119580
ggaagttagg aacatagctt catacctggg tcaaaaatga ggattttgtg aatgaagcaa 119640
ttaatatgtt tggaagaaag catgcttata ttcagtttct acctaaagtt ttctgccatt 119700
tggttatccc aaaacaacct aatgttttga agaagtagaa gagaggtttt taattacaaa 119760
actgatagag cagaaatagt taactatata tgtgctgatt ttaatgtatg atcgatttgt 119820
caaataattg aaaattaatt ttcttgcatg gtaaaaacta gtcattatta tacgtttgat 119880
tttttaaaaa atatattaaa aagggaacag gccaggcgct gtggcttata catgtaatcc 119940
caacactttg ggaggccaag gtgggaggat cgcttgagcc caggaattca agaccagcct 120000
aggcaacacg gcaaaaccct gtctctgcaa aaaatacaaa aattagccag gtgtggtggt 120060
gcgtacctgt agttccagct actctggagg ctgaggtggg aggatcactt gagcctggga 120120
ggttaaggcc acagttagcc atgattgtgg cactgcactc tagcctgggc aacaagcaa 120180
gaccctgtct caaaaagaaa aaaaaggaac aaactttact caagaaggca tacagatggc 120240
aaatagtcac aagaaaagat tttcagtatt attagacttt agaaagatgc aaattaaaat 120300

cacaatgcaa tattgctaca ctaagtagaa tggctaaaaa taacaaaaca gaacaaaaca 120360
ccctggcaat actaagtgct gggagcatgc aggacaggga agtctcatac gttgctgatg 120420
gcagtgcaca gctgtccagt cactttgggt agcagttcgg cagttttggt aaacgtacag 120480
ttaacatttc cattcctagg tattcaccga agagtcatgg aagcttatgt cccacaaaaa 120540
gaccaataga tgaacgttat agcttcttta ttcataattc ccccaaactg aaaatgaccc 120600
aagcaacctt cagcaggaag aatatggata aaccaaagag tgggtataact acccaatcga 120660
atactactta gcagtaagaa agagtgaact tctgatagtg gtatacctac acagtcgaat 120720
actacttagc agtaagaaag agtgaacttc tgatagtggc ataactacac aatcgaatac 120780
tacttagcag taagaaagag tgaccttctg acgtagatag caacatggat cagtctcaga 120840
aacattatgg tgggtgacag aagccagtct caaaagccag gctgcatact gtgtattcca 120900
tatatccttg ccatcctgga aaaggggcaa aaatcatgtc tgtggttgct gggggctgta 120960
tggggagggg atgggctgca agggggcaca aaggcacttc atgagagatg gaagtgtcct 121020
gttccttgct gtggaggtgg tagacagtgg tatatgtttg tccacattca gatctaagtg 121080
taccctgtag tgggcagttt tgctgtgtg tagattccat ctcgacctag agaagcagcc 121140
agcagtagct ctgctcctgg gacgactgca gtgctttgcc cagggaagtg gaaaggccac 121200
attgctctgt ccctggtggt ttatggcgat agagacttcc agtgatggcc ttcaccggcc 121260
accacacaac cgtttcatta aagggacagc agggccgatg ttgctggggg gcctcgggtcc 121320
caactgtttg gcatactgcc cttctcccgt tggctctgtg ctcagccaac ttgcctcttg 121380
tccccagcag tggtcacagg accagcacag gtcctctgc agcctctagg accaactggg 121440
agcttatattg gaaacctgaa atgtgtactt acggaaacac tcttcctgta ttcacacgg 121500
atttccacat ttcaagagtt gtcattgtaa tgtgtatttg cagctcctgt gtgtgccgat 121560
gggtggtacc aacatgtgac ttgttaagaa atatcctttt cttttcccta aacctttaac 121620
cagtggttct gaaaatcaca gagataagca ggtttcaaga gagccaaaat acatgatttg 121680
atggtgatta ccaatttgaa tactgttggt aataggcaga tgaaaaccag ctctggggtg 121740
tggcttcatg tcgactgtca aggctaagtt agggccgtgt tacctcagct cctcctacca 121800
ccagctcaac gatactgagt tgcttggttg gtattttttt catcctgctg taaatttcaa 121860
aatgacagta aggcctcatt ttaagggaag caaccacata atgtcacttc cagtgagctg 121920
tgaaaatgct cacttccgac ttcctcacag gagctgctct agtttgaggat ggatgctcag 121980
gttcagctca tgttttgacg agctgccgag ctgccagcct tgctggggga ggtctttgga 122040
gggaatgagt agagccggaa gtctcatctg cctctggaga aggctgcca ggctaaccac 122100
cttccagggg tctttgcttt tcaccaggat tctggttctc acaggagtgg agtcatctca 122160
gcccattgtag ctgagcctca gggacagccc ttcctccagg tgctgcagcc cccttccctc 122220

WO 2006/119775

PCT/DK2006/000254

cccttgccca tctctgtgca gggccaggca aagagatttt agaaaaggta aaactgacaa 122280
 ggggttgact aagatTTTTg aaggcattgg gagaaaaagg ctgaaaagag atcgTTTTtc 122340
 tgtcatatTT cctctgggtg ctcaagttat cagcgactgt agcctttcaa gccagcagct 122400
 aaaagaatga aacatatcga acatTTtcgt tttcacccca aattagccat taacatttcc 122460
 tcagttactc ttttattact tataactttg atgaactagt ttttcttttc atatacgaaa 122520
 aaaggcagtc aaaactcagc ttagccctga caaagctttg ccatcaagac agtcctactt 122580
 ggaacccttt gaaaaactgc ctttgggctg ggcgcggtgc ctcacgcctg taatcccagc 122640
 actttgggag gctgatgtgg gtggatcaca aggtcaggag attgagacca tcctggccaa 122700
 catggtgaaa ccctgtctcc actaaaaata caaaaactag ctgggtgtga tggcatgcgt 122760
 gcctgtagtc ccagctactg aggaggctga ggcagaagaa tcacttgaac ctgggatgcg 122820
 gaggttgca gtagccgaga tcgtgccact gcactccagc ctgacgacag cgagactgtg 122880
 tctcaaaaaa aaaaacaatt gccttcagct ccagcctagc cactgtggct gcaggatctt 122940
 gacaaaccac gtaacttctc taagcctttg tttcctcatc tgtaaaatgg ggacacttat 123000
 tcctactagg taggggagac agtatttgga ggtgtaaaac acctagcagg tcgcctggca 123060
 catggtaggt gctcaggaag tggaaaagtt tttacagcta tgattgggaa gtaataaccc 123120
 agcacatgat ggatgacatc aaagcagaag gctttggtta gatgtcacct actccaggaa 123180
 gccagcccca gccctcagt ctgcgtgagg tgcccgttta ggcacctca gcccacccc 123240
 ttcctgagcc agtcagcctc ctttcctgac ccacagccca gtttctactt ggtacccaac 123300
 ctaaagaaca ttgagaagcc cccgttcact gatgtgtatg atcttattaa taataaaaat 123360
 agagaggctc attatctaac tacacggagt gcctttcaaa ctcttgagaa ggggtgtggt 123420
 aaacatgacc ccaggggcg ctcagggcct gggtggcatg gaggccattt gtgccacagc 123480
 caaggacagg ctgccccaa cccctgcct cccagctgggt tctctctcct aggtcgggaa 123540
 ctacaagcgg actgtgaagc ggatcgacga tggccaccgc ctgtgcagcg acctcatgaa 123600
 ctgcctgcat gagcgggcg gcacgagaa ggcgtatgcg cagcagctca ctgagtgggc 123660
 ccggcgctgg aggcagctcg tggagaaagg tatggggccc cggcacagggt cctgtggcct 123720
 ggaggcagct acaccccagc cgggggctct gggcacgttg tgaccctctc tatgccagtt 123780
 ttctgtcttt taagattggg gcaagagtgc atttcccagt tccagagcgg ggtgtgattt 123840
 ttatggggag cattagcgcc catcaccaga aggcacaggc tccatgagct tttgtccact 123900
 ttccttgggg acacaggact atagcccttg tcatccatag gacacacttg cacttgccga 123960
 gttaactcag aactcgggta agttcctatc acgtggaagg ccttgccgct tggcacctta 124020
 gtttctaaca ctgagagctg tttattttta aacttgcca thtagcagggt cagataagac 124080
 tagctggcaa cacctggcag tctttctgag tttacatttt aagggaataa ttacagggtg 124140

tgtatggcag aggtttcacc cccttctgtc cccaccgaaa agctcagaag attctgtagc 124200
 tttattagtg ccactgttgt gtggacaggg gccatgctta gaacacactt cccagcaggg 124260
 aaggggaact tgactgatta cagcctttcc ggcacagaag cagagcagtg tgggtttctc 124320
 aaggtgtgtt ttggggcagg ctttcctgtg ctgaccagct gtcctcagcc ccagctgtcc 124380
 tcagccaatc aggggagcca ggtcccaagc aagccaggct tagtccagct gcctgtcctt 124440
 tcccgtgatg ctgacatggg ttgttttgtgt gggcttagca cagtccttag ggcaggggtat 124500
 gtgcccaggg tgtggagcag gcacgtggca gaggaggggc ctaggatagg tcagctccct 124560
 tggtcggtgg cagtgtcgtc atttcgtttc cacctgctgc agcccctctc ccaggatcaa 124620
 gctgtcagtc agagcttgag gtttacagag cacccttacc accctcacct tgtgggtctg 124680
 gctgtaacgt tgtggacagg acaggaatgt gctcattttg tagacacgaa ggcacagaga 124740
 ggtccaactt ccggggctgt tctctgttgt ctccaagccc cagagtcacg tgtccccctc 124800
 tgggttctca cagcctgccc tcccctcagc ccacctctgt catgactcag cgcacttggt 124860
 cactgtccct gtgaccacct tgaagaccag ccatggatct tctcgatttc tgccccggct 124920
 ccgttggtggc tctccctccc tccagcctca cgccacccca ccctaccccc tgaatagaaa 124980
 tgccagcctg gattagggct tccttcctga gcctcgggga ctgagaacag tgagtaggga 125040
 gaggcgactc agcccctcct gccatactgg tccagagcag gtcctgctt cctcctctcc 125100
 ttcctcctgc aggatgggct tccattctgc tgaggggaca gtcctcattc cacagtgttc 125160
 tcaagcttgt gctcagagcc aggctctttc tgggagtga tgaacacagt cctgcagagg 125220
 aaggcagtc agggctagcc aggggccaag cagggtgat aggagccatg ccaggagaag 125280
 cgttggcatg gtgtgggtgg gtgccaaca cctgggagaa gcatctgctc acttcagaag 125340
 gtccccctt ctctggccac ctttaagtaa attaagactt ctaagatgtg gaaatttgag 125400
 tttttataga gttcagtggc tttaaataaa ggaaaggcat ggccgggctc ggtggctcac 125460
 gcctgtaatc cctgcacttt gggaggccga ggcgaatgga tcacctgagg tcaggagtgc 125520
 gagatcagcc agcctggcca acatggtgaa acccgtctc tactaaaaat acaaaaattt 125580
 agccaggcgt ggtggcgggc acctgtaatc ccagctactc gggaggctga ggcaggagaa 125640
 tctcttgaac ctgggagaca gagtttgag tgagccaaga tcacgtctct gcactccagc 125700
 cgggacaata agagcgaaac tccatcccaa aaataaaaat aaaggaaagg catgtaatcc 125760
 cccaattttc ctaacacgag gggcagagca aattcagctt tgggaactga aaaaggcaca 125820
 gagtccctga gagaaaaata tctgaaacct tgaaaggaaa ccctctgtgg tctaaatggc 125880
 cacattgaac agccgaggtg tgagcagagc agaccccacc atggccggcg ctgagtgacc 125940
 cgcagcttct ctcccctgca cagggcccca gtacgggacc gtggagaagg cctggatggc 126000
 cttcatgtcc gaggcagaga gggtagcgga gctgcacctc gaggtgaagg cctcactgat 126060

WO 2006/119775

PCT/DK2006/000254

```

gaacgatgac ttcgagaaga tcaagaactg gcagaaggaa gccttttcaca agcagatgat 126120
gggcggcttc aaggagacca aggaagctga ggacggcttt cggaaggcac agaagccctg 126180
ggccaagaag ctgaaagagg tacatctgtc ctgcaggctt ccctgcccgg ccttgctctg 126240
cctggcacc cactagcagc accacctcct gcacctagga cttcctgcca ctggctgtag 126300
gcacaggagg ccagacctgg tagagatgaa gccactcctc aggctagaag cccatggtag 126360
ttgacattta cccagggttg agatgctgaa tgctctatgt gcattcttta aattttttat 126420
ttttattttt tgagacaaag tctcactctc tctcccaggc tggagtgcag tggtgcaatc 126480
tcagctcact gcaacctgtg cctcctgggt tcaagcaatt ctgcctcagc cctccgagta 126540
gctgggacta ccggcacatg ccaccacacc ctgctaattt ttgtattttt agtagagacg 126600
gggtttcacc atgttggcca ggctgggtctc gaactcttga cctcagggtga tccaccacc 126660
ttggcctccc aaagtgctgg aattacaggg cattctttaa atttgatcct accaacacat 126720
tttacagatt tggtaactga ggatcagaac catacattgg tggtagagct caaagcccaa 126780
gctcttaacc actaacagta cttctcccaa ccctggggaa gtgtctgctg gagacttaaa 126840
agtccttgcc gggcgcggtg gctcacgcct gtaatcccag cactttggga ggctgaggca 126900
ggcggatcac gaggtcagga gattgagacc atcctggcca acatgatgaa accccatctc 126960
tactaaaagt acaaaaatta gcagggcggtg gtggcgggca cctgtagttc cagctactca 127020
ggaggctgag ggaggagaat cacttgaacc caggagggtg aagttgcagt gagctgagat 127080
cacgccactg cactccagcc tggtgacaga gggagactct gtctcaaaaa aaaaaaaaaa 127140
agtccctttg gctgctctgc gtgtgcagta gcctttcttt attcctttac tttctcaata 127200
aacttgcttt cactttattt ttttttcaag tccctttgac gatagaaaga actagtttac 127260
atttttcttt ctcttgcata ggtttttatc ccaaatttc aaacattgat acctgttttt 127320
ctaaaagtag caaccaaagc cattgagttt tgtcgagcct tggctagttg gaggctgtca 127380
tctgtccacc ccagccctca gcttgtgtgc cccgcccctg cctgcagcct cgccccccat 127440
tcctagctgg aatggagtag acggtagatg tgccaccccc aacaccaga tgttctctgg 127500
ccttgcgctg cttcctccg cctcagaatc atccagaaca tgctctgtc cttccggatg 127560
gttttctctg tgtctccttt tctctgtgtc ttagcttttg tcacatctcc tttctcttct 127620
gttttgctcc tggccccatt cataatgaaa ctgtctgtgt ctcagaggc ctttgctaat 127680
aaactcctcc ttccctctag aatgaaattt gtggcttttg ctcaaataag gaaggatctg 127740
ttaagagtgt gtgtgtgtgt gtgtgtgtgt gtaaaaacca ttaaaaccag aggaaatgtg 127800
gctgatggcc ctgcccccta gtatggtcct ctctctcatg ctgggttcct ccttggaact 127860
ttctttgtat ttgccgatcc tgcggcctgg catggtcttc ctgggatgcc cataaccct 127920
ccatgtgcgt gttccatgtg cgtgctccat gtgcgtgttc catgtcatct cagtggcttt 127980

```

WO 2006/119775

PCT/DK2006/000254

ctccacccac cccatgctcc gtctcagttg tcctccccta cacttaatgc cctctaattg 128040
actctggcat gtaacttttc ctatactgta tctgcctcct gcaactcaaa tgaaagcttc 128100
acaaagacct tgtgttttctt gctgtgttcc tgggtgcctcg aaccatgcac aggacaagca 128160
agtactccat gaatatgtgt tgagtgcggt cctgtttgtga ctggttgcta tgatagtata 128220
catacaattt tgtacctgcc tttttctccc ctcttggtag cacaagcatt cccatgggtct 128280
gtgtgacttc tagtgaggat cctgaactcc cagtgcacatg gcattcagga actcattctc 128340
cctgtaggta gaagcagcaa agaaagccca ccatgcagcg tgcaaagagg agaagctggc 128400
tatctcacga gaagccaaca gcaaggcaga cccatccctc aaccctgaac agctcaagaa 128460
attgcaagac aaaatagaaa agtgcaagca agatgttctt aaggtaaact gtacacttgt 128520
ttttacacgt cgagtgtcac cgtgtcagtt gttgacgtgt ctggcatata gccttctgta 128580
actgcttggg tgaatgggtg gatagaggga taaataacca ataaaacaca cccctgtcct 128640
caagaaggcc acaggtgagt gccctagcac ctgcctcac catagccact gtacctgctc 128700
agtcaaggga cagagccagg ctctgggcaa gccaggggta agagacagct cttggtgctc 128760
agagacacaa ggacagctgg gattgctgca gggagctgct tgtgccaaca tgggcgtcgc 128820
cagtgggtggg tgtgaatggg tgtggctggt ggcaggtgtg aatggttgtg gccagtgggtg 128880
ggtgtgaatg ggtgtggtgt gagctggtat ggagggttat gaggaccaa tcaccagtgg 128940
cctggacagc caggctgagt aaggagctta ggatgctgct gaccaggagc cactcaggaa 129000
ggcctgagcc tgggtgggat cagggaggag gtgaggaggg tgagatgggt tagaaggatt 129060
gctctgcttg ttagaggaga ggggctgggg gcagagacct gggggaggga ctgcaggctg 129120
ggagcgcttg tctgggggag ccaggtggga gggaggatgg aggggccctg cagtctccac 129180
agtgcaggc cccgaaagtg aggcagaggc tgtggagcca ggggtggcag aagggtgggtc 129240
tttgccaag gagagaccgt gaccaggcac tgtggcatca tgggtgccct tccatttctg 129300
ttgtgaagcg gagattttct gcccctatggg ttaatgagaa atcatttgtgt aaactattac 129360
atgctccgca ggtgtagata agtgattctg cctggaaggc cctggagtca caggacttaa 129420
aaccagacag cgtcatcgtc ttgtgattta aaagttcaca ggaaggagg ggctcagccc 129480
tcggccccac ctcccctcct ttggctcagcc atgaggggtg agggctgggg ggtgtgattg 129540
ccgagggcct gccctgcctc ggtgtttgga gaggagattc ccaaccagga tttgcagaga 129600
gagacggaat gaggtagaga gaggctgcac tccaggaaga ggacatgggc agaggggagg 129660
caagaggcct aagctgggtt cactaccac cagcctttcc tggcccctgt cagttagggtg 129720
gtggcatccc tgttttgtga atgaggaaac caagatcaga gaggttaagt cacgtgcat 129780
ccagcagggtg agagcaggaa cttctgacct caagcccata catatgccct ttctccactg 129840
ttcctcactg ctatgccgtc ggtgtagcaa ggggaacacg tcattctacac gagagtctgt 129900

WO 2006/119775

PCT/DK2006/000254

cactcagacc aagtgtcaga atcctctggg aaacccctct cactcccatt gctggagcca 129960
 ggaatctggg cttttgaagc agcccaggtg cctccacggt gggatcaggt gaggagcctg 130020
 gaaccaacta ctgagtggga ggcacctggg ctcaagaagc acaagggtccg ggagaactct 130080
 cgggcctttg ccagaatccc agggaatcag gtgcccacaaa aggtgcctag aggaaagatg 130140
 gggcttctcc atgttgggcg ggggcgtgcc tgtgtgagag aaagggagac agagatggag 130200
 aggaagaggc ctttcatggt tgacagccac ccagaacatc aactcaggct tttgcagaac 130260
 aggcccaggg ctcgagggtg gacataagcg cctggcagga tgactgggac agttttagt 130320
 gggttccctc ctgggaccct ccctgcctct ttgggtctggc cagtgcctgc aggtatag 130380
 aggtccttgt gtgcctagag ctctggagag gaggccagtc cctccccacc acccttcttg 130440
 acgcagtttc caaaatccag cacacagcct gggaaacata gtgagacctc atctctacaa 130500
 aaaatttaaa aacttagttg ggtgtggtag cacgtgcctg tagtcccagc tactttggag 130560
 gctgaggcgg gaggattgct tgagcccagg agttcaaggc tgctgcaagc tgtggtcaca 130620
 ccactgcact ccagcctggg tgacagagtg agacctcatc tctaaagtag gtaagtaggt 130680
 aggtacatac atacctacct acctacctac ctacctacct acctacctac atacatacca 130740
 ttgctgactg gttgcacatg tactttcatg tttcttatca gccatcaact acatcttgta 130800
 gcatccatca taaaagaact ttttgcattt agtgctaaaa tattacaatc cacctaaggg 130860
 tgaaggccgc ctcttgtttt actccatggg tgttcgctga gaactttcta tgccctgcaa 130920
 ggggatttcc agaacagtga tgaccctgct gagtacctgg cctgtgctgg gggcaggact 130980
 tggggctagc agggtcggca caggctgagg gcagggggcag cccagaggtg gagcatcttg 131040
 ttgccactga gtaccctgag agctcccggg aggctattgt gtcttgaagg ctcttgagat 131100
 ggggctgctc tgtgtcactt cagcagtatc ctccagatgc tgggcgctga acccacagct 131160
 atgtgtcttg ggcttctgag ctaggctccc atgcagggtg gccaggtagt gggcacgggc 131220
 agtgtggatg gctgtgggga ataggggaat agccacagcg ctgtggagct ctttctctgc 131280
 tgggcatgcc ccagccctgg caggggtggc cttttcctgg cacagaggac ccaagcccag 131340
 ccatcttagg aggctgagaa gcagagcaag gacccggata ccaggagagg agaaaggggc 131400
 aggagtcagg gccctagagg gcgggggtgga acctggtctc acgaggctgg ggaaggcttc 131460
 aagaagagga gttgctttgg agcctgaagc agagtctccc agggctctga gccacgtgag 131520
 tcagaggaga gcatctggga ggggagccgc atgggcaggg tagactcagt cccagctggg 131580
 ggcagagagg gtgcccagata ggcctgcagt gtctcgggaa gcgctccatc cttcctcaga 131640
 aggccctgca catgttggtta gggacagcaa agcttagacc tgagtggagc ccgtgactac 131700
 ccccatagta gaaagaattg caataaatag gtggccccag ccctgctatc tctcaccac 131760
 cccaggggga caacacgacc atgatccac caactaaca aaccacctgg ccaccaact 131820

WO 2006/119775

PCT/DK2006/000254

ggtggccctg catgctgaga tgaggggggtg cggggaagtg tgtaccagtg tgaggtcatt 131880
 gacctcgagg cacatcccca ccagatgaat gtgcccactt tagccatgtc ttaggaattt 131940
 cacacgtttt ctagctgccc ctggatattc ttgagatagg ccacacctgc tttttaacat 132000
 ttcttgcctt gtaaaacccc ttcttgggac aatctgggag ggagcagctg tgtgaggggg 132060
 atggagtggg ctgaaagaca agagaaggag gacttttttt gaatgccttc ttcttgcacc 132120
 ccatcatgta catatcagaa cactgtgaag gagagatgtc ctcatctgat aaaggacagt 132180
 gaggctccaa gagagctgca ggattcagcc ccagccccac ctttgctgtg ggccagatct 132240
 gaggcttgag ggtagccagg cacagccctt cccacaggg agaagaggac atgggcgtgt 132300
 ggcacgctca agaccataca cactgagttc tggggagact gccctttgcc ggcatctacc 132360
 tgtctgggcc ttgctccagc tgagagacag ggcagatggc aggagtgtga ttcaccttgc 132420
 aggagtggag gaactgggac ttgaacttga agtctgtcag aatccagagc tcctgttttt 132480
 tctttgcatt tggaatcttg gaaggcctga gacactccag gggcaggggtg agccaagggc 132540
 taaatggaaa tgaacctctg tcttgtgcct ggattttcct ttagaccaa gagaagtatg 132600
 agaagtcctt gaaggaactc gaccagggca caccacagta catggagaac atggagcagg 132660
 tgtttgagca gtgccagcag ttcgaggaga aacgccttcg cttcttccgg gaggttctgc 132720
 tggaggttca gaagcaccta gacctgtcca atgtggctgg gtaagtttcc tttgagctaa 132780
 aacacacacg tcattgaagt caacggaaag tgcttttgaa gcaaactcag atgctgctcc 132840
 caggtttggg atggcgcgtt catcaggagg tctcctcacc gccctcaggg cctcctgctt 132900
 ctttgtaatc tggatggcag ggagggcccc tctgctgagt gtttctgcag ctgcaggcat 132960
 ggccggctgg acctctggaa ggcttggttg acccggccaa ttgctggccc aagcagataa 133020
 ggccggggcc actctgccaa ctccatgtac aggtctgggt gggaatggag attgcccatt 133080
 ctggggccac agctaggacc aaggcagctg gcatactttt tcttttcttt tttttttttt 133140
 gagacggagt cttgctttgt cccagggctg gagtgcagtg gcgtgatctc agttcactgc 133200
 aacctcctcc accttctggt ttcaagcagt tctcctgcct cagcctcccg agtagctggg 133260
 attacaggca tgcgccacca caccagcta atttctgtat ttttagtaga gatgggggtt 133320
 caccatgttg tccaggatgg tctcgatctc ttgaccttgt gatccgcca cctcggcctt 133380
 ccaaagtgct gggattacag gcgtgagcca ccgcacctgg ccgcagctgg catatgtaag 133440
 atccactgga ggtgaatctg gcatttcttg gctggtcgtg cctgggacag tctttctttg 133500
 aggccctttg tgatcagggt actgcttagt ggctctcagc cccaggaca gcacataccc 133560
 tctgagactt ggttctgaat tgtgatatgg aagctgtgct gtttcttctg tgctttcctc 133620
 atctttgggt ctgtgatgcc ttcaccaggg cagaggtccc agctaagtca aggacgtgga 133680
 ttactgtcat ccatctagca gaaccgtgta gggatgcgtc acgggtcaag gcttgctgtt 133740

aataggagac acattccggt tgtctccctc tgtgtaacaa tcatcaaaag caacagttac 133800
tgcaccgctg ctgctctgcc accattccat ttttatgata gtcttgtaaa gtgggtggca 133860
ctgaccccat cttattgatg aggaagctga ggctgggtga gattaaatga cttgcccag 133920
gctgctctgc tggcaggagc agggtagaggc tggcctgtgt caggctcttc tgcattgtga 133980
gtgactccca tcgtgtggat gcttgtctgt ctctgtgtct ggatttcccc cttttaatac 134040
aggcaccagt catatggatt agggcccctg ctagtgacct catttttagct tgattacctc 134100
tataaagacc ttattttcaa ataaggccac attccgaggc actgggggtt ggactccagc 134160
atgtctccta tgaaggacac aattcaaccc ttgatggagg ggatcatgcc cactcaacgg 134220
aggaggaaac cgaggcccag agaaagtaca tgactcacca gccataaaga ttataagtga 134280
aaatatcagg tccagcagtg atccatactc ttccctggca gggcagcttc ctaggaaggt 134340
gagccgatga tggctgccag gcggaaggca gccccaccc cgtcttggtg gtgggtgtgt 134400
tcagcaccgg gtcattctca cgctctgcat tggcaagtct gggataaag cttttctgga 134460
cagagatgag gaaggagggg tactttctaa caccctccag ttttctcttc ctcaactctg 134520
ggcctgggtga aatgggacca cacagagtgg tggcccacac agcccacctg ctggtgcctt 134580
ttgcatatgt gacacagagc agggcttagt cattgccacc tgcttggtg gtttgattta 134640
aagctcattt gtcaggagcc atctgtgcac gtggcgggcc gtgaccttc acctgagcca 134700
agggaggagt gttctggctc tgtctgtgaa accaactcac agagggcaga gaccacggga 134760
cctgctgtgg ctggcactgg gctgggacaa agggtagctg gccctgcct tctaaaagag 134820
ccacgtttct ctctgattta gctacaaagc catttaccat gacctggagc agagcatcag 134880
agcagctgat gcagtggagg acctgagggt gttccgagcc aatcacgggc cgggcatggc 134940
catgaactgg ccgcagtttg aggtaaagag aggggtgtgcc catgagctca tcagcctgga 135000
agaggacctg ggcccacaga gttgctagat gatgtctttc tcattctaga ctctatgtta 135060
gtgtttgccc tgtttatcag gcagccttta ggcatagcta tagtacagac caagttctgg 135120
catagagtcc tggggggctg caggcattac cgaaggcctt ctgaacctct cctgaagtcg 135180
cggttatgtc ctctctagtt gggaatgttt gactgtcact ttttgggcac tggcattccc 135240
ttccttccta gtcagtctca ggagccgtag gcacacagga cctcagggtcc ctgaccacc 135300
ctgtgggatg gccaggaagg gagcagccac agtgctcccc tcacgggtgg ccaggggagg 135360
ggagggtggtc tcagtgggca gccctcccta cccttccaac aaaggacttt tccctactcc 135420
aggcagtaaa ttcagtgaat cctgattcaa aatgaatgtt cttcttttcc catcatgtgg 135480
gtctaagctt tcattatctt tacattcgaa atgcaagggc ctggggatct caggactcac 135540
ctagcccttg gattggtcac taggccccac ctggcaccct gaaggcagt ggcaaagggtc 135600
ctggtgcccc tgaggcttaa gcacgtggcg tcagctgccc cagcagggtat gctgaggggc 135660

actacctctg ctgtgggcgt gggcctcgtg gcttccccga ctcagcagcg gccagagccc 135720
gggctgcatt cagattctgc tcacctctga ccacctgtgg gaagctgagc aaactattcg 135780
ccctttttgt gactcagttt tcccacgtgt ggaatgggaa tgacaagagc acccatctcc 135840
tgtcgggggc tgtagggaa catggtgact ttgtccacgc agggctattg acccaggccc 135900
accactcaga aacctcctgg gcgcctcata gcctgcagtt ttgtcagtcc ttactcctgc 135960
acttctttta tggcccttaa caggaagggt tttccaaatg acagtgtggg ccctgctgcc 136020
gagcactccc ctgcgccagc cccaggactt gctcccagta ggaggcagta gagcagctgt 136080
gccccctcca cagaggcact gagagagcat ggagggccag agccgtcctg actggaagcg 136140
gccctagagg ggacagcaca aagcccagcc tctcctctgg agccacacac caggaagtct 136200
tggctgacgt gttttgtgca cagctgggggt ttcagacaca gcgaggattt gctgctcccg 136260
gagtggcaca ggaaaccacc tgtgaaattc gcctgggtcc cttagagggtg cccattatac 136320
ccaaagaggg gcacagtggg gaggactatt tgacagtctg tctcctggag ggtgtcccag 136380
ccagagccta gggggtcctt ctctctgtgg gtccatgggc tgggaaccag cccctgtgca 136440
ttcattctgc gcacatttgc agctgttacg ggaccatgcc tttctctaata gcaattcttt 136500
gtttttatatt ttcattgccac aaggatgaag taagtgtctg cgttcggcta cctcatagag 136560
tgccatacgc catgcagggt gtctgtggcc cgcattgggtg ggtgtgctgc ccccgctagg 136620
ctcctagcag ttcagctgag gctgcgctgc tgtcaggggt attgagagcc ggcccttccc 136680
atcagtgccg gcttggtggt gatccattg gcctttggaa gccaccaat caggatgaat 136740
ttatatctcag aggtcactgt ggagaactct gaggttccaa ttgagttctc agtcaatcta 136800
ttttggctcag aaaataatta atagatgaaa aggtgaccca gatttgggggt cgggcccag 136860
tctgccaccc aactgcgtg ggcaaatcac attgcctttc tgggcctcag tttcctaatac 136920
tgacaaatgg acaggttaga caggatgctt ctgagttggc ttccaacca aaagttctgt 136980
aagtacgtgt gagcagagcc gtggactcag ccacaccaag ttttaatcag agttgtctgg 137040
taactggtgg agatgggggg aaggagagct cattaaacac ctgcgggcat gcacagactc 137100
ttctagaaaa tcaagacaca gtctccattc atggttttgc cctagtccat tttcagtgcg 137160
gttggtgaag ccagttcgcc acccgagaac tcgtgaggct tccaggagga gcgtactctg 137220
ctttgcttcg ttagccccag catgtgctca gatcacagag gaaaagtccc agggcagcaa 137280
aacactgggg agggccctgc gctttggaga ctgagagtct gcttgcctgc cccttctgtc 137340
agcgggtagc cacctagccc tggtttttatt tgcagtggct ttattttgcac ctggttttat 137400
ttgcacctgc ctttccacat ggcaggcttg ctcaggctctc cgaagccaac tgtcttgacc 137460
ctgtggccca gggacaattt agggcatcga cgcaaaggat cttcagcact tcaggacagg 137520
tgccctcagag gctggctcct gtctactttg ggaggctcag cagacatctg gggaggccac 137580

aggggccccat gctggcctga ggggggctggg actgagaacc ttggtgggtg gggtcaggaa 137640
 caggaggccc cggaccttgc cctgggtcctc gccacccccg ctgggtcctct gggaggtggg 137700
 cagcatggcc tgggggtggt tgtagaggtg caagcctcca gatcccctga tgagcccctg 137760
 ctttgaacag tgaggagggg ctcaggtcgt ttgggttcat aagtggggca tggggagcct 137820
 gaggtcttct gggaaaccag ggagagctgg ggccctgagt ctacacatca gatgcacaga 137880
 gccaactgca ggctcgctgg gcaggaactg aggacagacg tggcttccag caagtagtga 137940
 acctttgggtt tctctccctc ggcacctgtt gcctaggagt ggtccgcaga cctgaatcga 138000
 accctcagcc ggagagagaa gaagaaggcc actgacggcg tcaccctgac gggcatcaac 138060
 cagacaggcg accagtctct gccgagtaag cccagcaggt gagtgggcbg cacctccatc 138120
 caaaagaggc cgttccatgg tgggggcccct gggcaatctg catggtccag ggcagccccg 138180
 gcctggacac ctgggaggtt cctgggctcc tgagctctgt gggccgtggc ccagcctgga 138240
 cacctgggag gttcctgggc tcctgagctc tgtgggcbgt ggcccagggg catctccagg 138300
 gaaggctcgtc ttcctgcctg gtgccagagc tgctgacaga tccagagagg gaggcttggc 138360
 cctgagcttc agacctgagt ggggaggagc tgatgacagg gctggcattc atgtccccag 138420
 ggatagcacc gccctgcccg gacactgggt gaggtgccac cctgggttccc tcaggcagct 138480
 gaggtcctgc cgaggttcat gctcactctg tattgtgtgc aggcgccagt cgagggaggg 138540
 acaggagagg actgccaggc atctcctttc caaattacct cttcactagg agcattgggt 138600
 atttctgact tagtcattgg ctgaaaatga gagatataag ggtggcactg tggtcattag 138660
 aaccttctgg aaagtgtgtg ctagaggagg tcttcctata ggacctgccc acaggcctcc 138720
 atgtgacact cagtaatagt catttcatac gggggcttcc taagaaggct tagcccaggt 138780
 tcccttggcc ttcctgcagc cccaggcact gaagccccca cccaggctcc ccagcaggtg 138840
 tggctgtcgt cattccgctg ctttgcacgc tgtctgggca ggtgggctcg gcagtaagtg 138900
 cctgtcccag ggatcaggcc agcacctgta tgcttggggg tgggaatggg aaagaagaaa 138960
 cactgcctca gaggagacc tgggggcccct gggctggcca gggttgctgc acaatccagc 139020
 cggccccacc agtgccgtct cacaacagtc tgctagagaa aggtgccaag gagggaggag 139080
 ggccggtgag ctcagacatg tggcctggat attggcctcc agcaggccta ggtcaagtac 139140
 cacttgccag gctgaagtgg atgaggcagt tactggccac tcaggatcac cctgggtggca 139200
 ttggaggcat gtcctgcct cccttccttg gttctggggg ggtgggtgct caggctgagc 139260
 agtccatcca gtgctttggg ggccctgtgcc ctgaggccct ggtcttgga gtagaagggtg 139320
 ctgctggcca gggtagcccc gtcctaggat cagttacttg gtgggggtgcc agggcccttg 139380
 gtgcctgtct gattggattg cttgatggtc tgctgtggt cttcgtagt tgtaactcag 139440
 ctcccggctc ccagccctca ggcagagctt ggcccagatc tctctggttc ccacactggg 139500

cctggcaaac agtaggcagt tgcacgactg gattttccct ctggagcacc aaaggaatct 139560
tctttctatc tagaaaagac ggtgccgcag gccctggccc tgattttgga ggcttcagct 139620
ccacagagca gaagagcgag ccctgaaccc cttagtagtc aggaggaggc cgtgtgggct 139680
gaagagctca gagccctctc ccactcatcc atgggagccg gcccctccca cactgtcact 139740
ctcagaggag accctgaggg gctttttgac ccccccagct cacattggtg ggagtcagga 139800
cagaggagaa gccctgaacc tggaaacgac ttttgctcca gctgtggctt gtcacccgcc 139860
cctggggccac cccagggccc ctgttctgga cggactgtct ccagggtgcaa ggaaggaaaa 139920
gggaaaccaa gtcctctcct aggctcctgg accgtgtttc acagctgaag tattgcccgc 139980
attctgccat agtaaccttt ctgttgaaac accacggtcc cctccggtca caggaggagc 140040
tgatttgtct cttgagatcc tgcagctagg agttgccagg accctcacgg gccctgcag 140100
agctgggatc ccctgaaagc tctctctctc catagcacc ttaatgtccc gagcaacccc 140160
gcccagtctg cgcagtcaca gtccagctac aacccttcg aggatgagga cgacacgggc 140220
agcaccgtca gtgagaagga cgacactaag gccaaaaagt tggatgaaca atgcttcctc 140280
tgctcccacc gcgctcactc tgtctctctt cctctctctc tgtccagggg tctcacccct 140340
cggcactcgg caatgcacgg agctctgcag gaagtgggca cctggcgtgc cgcaccctgg 140400
cccactacct gctctctggc acaggctctc tggcaccccc attcctcggg gagggcacag 140460
ctaacaatct taagttgctt tttctctgtc ctctgttaac ctttcaaggc cctagcagca 140520
ccatgacagg aaaacagaca acgcgcagtc tgtgctgggg actagagagt gagggcccag 140580
gttggcctct gctgtgggca tctccagccg cctgcccctg gaaagcctgt tgatggctca 140640
ggatgagtgt ggatggcatg aactcacact agcactggct ccacctgtgc tccttcggaa 140700
ggagccctca gccctgagcc ccaaggctcc agtgccctcc ccttgcgggg cgggggcctc 140760
acaccctctg ccgccctgtg ctgagggccc tgtggcctcc cttttcttgc agtgtgagca 140820
gctacgagaa gaccagagc tatcccaccg actggtcaga cgatgagtct aacaaccct 140880
tctcctccac ggatgccaat ggggactcga atccattcga cgacgacgcc acctcgggga 140940
cggaagtgcg agtccgggcc ctgtatgact atgaggggca ggagcatgat gagctgagct 141000
tcaaggctgg taggtgggct ccccagcat ccaggggagg agggcttcca acctgcaggc 141060
tgtgggcggt cttttttgtt ttttaagtct ggaaaagttc ttgctagttt cacttgtgag 141120
tactgcagg cccttcccca gtgccctggc ttcatgtca aaagcagcca attggctgga 141180
tgtggtggct catgcctata gtcccagagc tttgggaggc cgagggtgaa ggatcgcttg 141240
agcccaggag tatgagacca acctgggcaa aatggcaaaa cctgtctcta caaaaaatac 141300
aaaaaattag gccaggcatg gtggctcacg cctgtaatcc cagcaccttt gggaggccaa 141360
ggcaggcaga tcaattgagg tcaggagttc gagcccagcc taaccaacat ggtgaaaccc 141420

tgtctctact aaaaatacaa aaatgagcca ggcgtgtggc ggggcatcta taatcacagt 141480
tactcaggag gctgcggcag gagaatcact tgaacctggg aggtggaggt tgcagtgagc 141540
caaggtcaca ccactgtacc ccaggctggg tgacagagca agactttgtg tcaaaaaaaaa 141600
aaaattagct gggctaattt tagtggcaca tgcctgtagt cccagctctc tgggaggctg 141660
aggtgggagg atcacctgag cccagtagat tgaggttgca atgaatcatg attgtaccac 141720
tatactccaa cctggacaac agagcgagac cctgtcgcaa acaacaacaa aaataaataa 141780
cctgggcaac agagcgagat cctgtctcaa ataaataaac aaacaaaagt agcagattag 141840
ctgggcggtg tgttgcatac ctatagtccc agctgcttgg gaggctgagg cagaggatca 141900
cttaaaccac agaggataca gtgagccatg ttcataccac gtactctagc ctagaagaga 141960
gagcaagacc ccatctctta aaaaaaaaaa aaagtaggct gggcatgggt gctcacgcct 142020
gtaatcccag cactttggga ggccgaggca cacagatcac gaggtcagga gattgagacc 142080
atcctggcta agagtgtgaa acccgtctc tactaaaaat acaaaaaatt agccgggctg 142140
ggtggcagcc gcctgtagtc ccagctactc gggaggctga cgcaggaaaa ttgcttgaac 142200
ccgggaggcg gagcttgcag tgagccgaga ttgcgccact gcactccagc ctgggcgaca 142260
gagcaaaact ctttttcaaa aaaaaaaaaa aaagcggatg cccaatgcca agactcagaa 142320
atagtgaggt gctgccaagc cagtgcctgg actcctcacc ttgctgggcc aaggctgtga 142380
attgcagggc actggctagt cacacagaag ggcattctga gaagctgttg gccaggacca 142440
ccaggttccc ccatatgtag tctgggagct gctagccccg cagggccatg agagaagcct 142500
ttctctaggg gtccccaact gccagagaag ccctcaggag agccacagct ttgcccagac 142560
cccagagtag cacttggcat gcccctgggc cagggtggggc cttgtctcag gcagcaactt 142620
gtgtgtgtgc acacctacct gcccacacct accatgcaca caccaacca cacacatata 142680
tgtgttgcta ggagcagaaa ccatagtttt ataaaacagc ctggatcctg attacatgtg 142740
ctctgcctgt gaagactttc tgtgcaatgt ctttttcttc tttgaatggc cttcaatgat 142800
aatctttttt tttttttttt ttaatagaga cacggtctca ctctgttagc caagctgggt 142860
tacagtgggt cgatcatagc tcacagcagc tgcagccttg atctcctggg ctcaagcagt 142920
cctcccacct cagcctccta cgtagctggg accacagatg tgtgctacca cccacccac 142980
tataggaatc ttttttaatg catttgcagt ttaccaaata ggcttttcac atcttcaatt 143040
tgaaaagtgt tatcttgtcc gggcacgatg gctcacacct gtaatcccag cactttgcga 143100
ggccaaggcc gaaggatcac ttgaggtcag gagtttgaga ccagcctggc caacatgggt 143160
aaaccctgcc tctattaaaa aataaaaaaaaa ttattcaggc atggtggcac acacctgtaa 143220
tcccagctac tggggaggcc aagagaatca cttgaacctg ggagggtggag gttgcagtga 143280
gctgagatca taccactgca ctccagcctg ggtgatggag tgacactctg tctcaaaaaa 143340

aaaaaaagga aagaaaattg ctatcttagc acctaaggaa gctggagtgg cctagggcca 143400
aagtggaccc acagcaaccc agattcttcc ctctcttgcc ctctttgtga agctcccca 143460
gaagcagcat ctaccactcc aagaggggaag ctggggcaga actcaacca cacaggctga 143520
agtgtgccct tgcccctagg ctacacttgt gggaaaggct gccttaccac caccaccca 143580
accctgact cagcaggga gccagggcct gggggctgtg cccaatggcc tagcttctaa 143640
ccagtcttca gcctcctccc ctacgacccc agagtgccag cctgctgtcc ccagacactc 143700
ttcccctggg ctggttatca gctgcctctg cccagtttga gaagttgggt ttttttggtt 143760
ttttggattt ttttttaatg tgcctttaag ccatcttcat ccagtttctg atatctttgg 143820
ttggggttgt tacccccgcac ttgtttggtt tattttgttta taatttctaa ctccttttca 143880
atcctgccta ttgggtttga cagtaatttt gtagatttaa aatgtttctt ccagatccaa 143940
gtaaattgtc aaaagcaaac tgattgaaat gaaccttcta atcaccagtt cacatttgca 144000
caatgtgaaa gctttcctgt ccgttctatt ttctgtgtta tcctgctttt taacaatgca 144060
aaattagctc ccagccagtg tggcttgact tacgtctctt gcggtgttct ggctgtttgc 144120
cacatggtag tgtgggggta caggcctcat cctagcctcg gtagcctgga tccaggcctg 144180
tgagagcaca gcaaccaatg aaggcaggac aaaagaaacg ccagggggccc gaggcacgaa 144240
gcggtggaaa gagatggtgg ttcacgtgg ctgctgccct gtggcccctg catctgctcc 144300
tgatgcctct gggttagatc tgcaggcgca ctggtctggc agaagcggca cgacagtggg 144360
aggctgagcc agtgcttacg catgtcccac cctcccaggg ttcttttagtg tcttaacact 144420
taatcttcac agtaccacca gtccacagat gagggccaca ggcccaggga gtggacgcag 144480
ttggtccagg tctcccagcc agtgagtgat ggctgtggga tttagacca ggcagtccgg 144540
ggagcactgg aagccccagg tattagctcc gcaggcatgt tgaccagcat tgctgccata 144600
caaatccaa agcatcaggg aaatgctaaa ccagggggga aactgcatct aattctgaaa 144660
aatggaccaa aaaatgcaa agagagttac aaaacttcaa acttagcaat aaaggcaaaa 144720
cagcactggg actatttctt tcccaccaat tgcttgctcc tgacgaggtc cgagccagca 144780
ggccaactg gccttcctg cagtgagtta ccaaccagaa gcacacaggc ccctgcccac 144840
ccctgggctc tgagctgctc tgcttccta acctggcaaa tgccagacct ggggcccctac 144900
cctagtcca cctccccctc accaccctag tctgcctcct tgagaccagg atccatcagt 144960
tcctctcatc tctgtatctc ctcttctacc caaggcctgg ccagaagtag gaggtccttc 145020
aggttggaac aaccaaagat aaatggaacc acgttaatac attagagcca caggagatt 145080
ccagaacagc gccaggggct ctttctcagg gtctcaggaa gggttgacac atggcaccgt 145140
gcacaagtca tcagccaggt gctggggata aaatgacagg tatgactgtc agacatggtc 145200
ccttcccctc gcagcctgga agttgcagag gagagggatt agcaaaggca gagagacccc 145260

agagacagct gctgggagca aggcggccgg ggccctgggg aggggacaga agaggcccaa 145320
 cgtgaggag gggcacaacc gagggacgat tcctgccacc tggccagggt ggtgaggaag 145380
 gcacagccct atggagctgg atgccgcagt ctgagtggga gccactgtgg cagctgagag 145440
 cagcttgttc caaattactg tgtgagactg tcatctctag gtctgtcaga ggcccatgac 145500
 caaagaggca acttaaaatg catgcattta gtagagaaat gaccaggccc tcggcctgcc 145560
 cgccacagaa tgcctggccc tccaccctct gtgggctcaa agctcgtgaa tgcggctctg 145620
 tagcccaccg tcggtgtcat acctctcatg gagacggctc cctctttgtc ttgcagggga 145680
 tgagctgacc aagatggagg acgaggatga gcagggctgg tgcaaggagac gcttggacaa 145740
 cgggcaagtt ggcctatacc cggcaaatta tgtggaggcg atccagtgat gagtcgggga 145800
 caggccagcg gggggacgga ggcgggcgggc ccaggagcct cagccagcca cgtgggcatc 145860
 cactcctttt cctgcaagag atgatgggtc cattgctctt ggcttcatgg tgttcctgga 145920
 aggcagatga gctggtcatt tcgcctggga ctgggcacct ttccgagtgc agctggaggg 145980
 atctgagcgc aggaagacgc agaacaacag aaatagccgc ccctccccgc ccactgtgcc 146040
 tgttggccta tcatagatct ctatgttctt gactttgtct ctctttccg agtcaatggt 146100
 gggttacact gatcttggtc cactgattac tctctctgac gagtccatca cctgcaactt 146160
 aaatgaacaa gcttacatcc cattttgagt gaagattttg aggtttttta tttaaaggct 146220
 gtgtacagtt atactttttt atacacctgt tcatttctac ttaaattatg gcacagattg 146280
 atgcgcacca gtcttgagga aacgatctcc ctattccctt accctgttac tcagccacgc 146340
 cgtgtgtagg cttagcctca ggtggcagat gtttgaggaa aggaattatg ccaggaagggt 146400
 gggaccgggt tatggtcggg tttctattgg gaatgctctt tgtgcttttg ggcatctgaa 146460
 tggaagcttt acatagaacc ttaggtagaa ctccccaaa tcgccatatt taaaaattat 146520
 tttcactcta ttcttgctta aaactgtact cttttgcaaa ttaacaattt tatcactgat 146580
 tcagagttaa aaagaagact aacttttcaa gcaaatgcat ctgtaaagat gcttttagatt 146640
 agactgtcat gtctcagtgt ctatctgtat atattatttg atattcagag aatctaaagc 146700
 actcgtctac tgttttaatg agatttaaca gcttttaaca gtgagtttcg tttgtaaact 146760
 gcttgaagtc tgtggcattc aggcacacgt ctggctggcc ggctgggtct cctcccgggc 146820
 tcagtgggcc tggggcctct ctgacgtggt gcctgctgga gggaggctcg tcgtcaccag 146880
 ctgactgctg gtccggcttc tgaccggcct ttgtcctggc tccgtagcag aacactgtaa 146940
 aagtgccgc gtctttgcag tagttgcaga tttcagtcgt cgtgttactt gtgcacaaac 147000
 agaagctggg tcttaccgc agcacgagtg tctcgggctg cccggagtcg cccgggagca 147060
 ggtgctgcag ccagagttac gcgggggcca cgcgggccgg cgggggtggg gggaacgtgg 147120
 gggaacctgt gtttcacgtg actcagcagt gcccgccgcc gtcaccagct atgcattcac 147180

tccgtttcca gtgagcagat gtcttgcttg gaaagtggac ctgtgtctgt gtctgtcctg 147240
 agaacttacc agcagaaatc ctcatttctg tgctacggat ttaccaaaaa ttgtcaagtc 147300
 tttttcagtt taacagttcc ttacatgtg tagtatttga ggaaaaaaat caataaacag 147360
 ttgatctcgt gcatatggaa gtcccttcgc catcatctgt cttcatgccc acttcacttg 147420
 gcgggggttg cctccctggg gcttactagc tttggagctg ggcaagatcc agggcacagg 147480
 acccctgccc aaaaggccac ggcccactgc ccctgccaaa ctggagggttg gggatttgag 147540
 gcacctgagc cccttggggg tcccttctcc ccgagacctg ctgctgtttg acctgggctt 147600
 ccacatgggt cgggcatcgg ctggcgctc tctattgccc cagcccaaac agcaggacaa 147660
 gactgacatc acccataact ctctatgag cacattttcc tgagcaccta ctgtgtgctc 147720
 tccacttttt acttacatta tctgcgtctc caaacatctc cacagggtat gggcaagtct 147780
 ccccatttcc attctaccga ggaggaaaca gaggtccaa gagccagcca gggtttcttg 147840
 aggagtctca gccagcccaa agcactgagc gaaatccctg aggaaccacc ccctccctcg 147900
 ctggagaaac tgaggaggga ggaggcttg actggcttga agacaggaga gaagcaagac 147960
 gtgagggcgg ctgtggccac gctcaccgag ccctgggaag gctgcccaga aaagcacgcc 148020
 agctgggcat ggggcctcac agagcacctg ggaggagcaa ggcaggcagc tgcccctgca 148080
 gtgtggcttt gtcccagctc agcaagggtg tgggacagat ccttgtcacg cagcctcctt 148140
 gtctgagggg tggcgggtcc cctgctcca gcttctgagg agttggagcc ctacagcgca 148200
 gggcggacac ccggctcggc ccacagcccc aggctgcaga gaaacaggca gaacgggggtt 148260
 gggaggggaag gtgggggggtc agaagcaggc tcctggtgcc agtggcctca gtcactgccc 148320
 aacaccatg gcaagaccac ggtccccgct gagctggatg aagcaacctc cactctgagg 148380
 ctggatcaca cggccccccc cgtgagagat ctggcagggc ctctaactgc ccatcttccc 148440
 tctcctgccc tgccctcagg ggccagcttg gaaacttcag cttgtctcta ggcctcagtt 148500
 tccccaagta gaagtgatac ttttccagtt ggagagttag agcaagaaga aaagggtggc 148560
 tgggcaacat agtgagacct tgtctctaca aaaagttaa aaattggtcg agtgtggtgg 148620
 tgcatgcatc tagtcccaac tacttgggag actgaggcag gaggattgct tgagcccagg 148680
 aatttgaggc tgcagtgagc catgattgca ccactgcact ccagcctgga cgacagagct 148740
 ctcgagacag tctcaaaaaa agaaaaaggc cgggcacggt cgctcacacc tgtaatccca 148800
 gcactttggg aggccgaggc ggggtggatca cgaggtcagg agatcgagac aatcctggct 148860
 aacacgatga aaccccgtct ctactaaaaa taaaaaaat tagctgggcg tgggtgcctgg 148920
 cacctgtagt cccagctact caggaggctg aggcaggaga atggcatgaa cccgggagggt 148980
 ggagcttgca gtgagccgag atcgcgccac tgcactccag cctgggtgac agagcaagac 149040
 tccatctcaa aaaaaaaaaa aaaaagaaaa acgccacagg cactcaggag ctggtcattg 149100

WO 2006/119775

PCT/DK2006/000254

tcacgattat gagtgacacc ataggcaggg gccaggggtt ctctgctgtc catggctgcc 149160
 cttcgctcct cccacagcca ccaccctgga tcagtgtgtc caccaactgt ccacacctct 149220
 tcgggtcccc aaggcccctg agttgacact aggggggagt ttagccccag agccctctcc 149280
 tgggccaaag ctacctggag ctcaggacag gggatatttca tcattctctct ccattccttcc 149340
 cacaaccag aatgtcggg aggggtgagca agtgc 149375

<210> 3
 <211> 15981
 <212> DNA
 <213> Homo sapiens

<400> 3
 cagtacattt gtatgtagaa gtatagtata cagagataag aatttacaac atagtgtgtg 60
 cgtcagtaat ttctaacaga gccttaaaac agaaacacag tctttccata acctatgatt 120
 agcaagattt taatcagtag taaaaactgc aacaaaagct ggttacaaac aatccatgga 180
 aacaggaagc tagacaaccg gttagaccag aaattctcag aagggtgat gccttaacc 240
 taaagaggcc tagaagagcc acggcaagat gagggcgttt atagccctat cttatccata 300
 tggacaggca ccccttatgc gtccatttat aggctctcca caagggtcgc attccattcc 360
 cagagctatg aacatctgct ttgctgggat acgaatcttg gtgatgtgaa acttccctga 420
 ctgcaagtcc attcataggc tctctgcagg gggaagcaca tcacgtgcta ttggctcatt 480
 ctggcagtcc aacctggcat tgtctttaca caatcctgca tgcaattttg tattttacaat 540
 aatcaggagc atttcatctt ttattccata gcaatagttt caggggggtct ccctacaatg 600
 aagttgcgcc acatcggcca ggctgggtct gaactcctga cctcaagcca cccacctgcc 660
 ttggcctccc aaagtgtgta gactacaagc ctgagccacc gtccctggcc aaaaaacaaa 720
 aaacaaaaca cacacacaca cacacacaca cacacacaca cacactttaa agatatgccc 780
 acaccatttt ctccccctat ttatagctgt actttgtcta tggtgttggc gaatgtaacc 840
 attacactcc atcctctcct ttttactcct catcttctct tggttctcca tgtaaataata 900
 tctttaatgt tcacaccagg ctaatgtcaa tgtcttttta gtctcgttgg ttgtctgaag 960
 cttattttct agtaattcct caggaaaggc tttgggaaca atattctctg ggttcctgca 1020
 cattggtgag aatctgctat tttttttttc tttttaactc actctgtcac ccaggctgga 1080
 gtgcactgcc gcgatctggg cccactgcaa cctccatccc ccaagttcga gcgattctcc 1140
 tgcctcagcc tcctgagtag ctgaaattgc aagcgtgcac caccacgccc agctgttttt 1200
 ctatttttcag tagagacgca gtttagccat tttgaccagg ctggtcttga actcctgacc 1260
 tcaagtgatc tgcctgcttt ggccttccaa agtggttgaa ttacaggcgt gagccactgc 1320
 gtccggcctg acaatctgct atttttatac ttaaaagtta gtttgaccag gccgggttgg 1380
 gtggctcaca cctgtaatcc cagcattttg ggaggccgag gcgggctcac caccggatgt 1440

tgagagttcg	agaccagcct	gaccaacatg	gagaaaccct	gtctctacta	aaaatacaaa	1500
attagtcggg	cgtggtcgcg	cgtgactata	attacaggac	tgcccagtgg	tcgggaccct	1560
gctaggttcc	agccagtggg	agcggatggg	gtagtgggaa	ggcggggcct	ggcgggtgag	1620
ggcggggcct	ggcaggtggg	gcggaggcct	tccgggggtg	gagccggggc	cttcctgggg	1680
tggagccagc	ggccctgggc	acagggcccc	ccggggctct	ggcagggcgg	tgggaaggcg	1740
gggccttctg	aatggggggcg	gggacggctg	gagttgcggg	gactgctgga	gttgcggggc	1800
ttgcctgggg	tggggcgggg	cggcctaggg	gtggggcggg	gcggcctagg	ggtggggccg	1860
cctaggggtg	gggccggcgg	ccctcgggat	gggtcctgcc	agggcgcaga	gatggcgggtg	1920
gaaaggcggg	gcgaaggcgg	ggcgtgaggc	gtggaggtgt	ggcctgactg	gggcggggcg	1980
gggcggggcg	gaacagccta	gaagcagggc	ctggaactgc	aggatacact	cccctcctgc	2040
tacctaggca	ggcgtgaggg	tgtgacggcc	gcgcattcgc	cagacgagag	cgatggctga	2100
gaacgccgca	ccaggtcgcc	ttagactcac	caggtcgcct	tagacttaga	tgggacgggg	2160
aggcctcgcg	cgagcgcccc	attgcccacc	cacctggttg	ggactgtgga	ggagggcgga	2220
tggagtggag	ggttcaccga	ccgcgtcgcc	cttctgggtg	tcgcagctcc	ttcctcggca	2280
gtcccggggc	gtgagagcgc	gaccggccgg	gagttggggg	acccgggaca	cggccgcacg	2340
tggcctccca	gggttctagg	agaccctgag	ttagccgggc	gtgcgcgtgg	gtgtcggggg	2400
cgcgaggtc	tccagccttg	cagagacgga	ggaggctgcc	agctgtgggg	ggcggcggtg	2460
ggtcctggag	tcccaccgcc	tgccctaaat	ttccactccg	cccctgtcgc	ccagtgtgac	2520
cttgggcatg	tactgcagct	cccagtgcct	cagtttcctg	ctctgtggag	ggagagtgac	2580
agtgtccca	cctgctgcgt	gtgtgagggg	ttggcacccc	tctggccttt	gtgaaaggct	2640
taggaaacgg	gctgttttct	tactccgggt	ttacaggaaa	ggaaactgag	gctcaagacg	2700
ttttagtaac	ttctccaagg	ttacccagcc	agcaagtgat	ggggtccagg	cctgaccagc	2760
gcccccgccc	cagcacattc	tagaagggtt	tgccccagtt	taaatggggg	gcatgctctg	2820
aggcctgtca	cagctttttc	cggcttagct	tctctgtccc	tcctctggga	gtctccactc	2880
ccatgtctga	ctctctcccc	ataaccttct	ggggcagcta	agggtctgag	tccctctgga	2940
catgagcccc	tggctcaagc	acatcaggga	gagataaggg	gctcaggggag	caccagcttc	3000
atgggtgaat	gaatgaagga	atgaaagacc	cggctgtgtc	tttcaggtct	gatctcagag	3060
ctgaagctgg	ctgtgccctg	gggccacatc	gcagccaaag	cctggggctc	cctgcagggc	3120
cctccagttc	tctgcctgca	cggctggctg	gacaatgcca	gctccttcga	cagactcatc	3180
cctcttctcc	cgcaagggtg	gaggctgggg	ctgaagggaa	ggccaggggg	aaagtgggca	3240
ggaagggtgag	gggggaggtc	tgcaggggaa	tgaggagggc	ggggagacag	ctggcataga	3300
cacctcttcc	taccctccaa	cccgcttctc	tctgtttcag	acttttatta	cgttgccatg	3360

gatttcggag gtcatgggct ctcgtcccat tacagcccag gtgtcccata ttacctccag	3420
acttttgtga gtgagatccg aagagttgtg gcaggtaaga aacagagtgt gtattttcgg	3480
cgggtgtgggg ggtgctctag ggcaccccct cttatcactg gaggcgggggt aggatccagg	3540
aagcagcgct ggcctgagag tggggcctgg gctctggccc caggtctgcc agcaattcac	3600
gggagatctt gggaaagttg aagagttcgc ttcctccagt ggggtcccca agctagtgtc	3660
ctgggtggac accccagatg ctgggaaggg atttcctgag cctggggact agaagcaaga	3720
tcgtgggctg agcctgtgcc cagccggacc ttctggctcc ttgcatttcc agccttgaaa	3780
tggaatcgat tctccattct gggccacagc ttcggtgagt acagtggcca ggagctgacc	3840
gggcccggag tagggctggg agggaaggat aggagctgct gccccacccc tttcccctgc	3900
acttatgaga tctgcaaaga gagatgagag ggtcagtcag aggttgacca ctgtaccctc	3960
tgacctcagg caactgccgc aagggaacaca gggatgaggg aaatgtgggtc ttttgagtgg	4020
tcaggtctga tgaaggggac atcatggccc agatgctgag cccccagcca cccatatggt	4080
gccccccccc caactggtgg gaagaagctt tagggtatgc agctggggag ggagtgaggg	4140
aatgtcaccc acagctcttt tgtccccag gtggcgtcgt gggcggaatg gtgagtagat	4200
ggctttgtct ggccaactgg ggctcccttg ggtggagtgg ggaggggagc acaaaggagg	4260
aaggagaagg tacctgggac agaaagtgcc tcctcctgtc acagtgatag tccttatacc	4320
aggaagttgg tgtaaacggt ggagactcct ggtagagttc atgtttccac gccccatac	4380
ccggtcccac cgcaccctgg agtccagcca tcaccaaggt ctggagctca gtgctgggtg	4440
cagacttagg ggtggtgggg gaaggcacac catagataga agattctcct ctggcctctt	4500
ggtcagcacc ctccccccag cctaaaagat gtttcgggga actccagggt cccagggaa	4560
aggcgggagt gtcctttcct tcttcacct attcattctc cccttttggc atcctcacag	4620
tttttctgta ccttccccga gatggtggat aaacttatct tgctggacac gccgctcttt	4680
ctcctggaat cagatgtgag aagcgggctt tcgtgcatgc tgtcattagg gaggacctgg	4740
gccccgggca gcgctcccag ccctgtctcc tttggtggac actagggaag cacgaggagg	4800
cggcagaggg tgggaaaggg gagtcaggga cccgctcccc aacagatgac acccttctgc	4860
tgcttatacc cctagcagtg gtcaggacaa ggggtgtggg ggacccttg gaaactggat	4920
gtgtggctcg ggatctgcag aatgccaggc tgatcacgca gtggcaggca gtggaagagg	4980
gcagcggcca tgggtggaag atgttctagg gcacacagct gggccggcac actgcactca	5040
caccagtcg aatgcctttg cactgtgtcc acattgcacc accacacatg acagccctaa	5100
atgtgaactg aaggaacaag gactgacccc tctgaccacc actaccata accagtactg	5160
tctctgcgtc tgagtcctgg cctggtttac ctgctgggtg acccctggca agttactttg	5220
cctctctggg ccttagttcc cgcacgtgta aaatgaggtt gggctcaatg acctggccag	5280

ctctgacatc	ctgtcatttt	cttgccaagg	tctcagccta	tcagttgtga	cacaagtctc	5340
atcagtttca	gttttcttgt	ccatagaatg	accacagatc	ggccagatgc	ggtggctcac	5400
gcctgtaatc	ccagcacttt	gggaggccga	ggctggcgga	tcacctgagg	tcaggatttc	5460
gagaccagcc	tgaccaacat	ggagaaatcc	cgtctctact	gaaagtacaa	aattagctgg	5520
gcgtgggtggc	acatgcctgt	catcccagct	acttgggagg	ccgaggcagg	agaatcgctt	5580
gaacccagga	ggcggagggt	gcagtgagcc	gagattgcgc	cattgcactc	cagcctgggc	5640
aacaagaggg	aaactccgtc	tccaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	gaatgacggc	5700
agatcccata	aggataagga	tcatgtccat	ctctgtcacc	acatccccaa	tcctggcacc	5760
tagtaagggc	tcagtaaagt	ttccttgcat	gagctcagct	agctcagggg	ctattataag	5820
gatccaattt	acagatggca	gacagcagac	actcccttct	ttctgcattt	caccactga	5880
attgacaagt	gcagtgcccc	caggggcttc	ctcctgggcg	agtcaagagg	ggcatttggt	5940
cctcttgtcc	tgcatttctca	ggcatgggta	gatgggagat	gccccctct	gcagccccag	6000
gcagtgggcc	tgaggctgtc	ctcacagagt	gagccacaca	tcatggggga	gcttggaagc	6060
cccagaaatg	cgccctttct	gccagtcctt	ttctgccact	gagaatgatg	gcattcgttg	6120
aaccagatga	aactgccaac	cccagatggt	tcgtggcctc	cggaccctaa	tttcatgcct	6180
ttgcaaagca	cttggcagtt	gacagagatt	ccccttcctc	tccatcatcc	cacctgatct	6240
cttcaactgt	ccctggccgc	aggcaagacc	aggctggctg	tgactgaggc	gaagcaaaga	6300
ctcgttcaga	atcgtggggc	tgactagcca	gggctggggc	tgtgtgcccc	tttgccgaga	6360
ggggctcagc	tgagagcaga	gggcagggct	aggaggggca	gagacctcgt	ctgcagcctg	6420
gatccctgcc	caccagccaa	ctgtgcctcc	ttaggaaatg	gagaacttgc	tgacctacaa	6480
gcggagagcc	atagagcacg	tgctgcaggt	agaggcctcc	caggagccct	cgcacgtggt	6540
cagcctgaag	cagctgctgc	agaggtgggt	gcccgggtgtg	gacggggagc	tggtggggtc	6600
caggggagga	catcccgtga	gacgtggaga	agcttctttc	tccctctcac	ccaacaagcc	6660
tgtttctccc	tggggtcacg	ctgagcacca	ggcttcccag	ggtcttcagg	gacgttctct	6720
ctttccctct	gcccctgctc	cagccatgcc	ccttacttcc	aataccctcc	ccacacctct	6780
cacgcacccg	cctctctccc	aagcacatgc	tgcttccacg	ccttctctcc	cctctgggac	6840
tccagcctgg	gctcccttac	ctggaattct	cccctgtgct	tgctcactgg	cttccccact	6900
ccttggggcc	tcagctggga	accctgccct	gacaccccca	ggcaagctgt	ccctgccctg	6960
caggtaacca	cactgtccaa	taagagccta	ttgattcctc	tgtttcccga	ggcatgtgaa	7020
ctccgcaacc	cccactatgg	ctgatttggt	catctttgtg	cctcctgcat	ctggcagtg	7080
gcctggtagg	tcagagctcc	acacatat	gttgtctgaa	ggaggagaag	ccccccactg	7140
ccccaccaca	cgcacgcgca	cacacacca	cacacacaca	ccctgccgct	gtaaakatgc	7200

aatcgaaaca	cgggacacag	gccaggcgca	gtggctcaca	cctgtaatcc	cagcacattg	7260
ggaagctgag	gtgggaagat	agcttgagcc	caggagttag	agaccagcct	gagcgagata	7320
gcgagaccct	gtatttaaaa	aaaaaaaaaa	aaaaaaatta	gctgggtcgtg	gtggcacacg	7380
cctgtggtcc	cagctactca	gcagccttag	ttgggaggag	tgcttagccc	caggaggctc	7440
aggctgcaga	cagccatgat	catgccactg	cactccagtt	tgggtgacag	tgagaccctg	7500
tttcaaaagg	aaaacaaggg	ccgggcacgg	tggctcacgc	ctgtaatccc	agctactcag	7560
gaggctgagg	caggagaatc	gcttgaatgg	gaggcagagg	ttgccgtgag	ccgagagagc	7620
accactgcac	tgagcctgg	gcgacagagc	gagactccat	ctcaagaaaa	aaaaaaaaaa	7680
agaaattaca	taggtgacat	aatacataat	gattgtcaat	ttaaccgttt	gtaagtgtgc	7740
acttcagtgg	cattaaatac	agtcgatgga	ggtgaggaga	ggagggagac	atcccaggca	7800
ctggctgctg	cacagggaaa	agctcgaagt	cagctcggtg	catctgttga	gagggggatt	7860
cctgatagag	gctgtgaaga	ttaaagaggg	gaggggcagg	atgacaaagg	ttctcaaaaa	7920
acagcgtccc	atctggaggc	ctgtgtcact	ggagacttgt	attaatagaa	acttggccat	7980
ctctgccagt	gaggggctgt	aggaccttgg	gcaagggtgct	taccctctct	gagccttggt	8040
tgtctcacgt	gcaaaaagaa	agtgccttca	ttcgctcagc	tggtttatga	gcagcttggc	8100
taggggagcc	aagccctggg	cctgcctttg	aagggttggt	gtcgagacta	ggtgagggtca	8160
tgtgtgttga	gtggcctcgc	acagtgctgt	cgcttcaggt	gctcatggcc	tgtgctattt	8220
tgtgattggt	aaaatatggt	agaacattgg	acacatgtga	gacactgtca	ttagagggtg	8280
ggatgctcta	actgcaggga	gaccgctgag	agctgctgcc	caggcccagg	ccaagagatg	8340
aggcttggac	cagggcgctca	gccaatgggtc	tgggaatggt	aacgtgtttg	tctttgatcc	8400
gtgtgggatt	tgttcttgag	catggtgtca	gatagtgatt	cattttcttt	tttttttttc	8460
atgttttaaa	agttctattt	ttctttatta	ttattattat	tattattata	ttattattat	8520
tattgataca	gagtctcgct	cttgttgcaa	aggctggagt	gcaatggcac	catctcagct	8580
cactgcaacg	ttcgccctccc	aggttcgagt	gattctcctg	tctcagcccc	caagtagctg	8640
ggattacagg	cacccaccac	cacacctggc	taatttttgt	gttttttagtg	gagatgggggt	8700
ttcaccatgt	tgtccaggct	ggtctcgaac	tcctgacctc	aggatgatctg	cctgcctcga	8760
cctcccaaag	tgctgggatt	acaggcatga	gccacagtgt	tcgggtctatt	tgaattgtag	8820
agacagggtc	tccctctgtc	accagactg	gagtgccatg	gctcaatcat	ggctcacgca	8880
gcctcaacct	cctgggctga	agtgatcttc	ccacctcagc	ctcccagagta	gctaggacta	8940
caggtgtgcg	ctatcacacc	cagctaagtt	ttttactttt	gtagagacag	ggtcgtgcca	9000
tgttgcccag	gctgggtcttg	aatgcctggg	ctcaagctat	cctcccgccct	gtttcccaaa	9060
gtactgggat	tatagggtgtg	agccactgca	gccagccatt	gtaattgttt	ttctagatga	9120

acagctagtt attccaatac cagctgttgg gtggtcctcc ctaccacta aggctcactc 9180
 actggtggtt tgtcccttgt gtgtgctgac cttcgtgggtg ggagttttatt gccagatcct 9240
 agtcaatctg actttttggag atgccccttt ccctgtcctt agggaccatg tcttatgtga 9300
 cctgatctgt agtgggatct gtgatcacc ataatgctgg aaggaggagg cttcccaggg 9360
 atggaaagac ttcgttagga cataaactaa attgcagaag atacacattc caagatgggg 9420
 ctggcccaaa attctttttt tttattttta ttttttaaata ttctggccag gtgcagtggc 9480
 tcacgcctgt aatcccagca ctttaggagg ccgaggcagg cagatcactt gaggccagga 9540
 gttaaagacc agcctggcca acatggtgaa accctgtctc taccaaaaaa taaaaaatt 9600
 agccgggagt ggtgacacac acctgtagtc ccagccactt gggaggctga ggcaggagaa 9660
 tcgcttgaac ccgggaggcg gaggttgcgg taagctgaga tcacgccatt gcactccagc 9720
 ctgggcaaca agagcaaaaa aaaaaaaaaa aaaaggttct ataaattctt ctcatatttt 9780
 tatttaacat ccacaatgtg atttaagatg taatcaacat aaagcttgat tgcattattt 9840
 tgcatgctgt tttccaaatc cagcttcgtg tcacacctac agcacatctc actcaagctg 9900
 gccacatccc cgccatccag acgtaaaaca gtcacaagac agggcaggca gggccgcgga 9960
 ggaggctggc tggggccatc acggagtgcc catcctgcac tgtggtccca gcaagtttct 10020
 ttctcctggc aagaagcctg tcccaggctg gcaggggaca gcgtgagggtg cagcctatgg 10080
 actgggaaag ggggtgtgaa gggccacacc taagtcctaa aatccaggcc caaaagtggc 10140
 ccaactcact tctctgactt taatcacaca gccatacctg gtggcaaagg agtattggaa 10200
 atggagtcag gctgggtagc cacgagccca ggaagagggg agaacagact tggagagggc 10260
 aggagtctct ggccaccagg ggctaaagag ctttcgatga ggcagtgatg tggggtcctg 10320
 ggctcagacc cagggtgggt ggctaagggt ccctcgccag ggcttagcca cccaacaga 10380
 gatgggtttc gtgcccacga gagtgcctgt gccttgtgac gagaattcac catgtttttg 10440
 tctctgcagg cagagaacag cattgacttc gtcagcaggg agctgtgtgc gcattccatc 10500
 aggaagctgc aggcccatgt cctgttgatc aagtgagtct ggacccatcc ctttcagtca 10560
 cccccaagg agacatgggc gccaggaatc tctgggaggg ggccctggca tgaggctcca 10620
 agttctctgc gtgtcgacca catcgctaag actcaagatc ttttttggga agccccctg 10680
 gcagcagggt catggaagga ggaaggctcag aggaggggag ggctcaggca gcaggggatg 10740
 ggccggggct gtcccatgcc tttccacagg tgtcagcggg gggcatgccc aggtaaggct 10800
 ccataaccag tgagcccagt ctcgactcac tgcaacctct gcctcctgga ttcaaacgat 10860
 tctcctgcct cagcctcccg agtagctggg actacaggcg cccgccacca cgcttggtt 10920
 atttttgtat ttttagtaga gacgagggtt cgccatgttg gccaggctgc tctccatctc 10980
 ctgacctcat gatccgcctg cctcggcctc ccagtgttgg gattacaggc gagagccacc 11040

gtgcgtggcc caccatagac gattttttaag ccataaaaag aaacgaagca ctgacatggg 11100
 ctccagcatg gatgagcctt gaaaacatcg cactaagtgg atgaagtcag acacaaccgt 11160
 ctacgtgttg tatggctgca ttcacgttca agtccaaata ggacacacgg cagagacaaa 11220
 gccaggatcat gggtgctcag gaccggaagc ctgttggggg tggggggggtg ggggtgtgtgt 11280
 gtgtgtgtgt aagctcagag gtttgcgatt tctttggggg tgatgaaaat gtaattgtga 11340
 cgatgggtctc acaactcaat gtattaaaaa ccacgggatt gtagactttt tttttttttt 11400
 ttttgagatg gagtctggct ctgtcgccca ggctggagtg tagtggtgca atctcagctc 11460
 actgcaagggt ccgcctcccg gggtcacgcc attctcctgc ctcagcctcc ggagtagctg 11520
 ggactacagg caccaccac caagcccggc taatttttgt attttttagta cagacagggt 11580
 ttcaccatgt tagccaggat ggtctcatct cctgacctcg tgatccacc gcctcagcct 11640
 cccaaagtgc tgggattaca ggcgtgagcc accatacccg gccaggattg tagactttca 11700
 atgggaggat tgtgaagaca taaattctct ctcagagcaa gtgtgggctc cagcgcctgt 11760
 tcagatccca gctctgctgt tcagggtga ctgtgaccaa gtcctgaact ccctgcgccc 11820
 cagcaaggct gtgccttcag cacttgcttg cgcctggctt gtgccagctg agcgtgagct 11880
 gagattaact gagcctcaaa tccaaccag ggtcaaggga cccagctgtg ggagttgggg 11940
 tagcccaggc tttggtttct cctcagaggc ctgaaccag ggcaggatgg gggcaccgct 12000
 gggaggcgggt tctaagatga accccaacaa cccccaactc ctcactttcc agagcagtcc 12060
 acggatatatt tgatccaaga gagaattact ctgacaagga gtccctgtcg ttcattgatag 12120
 acacaatgaa atccaccctc aaagaggtaa ggcggggctc aggcagctgg tgtccagcca 12180
 ctgtcgccca ctctgggtccc acctcaccca tctcttctca tgcactggga aagaccctg 12240
 gtctgcccc aggccaaca agtgaccacc aggatctatc aggcctcatg ctccagtgat 12300
 cagtccctag aggcctgggg acacatgtaa tcagaattta agaaacagag acctttgggtg 12360
 gggagggttct ttaccatcct caagtaccca cccctccttc ttggggccag gcacaattag 12420
 gtgacccaac tgccagagac agtcccacat tttttgagag acagggtctc attctgttgc 12480
 tcaggctcag tcatagctca ctgcagcctc aaactcctgg gctcaaatga tcctcccacc 12540
 tcagccacct gagtaggtgg cactacaggc gcatgccatc atgcccagct aatagtcaca 12600
 cttttttttt ttttttgaga tggagtctca ctctgtcgct caggctggag tgcagtggga 12660
 agatctcagc ttactgtaac ctccacctcc cgggttcaag cgactctcgt gcctcagcct 12720
 cctgagtagc tgggattaca ggcacctgcc accatgcca gctaattttt atatttttaa 12780
 gagatggggg ttcaccatgt tggccaggct ggtctcaaac tcctgacctc aggcgatccg 12840
 cccaccttg cctcccaaag tgctgggttt gcagatggga gccaccatgc ccagcctagt 12900
 ctcactttat aggctcccaa accacagcag aacacctggc tgcctcatcc ctctgcacc 12960

atcacgccag tgggcatgct ggctgtaggt ggggatggcc tttgcccggg ttcctgggag 13020
 cacctggtgt tttatggtca cctaccact tacacctgac cctggggcct ctgcagtttt 13080
 cttggcgttg tcaccacaaa aggaagaagc ccacctgtgg ctggctggct ggccgccagg 13140
 tcagagccaa ggcaggcagg gccttcctgt ctggcttctt gaaggcagac acaagtctgc 13200
 aggctgctca cgtcattctca ttatcttttc tcacgaggac tacttcggat acaatcacag 13260
 caaccctggc ctcggccctg ccctgtccct gccatgcaac ttcacaactc agctggccta 13320
 gaccctggg aggcctccaa gtccctaagg ttagacatct cctgggggtgc tatggactgt 13380
 cggggctcca aggagccgag tgtgggggaa actcactgtg ggaggcgctc ctgacctgca 13440
 gggagctgga atgctgtggg agggccctga ccccggggcc catggagctc cctaggctcc 13500
 tctggccaca tctcactacc caccctcc cctctccagc ggttccagtt tgtggaagtc 13560
 ccaggcaatc actgtgtcca catgagcgaa cccagcacg tggccagtat catcagctcc 13620
 ttcttacagt gcacacacac gctcccagcc cagctgtagc tctgggcctg gaactatgaa 13680
 gacctagtgc tcccagactc aacactggga ctctgagttc ctgagcccca caacaaggcc 13740
 agggatggtg gggacaggcc tcactagtct tgaggcccag cctaggatgg tagtcagggg 13800
 aaggagcgag attccaactt caacatctgt gacctcaagg gggagacaga gtctgggttc 13860
 cagggtgct gtctcctggc taataatctc cagccagctg gaggaaggaa gggcgggctg 13920
 ggcccaccta gcctttccct gctgccaac tggatggaaa ataaaagggtt cttgtattct 13980
 cactgctttt gaggcttttt cgttgccagc aagggtttt gcattgaggg aaaaggaagc 14040
 acagaacgga ttcattccagg gtcctggaga ggagggtgtg gaggcaagaa gcaggaagga 14100
 caagtcctcc tctcggcacg cctgggtccc caggactgct gaagcacaag acccgagggg 14160
 tggcctggtc cttccctcct ggcctgtgtg gactctgatg gggctgttg acgcggtggc 14220
 cttcaacctg gggcccgttg gccagagctc ggcctctcag agaccgctgc aggccctgcc 14280
 tcgcctcctc ctggccctgc acctccagcc attcactgcg gctctcacag ctcagtacgg 14340
 tcggaatttg cgctggggccc gcagcagctc gatcctctgc ttgtcctcct cgaacttctt 14400
 gcgcagctgc gctagatctg ggggtgagag gaggcgcggg gcagggtcag acagcgcgcc 14460
 ccggggaagc tggcttgctc atgccccacc cctcctccca gttactgtgg ggggtggggct 14520
 atgtctcccc caccagcca tccccctccc agctctgctg tgggggcagg gtctggatct 14580
 tgtcaggaaa cagaaccgac ggcaggccga gtccaagggc aaggccaaag tgggagatgc 14640
 tggcaccccc agcctggaga tgaagcccta gcgacaaagc aaaggtaccc gccccaccc 14700
 ccatccctgc acagggggaa ggaaggagtc ttgagcgctc ggcctggaga accctgagaa 14760
 tcagtaaagg caccagccc cttgtgctgg tgaatgcaga tggcaaccct ggtcagtggg 14820
 aggctgatgg cactgagggg gaagctgagg gtctagagcg agcacgagcg ggagccagga 14880

```

gggccagcca ggcagggaaa gcaccagtga ctggggcgggc ggctcagggc ctgccgcaga 14940
gcccgtccct gtgccctagc cagggccagg agaaccctcc cctggggcca cagggaaagct 15000
gttctgtaga cagggacacc aggcagtggg gcaggaggct cagccgcagc accaggagtg 15060
ggccacaggc gatcggggccc acaggcatgg gagggtcctg tgcggggcca cctgacagac 15120
acacacacac acacacccac caccaaccca acacacatac tctgcagcac cttctctgca 15180
cccacagtgg gttcagaagc cccctctgca cttgaactcc tcgttagctg gaggctgacg 15240
agcgccagca ctaagtgcac agggacccag aggtggggcac tgatgaccta ggggtccccac 15300
ccactgagca aatgcatctg cccctagggg gtgatgggga accgccacac agggtgctgc 15360
caggctgggc ctctttccac agagtctggt gcgagccccg tgcttaggga cgccctccct 15420
cctcgaccac caagtccagc gcccgccctc aggagcactg gccacaggcc tcatccgtgg 15480
ctggggccaa ggccgagggg tctgaggatg gactgggggt ggcttacgct ccatcttgct 15540
ctctcgatgc tgccaggcgt agaagttgag cagctctttt cggctgcgct tccgtctctc 15600
cctctccagc acccgcaagc tggctgcctc agtccggggg agcacaggcc gccggccccg 15660
gcgggtcacc ttcacccagc cctcctcgtc agggaccccc tcctcctcct tggccttagc 15720
ttcttcctgc aaggaagggt atcccatcct ccggtgggac ctctgcagc ctccagcagc 15780
gcgccctcca cgccctcctc ccaaagtcag aactgccctg agcccccgca ggggagggct 15840
gctcccctcc tcccagcaca gtgttggggc taaggatggc ggggagggca cggacctgcc 15900
ggcactgcct ctttctccac tgtaggacat gtggagacta catgtgtcac ccagccaggc 15960
agcccaggat ccagcaggga a 15981

```

```

<210> 4
<211> 19542
<212> DNA
<213> Homo sapiens

```

```

<400> 4
aggcatggtg gtgcacccct gtagtcccag ctactcatga ggctgaggcg ggaggatcat 60
ttgagcccag gagttggaag ctgcagtgaa ctgtgattga gccactgcat tctagcctgg 120
gtaacaaagc aagaccctgt ttcaaaaaag aaaaaataaa taaataaaaa ggaggaaaga 180
gagaggagaa tttaaagggg aaggagggt cctctgggta gagttctcag gtggggctgt 240
gaaataaaat cagaaccctg gctgggcgcg gtggctcacg tctgtaatcc cagcactttg 300
ggaggctgag gtgggtggat cacgaggtcg ggagatcgag accattctgg ccaacatagt 360
gaaaccctgt ctctactaaa aatacaaaaa aattagctgg gtgtggtggt gtgtgtctgt 420
agtcccagct actcaggagg ctgaggcagg agaatcgctt gaaccggga agcggagggt 480
gcagcgagct gagatcaagc cactgccctc cagcctgggc gacacagcaa tactctctaa 540

```

WO 2006/119775

PCT/DK2006/000254

aaaaaataat	aataattcag	aacatcccca	ggtaaccaca	ggcagcattt	ctcaacagtc	600
tctcctgcag	aaccagaaac	ccagactgtc	tgagcctgat	tgtcctgaga	ggttctggaa	660
gggcaggggc	gtgtggtgac	actgggtgct	cctgggttcc	agtacctctc	tagccagttg	720
gaagaatata	ggcaagtcaa	gtccttcctc	ctccctgagc	ctctaattcc	tcaccggtaa	780
agtagggatg	ataaaaataa	attattaaat	tgttgtgagg	tttcaatgag	ggcaaagctt	840
gctgcctgaa	acagagcaag	ctcttgtggc	gtcttagcta	tttgtattag	ttacttgtat	900
ttttctctgc	ggtttccctt	ttaaactcca	cctccttttt	attttaaaaa	gactcctttt	960
gaaaatgcca	gcctgcctca	gtagtaccag	gagacaccag	ggaaaggctt	ccttctttcc	1020
aagctgaggc	tgagaaggaa	tgggaagaac	gaagagaagc	acctttgtgg	aggcctcctc	1080
cagctggccc	tgggttttagg	aacatcacac	actcaacctc	caaaggccct	caggggctgg	1140
cccaagggtca	cacagcttga	cagtgcggag	cacttcgaag	ccagggcggtg	tggtgtagaa	1200
caggctgggt	gcctgaattc	aatcagccct	ctctacttct	ctactcagag	cctcccaggc	1260
tgtgggggtgt	tcacctgtgc	accttcaagc	cgatcacttc	cttgggggtcc	tcgtcaattc	1320
aattccatgg	aaatattaca	aggggtagac	agagatgagc	cgtgtcagtg	tctgtactag	1380
cagagggcag	gaccactgag	cgtccggagg	ggagacgagt	gtctgggtgt	ccctgggaac	1440
ctggcttgtc	tcgacccctc	tgggggggctc	ctcttggctc	tgacctccat	tctcccagac	1500
tcacccccac	tcatgcccac	ccaccccagag	gggcccccca	gtgtgcactc	acacactcac	1560
actcaaactc	accaccccac	ttgctctcta	attaccagga	acgtggcgct	gtctcagggtt	1620
cttcctcccc	tgcattcctc	agcagctgcc	ctcactgggc	agggcaaatt	ataccgtggc	1680
ccgaccataa	aagccacttt	ccaaaaaaag	aagctttaag	gcccccgaga	gacgcagctg	1740
agcctctcac	ctggcagtg	ctgactcagg	aggtgatgtc	tgggccatga	gccacagagc	1800
ctggtgtccc	tttaagcagg	gagtacaaat	tctcatccct	ccctcctgct	ctccctgacg	1860
tcagctggaa	ggcgtgtgag	tgagtgaggg	gctgaggtac	cgcccgcgcg	agcacacaca	1920
gacacacaca	cacacacaca	cacacacaca	cacacacatc	ttccagaccc	atcccctgcc	1980
tgccagctcc	acgagccaga	gagagacaga	gcgcggacgc	cgcaggaggc	agtgcctgga	2040
ccccagctgc	ccaggagagg	tgaggggcta	gatggggcct	gaggacactg	tgccagaaac	2100
tgtggctttg	gccaggccga	aggagagagg	tacaacgtga	aggtgggggc	tggatcagcc	2160
cagttcagct	tacagatagg	gtgttggggg	tatagaggat	accacctgtc	tctccctccc	2220
tcattctctc	cccgtgaccc	ccaaatgagg	tgaggccgtg	tggctcccgt	gggagcatgt	2280
ctgtgtctgt	gtctgtgtct	gtgtctggga	cctgcgtctg	gctcaggagg	cagagagcag	2340
cagatacctg	ccaagccgca	gcctcctgtc	tgggaagcag	cgggagagag	gaccgccttg	2400
accagagaaa	ccccagtgcc	gcttccagtt	gctgaaagca	gagtgcagca	gggacccttg	2460

WO 2006/119775

PCT/DK2006/000254

ctgggtccttc	ccctggaatg	ttgggctggg	aagggaccgg	ctgggaaagg	cagacagttg	2520
gacctggagc	agctgtctta	tcctgactct	atTTTTTccc	cagtgtgatg	gggataattg	2580
tgttagttgg	cttagagagg	tggctccgtg	tggttattct	cttccccag	gtaccacccc	2640
accccgagcc	cttgcttctt	gcctgcctgc	tggctctgctc	tgacctccac	ccttgccagg	2700
gcacagcccc	cggagggcag	gggtgataag	tcagggcctg	ccaggcctct	atcttgggat	2760
acagggactt	aggggtggcct	tggctgggtgc	caagcccatg	aatggcctga	ccctggcaca	2820
aggagtccag	gcggagataa	ggggcgacgc	tggcacccac	aaggagacag	caggttcctg	2880
ccaagctgcc	agcctgcagg	ctcttcccag	ggagagggga	gggggacaca	gagcgagaac	2940
aataataaat	acagagtgtg	aactgctgca	tccactgggg	ctggggtagg	agaggaggaa	3000
gtatgcccag	tgggggggtg	gcaacctctc	tgagagcctg	gcatcatgct	gcctggctgt	3060
gatccggcag	gggcctctcc	ctgggaaagg	gggtctccag	ctttctgggg	gatggggaac	3120
tggtttccca	aacctgactc	catcccagag	tcctctggac	cctccaaatt	ctgaatgttt	3180
cctcctccca	gctttgccct	gccgcccccc	atttccttga	agtcttctcc	aaaaatgtct	3240
gttttctgtg	tctgcccata	ttgacctatc	ctcacccctt	tcaaaaaaaaa	aaaaaaagac	3300
aggtatagac	aggaactgcc	tgcccgtgtc	ctgggaaagg	ccagaggaca	gttttggcca	3360
cagatggctc	tcccaggctt	ggagagacat	ggctaggcaa	tgttcacgga	tgcgcagggg	3420
gattccctga	gggggacttg	gggacactct	cagatggctg	gtggggaggg	gacacctgcc	3480
agccagtctg	gagtagagcc	ccctaggggg	tagagggcgc	tcaggacctg	gctgagccag	3540
tcatgtccac	ggtcacaact	ggggcagcca	ggacctcagag	ccttgtagga	atggggcctg	3600
gggcagttca	gacctttgtc	catgctgagc	agcaaaatgg	gaggctaaaa	gtagtaggga	3660
gctcttctgc	cccacagcca	gcaaaatgac	actcctgggc	tggggggctg	gcatggcaca	3720
ccagaaagct	gtgggggtcac	actgattgga	gttcaagttc	agatgctgta	tgtatgtatg	3780
actttgacca	gtctattcct	ctctctaggg	ggtctcaaaa	tgggaaagtg	acttgcctag	3840
ggtcgcacag	ccaagcacag	gtgggactgg	ggtgcgtatt	tggctctccc	agtgccgtcc	3900
tggtttcccc	accaccaccc	aagccacaga	agcccaaagg	ccttaaccac	agcagttccc	3960
cagggcagcc	tgggcatgtg	acaaggagga	ctacagggtc	tcttaaataa	ttaagctcaa	4020
gtattggctg	gggcagccct	catgcacaga	cgcaggcaca	caggctgtgt	gggggaagga	4080
ccaggtttaa	tcaggtcacc	tcaggctgca	agagcagttc	aggcactgat	gtattttaag	4140
gagcatctgg	gaggctgggg	gagggggagg	agccagggaa	gtgtcagctc	ctgcccgggg	4200
gtcctgagtt	tgtttttctca	ctctcccagg	gcccaccatg	cacaagctcc	cgatgttcgg	4260
gaagcccttc	cagcgtcatc	cttgtgttcc	ccgtcctttc	tattttttaa	ccaccgtacc	4320
cattcctccc	atctgctccc	ctgtgatggg	ggcttggggc	tgggacattt	ttattggggg	4380

agttgtttttt aagcttagca gagttggtga gtgactcacc caaggtcatg gtctgagtca	4440
gcgcaaagcc agaaataaaa gccccagtcc tggctcccca ggaattccac acccgactcc	4500
tcctttctcca ggggaggggtg ggcgccaccc tcaaggaatg gctgggtcttc acctccccag	4560
acatctccag agggcagata tgctgcacct gacctgctgt gagagctcag cgaggctctt	4620
cacctctctg agcctcagtt tactcttttag aagatgggtg tactaagcca ggcgcagtgg	4680
ctcatgcctg taatcccagc actttgggag gctgaggcgg gcggatcacg aggtcgggag	4740
atcgagacca tcctggctaa cacggtgaaa ccccgctctt actaaaaata caaaaaaaaa	4800
aaaagaaaaa ttaaccgagc gtggtggcgg gcacctgtat tcccagctac tcgggaggct	4860
gagccaggag aatggcggga acctgggagg cggagcttgc agtgagccga gatcgcacca	4920
ctgcactcca gcctgggcga cacagcaaga ctccatctca attaaaaaaaa aaaaaaaaga	4980
agaagaagaa tgggtgtact aatggtagtt actgcatcgg gttcttatga gtacgaaatg	5040
aaaaaatggg gcaaagggct tggcacatgg tcagtatcaa tagatcacta gatggtggct	5100
gaggctcaca tcaggacttc caatgctgct tttttaaaat tattattggt tttgagacag	5160
agtcttgctg tgtcgcccag gctggagtgc agtggcgcgga tctcggctca ctgcaagctc	5220
cgtctcccgg gttcacacca tcctccggcc tcaggcctcc ccagtgactg ggactacagg	5280
tgcttgccac cacgcccagt taattttttg tattttttagc agagacgggg tttcaccgtg	5340
ttagccagga tgggtctcaat ctctgacct cgtgatccac ccgccttagc ctcccaaagt	5400
gctgggatta caggtgtgag ccaccactct tggcccccaa gctgcatttt aataatctgt	5460
ttctgttccc cccgctggac tggggtgctc taaggacagc tgttgccccca cgtttgttgt	5520
tctggtgcct gcactcagtt tgatttcatt tacatatgtg tttttgaaaa cctcggagga	5580
gaattggcct gaccttcatt agtcatctgg accaacgatg attcacgtag ctgaagagaa	5640
aggtgaggtc cagagaggag caggtgcctg cccaggtca cacagcactg ggtggctgac	5700
caggctgcct ttccctaata cgctcagact gagaggactc atggtcaaga gaagaccata	5760
cagtcaggcg cggtggctca cacttgtaat cccagcactt tgggaggcca aggggggttg	5820
tcacctgatg tcaggagttc aagatcagct ggccaacatg gtgaaactct gtctctacta	5880
aaaatacaaa aattagctgg gcatggtggt gagcgtctgt gatcccagct actcgagagg	5940
ctgaggcagg agaatcactt gaacccggga ggcagagggt tcagtgagcc gagatcgcg	6000
cactggactc cagcctggga gacagagtga gactcaaaaa aaaaaggatg ggtgcggtgg	6060
tttacgcctt taatcccagc actttgggag gctgaggcag gtggatcaca agatcaggag	6120
tttgagacca gcctggccaa catggtgaaa ccccatctct actaaaaata caaaaattaa	6180
tcgggtgtgg tgggtgcacgc ctgtaatccc agctactcag gaggctaagg aaggagaatt	6240
gcttgaaccc gggaggcaga tgttgcagtg agccgagatc atgccactgc actccagcct	6300

WO 2006/119775

PCT/DK2006/000254

gggtgacgga	gtaagattcc	atctcaaaaa	aaaaaaaaaa	aaaaaagagg	actatggcca	6360
ggcacagtga	ctcacactgt	aatcctagca	cttagggaga	ctgaggctgc	agtgagccat	6420
gatggtgcca	ctgcactccg	gcctgggtga	caaagtgata	ccctgtctca	aaaaattaaa	6480
aagtaagaga	gaaccatgac	aaccatttcc	tgccccattc	tgcatacctt	taggggagct	6540
gagaattgga	gcagagtccc	agcatgggct	gtggttcatc	tgccaagaag	gctgagtcac	6600
caccagagt	gccccatgac	actgcatcca	gcctcagggt	atctccaaac	agcctgacat	6660
ttgtgtttac	ttctaggccc	ctttgtctgc	agagtcagtc	atttggggtc	tggccttctg	6720
cagggaatg	aagatgagct	tggctcagag	cccaaactca	gggccggggg	acagggggat	6780
gcacactcgc	cggggggcct	ttgacacagc	tctgtggctc	ctcccgacaa	cctgggtagg	6840
atgttatgat	cagccccatt	tgacagatga	gggatctgag	gcacaaaaag	ggagagttgg	6900
ccgggcgcag	tggctcatgc	ctgaaatccc	agcacttttg	gaggccaagg	cgggtggatc	6960
acctgcagtc	aggagttcga	gaccagcctg	gccaacatgg	tgaaacccca	tctctgatta	7020
aaaaaaaaaa	aaaaaaaaag	gctgggcgca	gtggctcacg	cctgtaatcc	cagcactttg	7080
ggaagctgag	gtgggcggat	catgaggttg	ggagatcaag	accatcctgg	ccaacatggt	7140
gaaaccttgt	ctgtactaaa	atacaaaaaa	ttagctgggc	gtggtggcac	acgcctgtag	7200
tcccagctac	tcaggaggct	gaggcagaag	aatcgcttga	acccacacag	cagaggttgc	7260
agtgagccaa	gatcacgcca	ctgcactcca	gcctgttcac	agagcaagac	tccatctaaa	7320
aaaaaaaaaa	aaaaaaaaaa	aaacacacaa	aaattagctg	ggtgtggtgg	tgggcgtctg	7380
taatcccagc	cagctacttg	ggagtttgag	gcaggagaat	cacttgaacc	tgggaggcag	7440
aggttgcagt	gagccgagat	cacgccactg	cactccagcc	tgagggacag	agtgagactc	7500
tgtctcaaaa	aaaaaaaaaa	aaagggggag	ggtcaccgga	ctcagtggct	cacgcctgta	7560
atcccagcac	tttgggaggc	cgaggcgggc	agatcacgag	gtcaggagat	cgagaccaca	7620
gtgaaacccc	gtctctacta	aaaatacaaa	aaattagccg	tgcgtggtgg	caggcgcctg	7680
taatcccagc	tactcgggag	gctgaggcag	gagaaaggcg	tgaacccggg	aggcggagct	7740
tgcagtgagc	caagatcgcg	ccactgcact	ccagcctggg	cgacagagcg	atcctccgtc	7800
tcaaaaaaaaa	aaaaaaaaag	ggagggtcac	aaaactgggt	tgcagtggaa	caaggattta	7860
aactcaggca	gtctgggtct	caggctgccc	tcttacctga	ctgctcgttc	tgaagggcac	7920
tattcctggc	aggagatggc	cccagaagga	cacacgaaat	ccccctggag	ctggccttta	7980
gctctgtttg	agtggaaaca	aatgttctag	aaacatatca	gtttctctgc	tcttggggga	8040
gtttattttc	ccatttagtt	tctcagggaa	agttcttttg	gaaaaactcc	agcctcatct	8100
tggctgaacc	aggaacagaa	tttatattca	caggaagagc	actggtttac	tgttagacag	8160
atctgggact	ctgtcacctg	ctgtgtgact	ttgggcaagt	cactcaacct	ctctgagcct	8220

cagctatctc atctgtaaaa tggaaaaatt gtaatttcaa ctctataggg taactgggag	8280
gttgaaacgt agaaaataat ttctgtaaag acttttgcac agtacatggc attcagccaa	8340
tgaaagcctc attttccactt gccctctttg gcagcaatgc tctggagagg gggctctgaac	8400
tggaggggtg ctccgtatct ccatcttccc ttctctctgg tctctcgggc tgctgggcat	8460
tgtacagctc ctccaagccc acgccctggc agggcccaga gagggaaagg ccctgtacca	8520
gctcacacag taggtcaata gcaatacaga gcctggacac aggactcctg ccttcccact	8580
ggaaggctcc agatcatact gattcaaatt ctggctcaga ctttgattgg tgtgtgtcct	8640
tgagtgagtg atttaacctc tttgagcctc gattttcttg tctacaaagt ggggtaataa	8700
tagcaccagc cttgaagagc agttgcagga attaaataaa acaagagact gtggagcttt	8760
gaggggtggg cgatccttag ggaagaagag aggatgaaag tgaatggtgg gcttggccca	8820
gctcggctgg gtctctcctt cccccactcc ctcttccctc aacacccaat tccacagcct	8880
tgtctaggag gggcaggctc tcgcccagcg cccagcgcct actcagtgtc tcccacatcc	8940
cacctcaatt ctcatctcat cctcaaaaga gcccttggag gagagagtct tttttttctc	9000
tctccctctc ttttttttga gactgggtct tgctctgtcg cccaggctgt agtgcagtag	9060
cttgattacc actcactgca gccttgacct cctgggctca agcagtcctc ccaccttagc	9120
ctcctgagta gctgggacca caggcatgca ccactacacc cagctaattt tttttttaag	9180
acagtgtcta actttgtccc ccaggctgca ctgcagaggc gtgatctccg ctctctgcag	9240
cctcgacctc ctgggttcaa gcaatcctgc tgcctcagat gcccaagtag ctgggactac	9300
aggtgcgcca ccgccatgcc tagctaattt ttgtaatttt tatagacaca agctttcacc	9360
atgttgcccc agctgggtctt aagctcctgg gcttaaaaga tcctcctgcc ttggcctccc	9420
aaagtgctgg gattattggc gtaagccacc ttgcctgttg gaggaggaat tcttgtgatc	9480
cccatctccc agataaggaa actgaggctg caatgggcga caggggagaa gcctgttctc	9540
tctgccttgt ggaagctgcc tgtctcctgc ccctgcctga gctcttgccc tgaggccaac	9600
tcgctgcctc ctcttggtcc tgtctttgct gtggctcttc tgccctccctt ctgctctctg	9660
gtctctctgc cacctgagag cccatcatgg ggaggctgcc ctcgaccctg gcccagaatc	9720
atcagtgggt tcttccccca tccctgcagc ccacttagca tgaggccctg gagccgcaca	9780
gtcacttggg tgggaggtct ggctggccag agcaggcccc agctcagccg cgtctccctg	9840
gtaaccacca tgagctgcat ctccccctct ccatcctcct cttccctccc tgccccctca	9900
tggcctccca gcaactggctg tcaactgcca gcacaatacc aggcctggga aaactgcagt	9960
tgggaggcca ggaccgccct gctccctaga gagagggagg cagggcctct gtccccatcg	10020
cctgccaacc ctcacccagg tggggcaggt ggcactcaag gagtgcactg cagggcgtct	10080
ggtgcttccc agctgtgtca tcttgggtaa ctttgctgcc ccgtcttctt ttttttggtg	10140

WO 2006/119775

PCT/DK2006/000254

acagcttttac	tgagatatcg	ttcatataacc	ataaaaattca	cctgcttaag	gcgttcagtt	10200
cagttgtttt	tagtacattc	acagagttgt	gcagccatca	ccacagccaa	ttatagaaca	10260
tttccattcc	ccaaaaagaa	acgttctacc	cattggcagc	cctcccctct	ccctccttcc	10320
cagccctggg	gaccaccagt	ggactccctg	tctatgagga	tttgcctcag	agatggaatc	10380
atgccctctg	tgttttttggg	actggcttct	tctctgtgac	cttgcaaaca	ttctcgcata	10440
atgtgtgcag	ggttcatcta	gggtgtagcg	tgaatcagta	tttcttcttt	tctttttctt	10500
atggccaagt	gctagtctac	catgtgtcta	gaccataatt	tattatccat	tcatctgctg	10560
atgggctcgg	gctgttttca	cttgttggct	gttgtgaatg	agcactgggtg	tgaccattca	10620
tgaacattta	tgtccaagtg	tttgtgtggg	tgcttttttt	tttttttttt	tttttgacag	10680
agtctgggtc	agtcaccac	cctggagtgc	aatgggtgtg	tcttggcttg	ctgcaacctc	10740
tgccctcccgg	gttgaagcaa	ttctcatgcc	tcagccttct	gagtagctgg	gactacaggt	10800
ggatgccatc	acacccgatt	aattttttgta	tttttaggag	agacgggggt	tcaccatggt	10860
ggccaggctg	gtctcgaact	cctgacctca	agtgatctgc	ccaccttggc	ctcccaaagt	10920
gctgggatta	caggtgtgcg	ccaccacacc	cggctgtttg	tgtggttgct	tcttgaagct	10980
tccgtattct	tctctctaaa	atgaggataa	tgggtgaaaa	ggccaaacct	ggtttgccct	11040
gccaacctgg	catgagctta	gacagaaaag	ccccacttca	gcctttttcag	ggttgagggg	11100
ctgtgaggca	agccaagccc	ggggctccag	tgcccttctg	actggagggg	gagggttttg	11160
ggaggaggga	ggatactact	ccacctccac	ccctaggcat	ggcagagcca	gctcagcatc	11220
ttggttcaga	aaaattccac	ttgctgtata	atcttgagca	gaggccagaa	gacctgctcc	11280
agtggactcc	tggactcatc	caggcctgcc	atgcaggggtg	ttcaggcttt	ctgagctctg	11340
gactccaagg	aaccttacag	gaaccttgtg	ggtgattttc	cttccctggg	aactccaagc	11400
tctgggggtg	gaggggtata	aattggaggc	cacaggcctg	taaccccagc	attttgggag	11460
gccgagggag	gtagatcaac	tgaggtcagg	agttcgagat	cagcctggcc	aacatgggtga	11520
aaccccgtct	ctactaaaaa	tacacaaatt	agccaggcgt	gttggcaggc	acctgtaatc	11580
ccagctactt	gggagggttg	ggcaggagaa	tcacttgaac	ctgggaggca	gaagtgacag	11640
tgagccgaga	tcacaccact	gtactccagc	ctgggtgaca	gagtgagact	ctgtctcaaa	11700
aaaaaaaaaa	aaaaaattag	ctaggtgtga	gggtgcatgc	ttgtaaattc	cagctacttg	11760
ggaggctgag	gtgggaggat	cgcttcagcc	caggagatgg	aggctgcagt	gagtcatgac	11820
catgtcagtg	ttccccagac	tggatgagac	catgtctcaa	aaaaaaaaaa	aaaaacaaag	11880
gatcaatgct	agccccattt	tacagaacag	aaaataagag	tcagaagtta	agagtgttga	11940
ccaatgcgct	ggacgagttg	gctcatgcct	gtaatcccag	cactttggga	gaccaagggtg	12000
ggcagatcac	ctgagggtcag	gagttcgaga	cctgcctggc	caacatgggtg	agaccccatc	12060

WO 2006/119775

PCT/DK2006/000254

tctactaaaa	atacaaaaaa	tcagccaggt	gtggtggcag	gcgccctgtaa	tcccagctac	12120
tcaggaggct	gaggcaggag	aatcgcttat	acccaggagg	cagaggctgc	agtgagccga	12180
gatcgtgcca	atgcactcca	gcctgggcaa	caagagtgag	actctgtcta	aaacaaaaaa	12240
aaaaatttga	ccagagccaa	ggctatgtag	ccagaagggtg	gaggatgcag	gattcaaacc	12300
caggcctggg	tgtggagccc	atactcccac	ctgccatggt	ggagctgatg	tgtggcttgg	12360
cttgcttgaa	gtcaccaagt	tgtaggtgac	aaagctaagg	tttgaaccca	gatgtgtgac	12420
tctgaaacaa	ctgaccctta	tgccgaacca	tattctagac	catcactggc	caccctaccc	12480
agccacctcc	caggcccctg	gagttaagag	aattgaactg	caaactggaa	ggacttaggg	12540
gtcatatgac	ccctagtttt	gtttttcttg	tgaggactct	gaaccacgat	ggggtcgttt	12600
gtgacttgaa	gtaacgccgt	gaggatagca	ggtgggacca	gccctcctca	ctgtggccag	12660
ggttcaggag	ctgctcccct	ggactaggct	cacctgtgcc	tctttcttcc	tcttccagac	12720
ggagcccgtg	gctggtgagg	ccgcagagca	gggccaggcc	gggccctgcc	cgcaaggacg	12780
accaaaccct	tcaccggccc	ctgggccccca	gagcccacca	ggctctcctg	ctggccccac	12840
catgacttca	gagtcgacgt	cacccccagt	tgtgcccccg	ctccactccc	ccaagtcccc	12900
agtctggccc	accttcccct	tccacaggga	gggcagcagg	gtctgggagc	ggggtggtgt	12960
cccacctcgg	gacctacca	gccctctgcc	caccaagcgg	accaggacct	attcagcgtg	13020
agtacctgcc	ccttgcccag	gcccctgacc	cttgtcagga	caggggctga	ccatccccat	13080
cctcttccaa	ctcctgttgt	cctcacctgg	gacaggggaa	ctgtgtgtgg	accttccagt	13140
aactggatgg	atgtttgtcc	catacatcca	ttcatccatc	cacccatcca	cccatccatc	13200
catccatcca	tccatccatc	caccaccca	cccaccacc	caccaccca	cccaccacc	13260
cacccatcca	tccatccatc	catccatcca	tccatccatc	catccatcca	cctgtccatc	13320
catccatcca	tccatccacc	cacccatcca	tccgtccatc	catccgtctg	tccgtccatc	13380
catccaccca	tccattcatc	tatccatcta	ttccacaaag	atttactgag	ctcctgacca	13440
gtcaggatcc	ataagctgct	tacaatccag	agggaacttg	ttcaccatcc	attaccacc	13500
agcagggtta	ggggtgagga	cccagcatgg	ggtaccagag	gagatgatgt	cctgggttggg	13560
tatggaagat	gtatcaggtg	aagaaggagg	ctggcgattc	tgggcagatg	gtacagcatg	13620
ggcaaaacc	aagacacgga	gttaaccttt	ggagaggggt	aggtaggcgg	gacaaagact	13680
aggcctgggc	ctttggggta	ggtcaccgat	gaggtggcat	aggggagtaa	tacgggtcca	13740
tatgggggag	gccaaaggat	cttttgggaa	ggacagggtc	tcgctctggt	gcccaggctg	13800
gagtgcagtg	gtgcagtcac	ggctcactgc	agcctcaact	tcctggcctt	cagcaatcct	13860
cccacctcag	cctcctgagt	agctgggacg	actggtgcgt	gccaccacc	ccagctcatt	13920
tttttaaaaa	aatgttttgg	ccagacatgg	tggctcatgc	ctgtaatccc	agcacttttg	13980

WO 2006/119775

PCT/DK2006/000254

gagggtcgaga	caggtggatc	agttaaggac	aggagatcaa	gccagtatgg	ccaacatggc	14040
caaacactgc	aaaaccctgt	ctctactaaa	aatgccaaaa	cttagctggg	tgtggtggtg	14100
ggtgcctgta	atcccagcta	ctcaggaggc	tgaggcatga	gaattgcttg	aacctgggag	14160
gtgtaggttg	cagtgagctg	agattgcacc	cctgcactcc	agcctggggtg	acagagaaaag	14220
accctgtctc	aaaaaaaaaa	aaaaattttt	tttgtagaga	tgggagtctc	cctatgttgt	14280
ccaggctggg	ctcaaactcc	taggccaag	tgatcctccc	gcctcagcct	cccaaagtgc	14340
tgggattaca	ggtgtgagcc	accgtgcctg	gcctcattat	cgttgaatgg	ggatgactgt	14400
gtggtttctg	gcctgggtca	agtttctcac	agccctgagg	cacgctcagc	agaaagggcc	14460
cagaatggct	gtagggatcc	ggggaggagt	tcctcaggcc	acagctccaa	tcccactacc	14520
ctatctccag	gacagcccgg	gcctcagctg	gccccgtgtt	caagggcgctc	tgtaagcagt	14580
tctcacgctc	acagggccat	ggcttcatca	cccccgagaa	cgggtccgag	gacatcttcg	14640
tacatgtgtc	tgagtgagtc	ccctccacct	ccctgttctt	ggcccctgcc	ctagtgtccc	14700
caatggtgcc	tcaaaaatgg	ctctccaggc	ctctcctctc	actggctgag	gtctgtattc	14760
atggtgcctt	ttctgtcctc	tgagtagggg	ctgaggggtgc	gttgggagag	gagggctggg	14820
gccagggagg	gagtgggcac	gcagacccca	gaagcctggg	agccagcggc	tgggccctgg	14880
atcctgttac	agcatcgtgg	aggtctgaga	gggctaggcc	agctcctgag	cctcttggcc	14940
tggaaagggc	cttcatgtgg	gtgggcccagg	aggtagagat	ggatatcatg	tctaggcctg	15000
gccccaccca	cttctaatac	caggaggacc	gggggaggtc	actgagcttc	tcagagcttt	15060
agggacctcc	ccgtgccctt	tcacatctcg	gtccccaagg	gaccagaaaa	gggaggtaca	15120
gtttatgagt	cccattttat	tttattttat	ttattttttg	agagggagtc	tcgctctgtc	15180
gcccaggctg	gagtgcaatg	gcacaatctc	ggctcagtg	aacctccact	tcctgggttc	15240
aagcgattct	cccacctcag	cctcccaagt	agctggaatt	acaggtgcca	gccaccatgc	15300
ctggctaatt	tttgtacttt	tagtagagcc	agggtttcgc	catgttggcc	aggttggtgt	15360
caaactcctg	actcaggtga	tccgcccggc	tcagactccc	aaagtgctgg	gattacaggc	15420
gtgagccact	gtgcccagcc	atgagtcccc	ttttacagat	tcgagagctg	aggttctgag	15480
aagcaaggcc	tggccaggcc	tcctggccta	gggaggctta	gggccaggct	gcccccaagg	15540
ccaggccaca	ggattgtctg	gggcacaggg	cctgctgctg	ggctaatacc	cctcttcctg	15600
cccgccagca	tcgaggggga	gtacgtgcca	gtggagggcg	acgaggtgac	ctacaagatg	15660
tgccctatcc	ctcccaagaa	ccagaagttc	caggccgtgg	agggtggtgct	cactcagctg	15720
gccccccaca	ctccccacga	gacgtggtct	ggccaggctc	tgggctccta	ggctgagtgg	15780
ttcacaggcc	agctggccgg	gggttgggga	gccacacagg	gtgaacgggc	agcagccggc	15840
tccatgcccc	actgccctgg	ctgatgagtc	cttcggtggc	ctcagtgctg	acgtctgtct	15900

WO 2006/119775

PCT/DK2006/000254

gtccgtctgt gcttgtggct atgagcgtgt gcctccaccc accccacgac ccgtgactac 15960
 tgtgcaagct ggccaggagg taggtggagg gcaaggccac cagacctggc ttcgcgctgt 16020
 ccactgcctc tctcctcccc tccctgccgc agacacgcag gacccgctcg ccctcctgct 16080
 tacccgctccc cacggtgact gagctgcgag agcctgcgct gggctcactc cctggcctcc 16140
 gcccctgagc tgtcccgtgt cagtccctgg cccggagcca ggctctctgt gctacttaca 16200
 tctccctccc ttggctgagg tccccgcca cagcctggac cttccctca gaagcctggc 16260
 cgatgggtgc ttttggggag cctctcagcc tctggcaggg gaggaggact gacacggagg 16320
 cacacaggcc tctcactgcc gaggccagcc ttccttcctg ctctgctctc tcagtgccat 16380
 ggagaggcag gaagaggggc ctgagggcgg aaggggtttg gggctcccat ttcgccggcc 16440
 agtccctcct cctcagcctg gcagtggcgc tgggctcctt cccctgggct gtactgagcc 16500
 gagcccaggg gtttgcagag ggtgggggtc catctctcca gcttggctgc agacctcctt 16560
 taccctgact cacaagcccc actgatgctc tgggccatgc ccggctgacg gctgctgtgg 16620
 atgccgagcag ggacggacac acgtccgggg caccacagag gagggcccc agcctgggag 16680
 actggccctg cggcctccac gtcaaactct cttcccaaag ccctaacaga ccagtggccg 16740
 aggtgtggct cttattgcat ccatccctga agatgtgtgg ctgttgctgt cacctggagt 16800
 ccttttgggg ccaagatgtg tgtgcacccg gggtcgtggc cattcactcc caggcagggg 16860
 tgaggggtggc ctggcccagg aggccaggaa ggagggccct gtctgcctcc acctctgggt 16920
 gcaccccctg cctaccaccc tcccttctag agagcacatc gcctgaccgg ggagaagtgg 16980
 ggccgtgggt cgagggaggg ctggccaggg gtgggaccct tatgagactc agtctgtgag 17040
 taaaactggg ggctcaaagt cccaggatga ggggatcagt gactgtctag gaggatccct 17100
 tgcctttagt gtgccccaa accgcagggt agaaatcagc cgggatgcct gcatcccacc 17160
 cccggcccca gggcccacca cataaaatct gggagcccag agctgctgag gtgtgggtcag 17220
 ctcccctaaa atgggcacgg cccagcctgt cccatgagga ataaaggccc ctggccccct 17280
 tgagagaggg cttgttggtg agggctgact cctggggggc cccaaggctc ccctcctgtg 17340
 gggaggccac cttttaaggc accactagca gtcaggatat ggtggcatca gccctggggc 17400
 ctcttggggtg gcaggggggc agcagctctc ctccccactc acaggccctg cagtgtcat 17460
 ttggaattcc tccaagacc cctggccacc cagaccccc attcttccta aactggcaa 17520
 taaaccctca actgtgaccc accgtggcct gaagctgcct ctgtgtcagg gtccccaagg 17580
 aaggggctgg ggggtggacac accaaaccag ctcttctttg gggattcaat ggccccctcc 17640
 tcgactgcat cagacttcct ttgtgacctg ggccaagtcc ctgccctctc tgtgctgttt 17700
 acttgtcggc aaagtagtgg ctgtgacagc atgtgactgg ggccaattat ccattcagac 17760
 acagctgaca cctactttgt tctgggaact ttaatactgt gacaaagttc tctaaaatag 17820

WO 2006/119775

PCT/DK2006/000254

```

gcaccttccc caccgtacct catcgcccag ggcaggcagg cagggcaggc tagatctcgt 17880
accgatactt gagcacgcct cctcctggtg cagaaagaaa cctcttctgt accgaaatac 17940
aagcagcagc tgtggcctgg gccaccagggt ggagcatggg gaacactctg ggccctggga 18000
ggacgaagcc agtgccacta ggagcagact ggctggggac ggttgtccac acagactctg 18060
gccccatctg ggtgggcttg cagcaggcgt cctggggccag aggagggggc ctggcatcta 18120
tccaacacca ggacctctct ttaggcctag gccaaactct tcagaagtca cgacacacag 18180
atgtgggccc cggtcacgcc tcatgagctg tctctgggaa gaaaatctcc cggcatcgct 18240
gcaccgtgtc ctgaggagac accacgctgt ggccaacagt ccggggggtcg gcaaagatct 18300
caaagtcgtt cccaccctgc acacagagga tgggcagagg agggatattt aggtggagtc 18360
cctctccaca ttggctgcta cggacctgcc cctcctcctg ccccttcaga actgccagga 18420
agcccacctt ctcacccctt gtcagacctg ccctgacttg ctgtgtgatg tagaacaagt 18480
gcccagcctc ttggggcccc agctttccaa tctgtgtaac ggggctggtg agcaatggtt 18540
gttctccagg agccccctaa gaggtctggt gtgttgcccc atgttctggc cttgacctca 18600
ggtgatttgc attggagaag ccaccagacc tggcttctca ctgtccactg cgttagggga 18660
aacctgggggt gggccacact cacagggcta gtctcgttcc caaagaagtg gatggtgtcg 18720
aagctgtcct ggtccaggct atccaggcag tagcgcttgt cccagccctc ggggaagacg 18780
tcaaagctga tcatgcctcc tgtgggcagg ggtggggact gttattccct gctgggaggg 18840
gcagacaagg ctaacagaag ggctgcttgt tcattcattc attcaacaac ccttgagtgc 18900
ctactgaggg ccaggaatag gacacgggac agacttagga tcctcaagga agttatgaca 18960
caggttggag aatgaacaag accatgaagt ggggtggggga agctcatccg gccagggtga 19020
tcaggggaga cctctgagga ggtggcaatg agctgggacc tgagccgagt ggggatggct 19080
ggggcgcaaa gtacagcccc tgaggcaggg aagagaacga gcgtcccagc aacggaggag 19140
gccaggagag gtagggggag gacgggcagg ccctgggcat cacatggaga gggaggaaga 19200
gcctggggag gactttgggc agaggcatgt catgactcca gctccagcct ctctggccac 19260
tggggaatgg gtgaggtggc agaggggagc ccagttgtgt cagggtcagt tggagcctgg 19320
gacaagctgc tggcagtgga agtgtgagca ccggtgggat tccagacagg gggccctgtg 19380
acttctagga gtttggactt gggaaggggt gggggagagc attcaaggga gaggctctgg 19440
ttttttggcc tgagaaactg ggtgggcatg gggcgtttac ctctgaggta tgaggtgggg 19500
aaaactgggg gccagtaaac actgaacccc tttctgcatc gg 19542

```

<210> 5
 <211> 58404
 <212> DNA
 <213> Homo sapiens

<400> 5
tcctcttctg tgaagggttt cgtgtctttt catcttctct ttcattgagac tgtgtcttac 60
ttttgctctg tgttggtggag aacctccagc ttgggtttcca tctggcttcc ctgaggtgga 120
ggagagaagc gttcgggctc agcttgcttg tttggccaca ggtaaacaga catctgtaga 180
atggaggtga cacgaagctc ctgggtctgg tgcagaagag gtcagggccc cagtggcctg 240
taccctctaa gctcaagggt gtgtctgcag ggagtcctag atcaggactg accactgggg 300
ccttggtctg ggtcctacct gcccttgagt cgtcacgtca caaccacac actgggtgcc 360
atgggcatga ggactggcca ggaaaaggcc tgtggggcct ggcccctatc agggacactt 420
gggcgaacaa ggccacccca ctccaataa ggtgttttgt gcagtgcccc gagcagtggg 480
gccgctcggc tgcccccggt gggacggtct gtcccagggt ggacagctgc tcttcctgag 540
gggcctgcct ggatgccaca ggctgtgctg ctggaacca gggattccac agcaccatcg 600
tcctgatcac aaattctcaa ggaaaccaag gagctcctgc ctcaccttaa aacattttggg 660
atttaatctt ggagtgaana ttgttatgac cccactgaat tttagagttt acacaaaagg 720
ccaaagtga ggcagagact caggcctctc aactcccacc cgcagggggc ttttcttgca 780
accttaaagg gcatctgctt cacgccggcc tcggcagggc cctcacactg gcaggactac 840
ttgcacaccg ggcctgacac tcactagggg aggctcccca gcccggtgc accaccacc 900
tctacctagt cgcgtgtccc ggccagtgcc ctttccactc ccacccagt cccccgcagg 960
gccccctctg tcctcttagg gacctccact cccggaccaa tagcctcctg cctcagggcc 1020
ccccactcat gggccaatcg ccccctgcct caggcctccc cccactcgtg ggccaatcgc 1080
cccctgcctc agggccccac caactcgtgg gccaatcgcc ccctgcctca gggcccccca 1140
ttcctggggc aatggccttc tgcctcaggg actcccactc tcagctgggt cccccgcct 1200
cagggacccc cactcctggc cgttcccccg tcaggggccc cgactggcgg ctgctcccc 1260
atcctgcctc ggggcccccc actcccgggc caatgtcctt ctgcctcagg taccctcact 1320
cccggcccgt cccccgcct gtgggcccc cactccaggc cgggtggtag ccctgcctca 1380
gagccgcca ctctggggc atctcctgcc tcagcgacc cactcctggg ccaatgccct 1440
tctgcctcag ggacccccac tcccagccgg taccctcctg cttcaaagcc cccgactccc 1500
ggctgctccc cacctgcctc ggggcccccc actcctaggc ctacgtcccc ctgcctcagg 1560
gcccctcatt ctccatctc ctgacccct cagggacctt ctggcccgg cctccctaac 1620
tcaagggccc ccactctcca tcctcccag cccccactcc cagcccgggt cccgcccct 1680
caggggcccc tttggcccc gctcctcgcg gcgccccct cccgcgcgc ggccgaggca 1740
gcggagtccg cgcgtcatgg cggcggcgg cggggcgcgt ctgcacgca gcgcgccggg 1800
cggagcgtc gccgcattgt ttgcttcgct ggggagcgag cgagcgacc ggccctggag 1860
tccgagccag cccgcgtcc cgcgcccgc gcccgccgc cgtccccgc cccgcgccg 1920

cgcccgccgg ccgggcccc cgagccgggc ctccagctgg gccccgagcc cgagcccgcg 1980
 ccgcgcgagc cgctgcccga gagtccgcgg ccgccgctgc cgggcccgga gccgccgcgg 2040
 ccgagagggg cccgcgcccga ggagcggcgg cggggcccgg cctgcgcggc cgttggcgga 2100
 ggagccgcgg gcgccattag caccgcctcg gccgcgccgg ccccgcgcc cgcccgcctt 2160
 ccgggctccc gcggccgcgg cggcccgaag gtgagtgggc gggggcctgg ccggggaggc 2220
 ctccccgccg cctcgcgggc cgcccggccg ccgccgccgc ctcgggccgg cctctcgggg 2280
 ttgtggtgcc gacgccatgt ttggcggcgg cggcggcggc ggcggtccg cattgtctgc 2340
 gggccgggag cccgggatgc cggggaggcc gggcggggag accggggagg tcgggcccgg 2400
 cggccgggga gacgggcggg aagacgggga ggccgggcgg ggagaccggg taggtcaggc 2460
 cggggagtcg agggaggccg agtcgggctg tgggtctcgg cggcctgggc cgcgggcgcc 2520
 tgcccggagt ccgcgctcga gtcgggtcgg gcctcgggca gccgccgagc cggccgggccc 2580
 gccccttatg taacagacag gcccgccgag ctctgccaag acgggaactc cctgtgagca 2640
 gggacgcgcc ccgagatcgc ccaggagccg cgcggcctgc ggcgccggg cttgggaacc 2700
 ggcaccgcgg cctccccggg cgtttgcagg gagacgcctc gcgacgaacg cgaggcttcc 2760
 ccgggccttt taaaacgagt gacacgattg aggcgcagtt cacgggaaac gcggctcttg 2820
 tgagcggacg cgggagttct gagaatccgt agcagcggga cccctgccc cgcggggtgc 2880
 ggccccgtct cccagaaaag cgcggccgcc cgtgtctgtt tttgagggtc tgtcgccagg 2940
 tggttgtgta gttagtggaa tcggcggatg tttttgtttg gccggcccag cgccgtcccg 3000
 agtgctgtgt gctgcacaca acttcgcaca gctcagcggc cgcggcccct gtggccttga 3060
 ccgagaggcc cggcgcgctg ccggcttctc cctgggcggg cgggtgggaag atggaagcgg 3120
 ccgtcttggt ggcagccttc cccgcctgtg gcccgcaggg ctcttcgatg tacttaaggg 3180
 aattattcct tgagaagcta cgtgtgtgcc cccgcggatc cctcccagtg gaaatgcttg 3240
 tcccagtgga agggccaggg acagaaccgg cccagggggg aagcggtgaa gcagaggagg 3300
 tcgtgggtcg ccctgccctg ctgagggcct actttccccg ggtcccagtt gcatgttgag 3360
 ggtatggagc tgctgtggcc gtgaagcctt cggttcagtt ctccccttgt agaaagtgtt 3420
 gcgatgaaga gcttcacggc tccctgggag agagtgggtt tccaggggct gcttgtcggt 3480
 agggcgactc aggactttgc tgtcgagggg aaaatgacag gttcgtttct gatctgggtga 3540
 gtaaccaact gttgcttctg cccacactca gtaatgacgc gtgtcggcgt cccgggtgca 3600
 ctagggtcgc tctctgcagg ctgtgggtgt acctctgcag ggctcacggg gccggctggg 3660
 ggggtggctgt gctgacagct cccaggaat ccagggttga gtgggtacct ggaggccctg 3720
 gagatgcagg gggctctgagg tttggccgac cacatcttct ctctattcc tttctgtagg 3780
 taattgcaat ctagtttcaa catacaatga ggatttctgt ctgatacaga gaaagatacc 3840

cagtgtctgg	agggctctgtc	ttaagaagtg	ctcttgaagt	ctgctgggat	ttcctttcgg	3900
gactcacaga	gggaagcctg	ctccctgagg	gctttttctc	aggcccaata	gatcagggta	3960
aggctgctgt	gggtgcccct	gatccccctg	ggccccgagg	ccaggcgtcc	ctgacgtggc	4020
agggctagat	gcagagtatg	ccttgggctt	cagctgagaa	tatttgtttt	gcgagtcagt	4080
cactgaaaaa	tcattttgct	tcctgtcagt	gaagtgtggc	gctttggagg	gaggcactgc	4140
ccagcaggcg	gtggtcttca	caggaggtgt	ccacacacca	ggcaccctgg	ggacagctag	4200
gaatagacgc	agtccgtggt	gttgggacag	cagaggcctc	tgccacgagg	agctggaggt	4260
atgggaggca	gctcactggt	tttgcctacc	tctggaaaag	cagattctcc	taagcctctg	4320
tggccaagtc	ttcaggcacc	tgagccagcg	ctggcgaccc	tcctggtgag	ggtctctgca	4380
agcagcgagg	ggttcctggg	ggtcaccctt	cgctgtggag	cctgagtgga	gcgtctgcct	4440
tcacgtgtgc	cgggggagaac	agagattccc	tttggcgcta	tgcattgtgc	tggattctct	4500
cagaaaggag	gaggacgtgg	aaggaggagg	acgtggccag	caccatgggtc	tcagtggggc	4560
ttgaggagtg	cccctgaggg	gaaggggcag	gttggctcag	cctgtgggtg	tttccatggt	4620
ggtcactgcc	aaggctctgc	agaagcgggg	ttggggcctc	ccagggggcca	agactgtccc	4680
ctcagaccag	tacagggtc	ccctgttggt	cccctgtgcc	gtgttggtgc	tgtggagggc	4740
ctgtctatag	aggatggaca	ggctgtgatc	tgtctcctac	ccctggcccc	agcaggaatg	4800
ctgcaaggac	tgtggtgact	ctgggctcat	gggagtgtc	tcctgggagt	agtggggagg	4860
ggaagagtgg	tttaggggga	gcatccgata	ggaccaggta	gccactgagc	atgcctgtct	4920
ggtagggcata	gcctctgtgc	ctggtggggg	ccagccctgt	atgtggactt	cctggggggac	4980
cgcttgcatt	ctgggaaagt	cctgcaactg	tgggctctgg	gccggacccc	tcctccttgg	5040
ggagcgttgc	tgtggtctcc	gcagcctgtt	gggtgtggtc	gagaggggac	tgggtgtgtg	5100
tggtagggac	gcaggtgtc	aggagggctc	ctcaggctgc	ctctctgcct	gttactgctg	5160
ctccctccct	ggtcctgggg	actcgctcag	gcctggatgt	gtggtttttg	acaccttgga	5220
aagaggttag	ctgtgagttt	tgctttgatt	gggctttgag	ctgctgagct	agatctgtcc	5280
aagcatgctt	ttgtggtttt	ctcgagtctt	tggtcaccct	gcgtagaatt	acctttactt	5340
ttgccttcat	tttaggtaat	cattaccaa	tgaggaggaa	aggacgatgt	catcgaggct	5400
ctgcagcgag	gcatccttct	tccccatgca	gtgttaaaca	ctcccctacg	cgagaaacgc	5460
tgacctacgc	tcaagctcaa	aggatggtag	agatagaaat	tgaagggcgc	ttgcacagga	5520
tcagtatttt	tgatcccctg	gagatcatat	tggaagatga	cctcactgct	caagagatga	5580
gtgagtgcaa	cagcaacaag	gaaaacagcg	agcggcctcc	tgtctgctta	agaactaagc	5640
gtcacaaaaa	caacagagtc	aaaaagaaaa	acgaggccct	ccccagcgcc	cacggcacgc	5700
cggcctcggc	cagtgccctc	ccggagccca	aggtgcgcat	cgtggagtac	agccctccgt	5760

ccgccccccag	gaggcctcct	gtgtactaca	agttcatcga	gaagtcggcc	gaggaactgg	5820
acaacgaggt	ggagtatgac	atggacgagg	aggactatgc	ctggctggag	atcgtcaatg	5880
agaagcgcaa	gggcgactgc	gtccccgccg	tgctgcagag	catgtttgag	ttcctgatgg	5940
accgcttcga	gaaggagtcg	cactgcgaga	accagaagca	gggcgagcag	cagtctctga	6000
tcgacgagga	cgccgtgtgc	tgcattctgca	tggacgggga	gtgtcagaac	agcaacgtga	6060
tcctcttctg	cgacatgtgc	aacctggccg	tgcaccagga	gtgctacggg	gtgccctaca	6120
tccccgaggg	ccagtggctc	tgccgccact	gcctgcagtc	gcggggcccg	cccgccgact	6180
gtgtgctgtg	ccccaacaag	ggtgggtgcct	tcaaaaagac	agatgacgac	cgctgggggtc	6240
acgtgggtgtg	tgccctgtgg	atcccagagg	tcggctttgc	caacacgggtg	ttcatcgagc	6300
ccatcgatgg	ggtgaggaac	atccctccag	cccgggtggaa	actgacatgc	tacctctgta	6360
agcagaaggg	cgtgggtgcc	tgcattccagt	gccacaaagc	aaactgctac	acagcattcc	6420
atgtgacgtg	tgcccagaag	gctggcctgt	acatgaaaat	ggagcccgtg	aaggaactga	6480
ctggcggttg	caccaccttc	tccgtcagaa	agaccgctta	ctgtgatgtc	cacacgcctc	6540
caggctgcac	ccggaggcct	ctgaatatatt	acgggggatgt	cgaaatgaaa	aatggcgtct	6600
gtcgaaaaga	gagctcggtt	aaaacgggtca	ggtccacatc	caaggtcagg	aagaaggcaa	6660
aaaaggctaa	gaaagctctg	gctgagccct	gcgcgggtcct	gccgaccgtg	tgcgctcctt	6720
atattcccc	gcagaggtaa	gcgtgtccaa	gcattgagtc	cacaaggagc	ctgcagtgtg	6780
ggaccctggg	aggacgtggg	cacagcctgc	agggcccggtg	ctgtgtgggtc	attgcgtttc	6840
tagttggcgc	agcttcctaa	tattcccctg	aagtgtgggtt	ctgcttggtt	accacgtctt	6900
tttggggaat	agatgttgta	aaagattttta	gtctgtaggc	ttggaataat	ctgtagtttt	6960
ctatgttttt	ttttttggag	acggagtttt	actctgtcgc	ccaggctgga	gtatagtggc	7020
gcaatctcag	ctcactgcaa	ccccgcctc	ctgggttcaa	gcaattctcc	tgccctcagac	7080
tcctgagtag	ctgggactgc	aggcgtgcgc	caccatgccc	gacaaatttt	tgtatttttt	7140
aagtagagac	ggggtttcac	catgttggtc	aggctgggtc	caaacgcctg	acctcaagtg	7200
atctgcccgc	cttggtcttc	caaagtgcctg	gaattacagc	acttgctcag	cctttttttt	7260
ttttaaatag	ggttttgctc	tgttgcccag	gctggagtggt	agtggcacga	tttcagctct	7320
ctgcaacctc	tgctgccagg	ttccagcaat	tcttggtgtc	cagcctcctg	agtagctggg	7380
actacaggcg	tgtgctccca	tgcccagcta	attttgtatt	tttagtagag	gcgggggtttc	7440
accatgttgc	ccagggtggg	ctggaactcc	tgagctcagg	cagtctgcct	gcctctgcat	7500
cccaaagtgc	taggactata	ggcgtgagcc	accatgccc	gccgtgcgtg	aagatttttta	7560
agaatttcag	tgactaaact	ttaagaagga	atgaagccc	cgcaggaact	ctgtatgact	7620
ttcaaccgga	ggggctctgg	cctgggcttc	ccagtgactg	ccggcatcgg	gacatccact	7680

gctgtggcct	tggccccgtt	tcccgtgctg	tgccccccac	ttgctgcaga	tgccctgagg	7740
aaaggggtggc	caaacaaggg	gaggagctgt	gagcacagat	ggggcctgcc	ggtcaacagg	7800
gcgggtagag	ccagcctggg	gctgagcttg	cagcttggct	gggggagacg	cctagagccc	7860
cagtaccaca	cagaccctcc	aggactgccg	ggggtcagtg	agcaggaccc	tctgaggtgt	7920
cagacctgcc	ccacatgtga	ggggaggaca	gcacagggtg	gacgcatgca	ggctgagctg	7980
catcgggcgg	tgaggctggg	aggttgcagt	gagctaagat	tgtgctactg	taccgcagcc	8040
tgcgtgacag	agtgaactg	tgtctcaaaa	aaaaaaaaaag	aagttttaag	tgatTTTTgt	8100
taatattact	ctactttacc	aaagtgaagt	ccagtttctt	ttcacagagc	tgtaaatact	8160
cagataatta	gtcaattcag	acagaatgaa	caaaaaatgc	tcataatcct	tcccatccag	8220
ggaaaaccct	gtcaaaaacca	aggtgtttgt	ctttgtaaat	taatctgact	ttgaagagat	8280
ggTTTTgctt	ggattttctac	ctatgagtca	aaaagtctag	ggtttTgtatt	aagagcttct	8340
gtataagagg	tctaaaaatt	tcagctttac	tgaaattgaa	gcaaaataaa	caaaatgtga	8400
tctaacaaaa	tgaagagcca	aaagtctgaa	gtgctccctt	ccctggccag	taccagatga	8460
tgacctgctc	tgggagaggg	cgggcctggc	cgggccgcac	gctctctggt	ctttggctcc	8520
tgatgttggg	taggcaggac	tgaaggaggc	agtgtttgca	tgaggacat	gggggctgaa	8580
gtggccaggg	ctgattgaga	agaaaactcg	acctaaatat	ttgaaagcat	gtattggagc	8640
taaaccgcgt	actgtgtggt	tctgtttctt	tcctaacttg	ggaataaact	attatttcag	8700
cttcgagtag	agagagagaa	acaaacctgg	acgagtgttc	ctggccgcgg	tctctgcagt	8760
ctgggtttcc	cagtctccag	gagatgctgg	tgcttcaggc	caccttgcca	ccagtggagg	8820
agctcccagt	gcgcctggcc	tgttggcacc	atggtgctgg	accacctct	gggggtggtt	8880
tgaggggagc	gctctcaata	cagtactgac	tgctgcctga	aaccgggttt	actgccttgc	8940
tgtggtctga	gaggcagctg	atctggcagc	cctgtttctt	ggccctgggc	actaccctg	9000
ggtgcatgta	ggaggcctca	gctgtgcatt	tgaattctgt	gtagatgagg	gtgaagcttc	9060
ttatTTTTta	TTTTatgttt	agagacaggg	tcttgttctg	ttgcccacgt	tggagtgcag	9120
tggcgtgatc	atggcttaca	gtagcctcaa	actcttggac	tcaagcgatc	ttccgcctc	9180
agcctcctcc	cagcctcagc	tgggaataca	ggcacgagcc	actgtgcctg	gcccatggtg	9240
aagcttcttg	atgtgcagcc	gtgggcgggg	atgctgagtc	tagttggtta	tgggtgtgcac	9300
taaccaatgg	ctccggtgta	ggaccacact	agagcatctg	aggatgttat	gcttcaggcc	9360
agccagtcac	gctggaaggg	agatgagacc	ccatctgagc	agcgttttag	agatTTtaga	9420
gaaagaggat	cttccttatt	ccctgtggac	agaacagaat	ctgccttcag	ggacggactg	9480
cttttataat	tctagtgggt	taaaaaatta	aaatttcaag	taaaaaagat	agttctTTTT	9540
tttccccct	TTTTTTtgag	gtggagtctt	gctttgttgc	ccaggctgca	gtgcagtgcg	9600

gtgatcttgg	ctcactgcaa	cctccacctc	ctgggttcaa	gtgattctcc	tgcctcagcc	9660
tcctgagtag	ctaggattac	aggcacctgc	caccacaccc	ggccgatttt	tgtattttta	9720
gtagagacgg	ggtttcacca	tggtggccag	gctggctctt	aactgctgac	ctcaggtgat	9780
ccacccgcct	cgaactccca	aagtgcctgg	attacaggcg	tgagccactg	cgccaggcta	9840
aaagagatag	ttttaagggt	ataatctgtt	tttttgtaag	tttgatatac	tatgattaaa	9900
taataacagc	ggtaaataaa	cactttgaaa	aacaggcagg	gctgggtgtg	tggctcacac	9960
ctgtaatgcc	agcactttgg	gaggccgagg	tgggtagatc	acgaggctcag	gagttcgaga	10020
ccagcctggc	taacatgggt	aaaacccatc	tgtactaaaa	atataaaaat	tagccaggcg	10080
tggtgggtgc	cgctgtaat	cccagctatt	caggaggctg	aggcaggaga	atctcctgaa	10140
cagagtcctg	ctcttctgct	cagacgggtg	cagtggcatg	gttatggctc	actacagcct	10200
cgacctcctg	ggctcaagcg	actccccac	ctcagcctct	tgagtagctg	ggaccacagg	10260
catgtgccac	catgccctgc	taatgttttt	atgttttggg	gagatggggg	caagcaatct	10320
gcctgcctcg	gcctcccaa	gtgctgagat	tacaggcgtg	agccaccatg	cctggcctta	10380
cattgattta	ttgttatttt	ttcttttctg	ttcttctttt	ttttcctgag	atgaagcctc	10440
gctctgttgc	ccaggctgga	gtgcaatggc	ataatctcgg	ctcactgcag	cctccacctc	10500
ccaggttcaa	gcgatcctcg	ttcctcagcc	tcccagagtag	ctgggactac	agggtgtgtgc	10560
cacgggtgtct	ggctaatttt	tgtatttttt	ttgtagagat	ggggtttcac	catgttgtcc	10620
aggatggtct	caaactcgtg	acttcagggt	atctgcctgc	cttggcctcc	cagagtgcctg	10680
agattacagg	cgtgaaccac	cgcacccagc	cctatttttt	cttttcactt	cctctctcca	10740
gcatacatca	gtttttatta	acggattgag	ttgcacctgt	atagaatgac	atgttttttt	10800
tgagacggag	tcttgctctg	tcacctaggc	tggagcgcag	tggcacgata	tcggctcact	10860
acaagctcca	cctcccgggt	tcacgccatt	ctcctgcctc	agcctcctga	gtagctggga	10920
ccacaggcgc	ccaccaccac	acctgcta	tttttgtatt	tttagtagag	acgggggtttc	10980
accatgttag	ccaggatggt	cttgatttcc	tgacctcgtg	atccatctgc	ctctgcctcc	11040
caaagtgcta	ggattagagg	cgtgagccac	cgagcccggc	caggaggaca	ttttgtcttt	11100
agtcttaaag	gaaaagggaa	agtcttttcc	atgtgcacca	gtttgaaatt	agggtccct	11160
taacatggat	tactgcagga	ctcattcaca	tgagagttca	tacatggggg	gaaaagacaa	11220
acagggtttc	ctgataactc	tttttgggtt	ttgcatgaaa	tgagggataa	atacacgcat	11280
aatctgttta	acaaaactcc	taggttcata	agaatgcaaa	gtagaaagtg	tcttaaaaga	11340
gactgtgatc	agcaggcaga	cggcctgtgt	cggtttgggt	cctttgactg	cttcagatct	11400
tctaggagga	gctcctctac	tacgagagcc	tggacttcag	cctcacggac	aagaacctca	11460
gataggcaag	gctttgtgat	gggtgctcct	cctaccttca	ttcgaatgag	gttttcctcc	11520

cgccttcacc ctgataggggt gctcctcccg ccttcactgg tgtcatttta ttttgagttt 11580
 aaaaggatga actgggtcttg gtcagcctta ggaagggaag catcactcag gtgaaaatac 11640
 gtggcgaact ttaaataacct agatattctc ctcatgcatt taaattagag tgacttgagc 11700
 tgttagaaaa gttgactggg tgcgggtggct caggctagta atcctaacac tggaaggctg 11760
 aggtgggtgg attgcttgag cccaggaggt cgaggctgca gtgagttgtg atcgcacctt 11820
 tgtattccag cctgggtgac agagtgagac cctgtctcta aggggttttg catggttatc 11880
 cagagcagtc tcttgaattt tggttgtctt tgacgcagct gcatctggag atcctctatg 11940
 tgaggcatgt ttagagaaaag agaaacattg acccctggaa ttgatattct tccaagaaat 12000
 ctttttagtct tgttggtaaa ggcagtcctc agatgggtacc ggtccctttg ttagtggtga 12060
 gcctacagag atgggtgcctg gtgtgcagag cgatgtgggt ggatggaggt gaacagagca 12120
 gtctagggcc cctgggccgc ctgctgggtcc tagcccatca cagcacagtc attcagcaga 12180
 catgctgcac cagctgattg tttgagcttg ggcttggtgc acatgatgga gtttgtctgg 12240
 atctcctgtc ttggtattcc ataggaagat gagttgggtcc agcccatgtc ctaatagatg 12300
 ttgttagcag gaagtctgaa cttgaactat ctacatgaa gaaagttcct tctgctgcct 12360
 ctttttattt ctctttggct tcatgaacag tcatgcaccc ccagtgccat ccttttgagc 12420
 tccccaaggc ctctgccaga cacatcacct ccaggggtga ctccttgacc caccgtgcca 12480
 tcaccactgg ctctgcaccc cttctgagcc attcctggag cttgaggcac tgtgaggtca 12540
 ggggaaaagg gtcagcgtgg caaggagaga tggtcaccca agctgggggtc atagagcagc 12600
 ttatggattt cctggaagtc agtggggaact catttgggaa ccccaaaaat gagtccctgt 12660
 ggggtgtctg aaatgggtggc agaaggggca ggaatgggggt ccggcagttc tcttccctca 12720
 ggtcctggcg gatggatagc aaagtcctgt gtagacaagg tcacccacag gtggtatatg 12780
 ccggctgcta agggccttgg gtggcgggcc ttggcccccg cagtccagag agtgtctgtt 12840
 tactattccc tgcagggtag gcaccttggg cccagaatt cggaaggcac ttttttccta 12900
 cctggactca cttcctgcag tgttgaggc cacctggccc tagtctattt tatagagaca 12960
 gggctctact atgttgtcca ggctgatctt gaattcctgg gcttaagcaa tcctcctgcc 13020
 taggcctccc aaagtgttga gattacaggt gtgagccacc tcgcccagcc ccagccttag 13080
 tctgtagtct aatgctggag ggtagtatgt tgactctgga ataaagatgt cactattctg 13140
 gaactttttc tggggtcact gtcgggaggc gttgtctccc ttgagcactt gctgcagccc 13200
 caggctgttg acctgcacag ccgtcctcag aggggtgggga agctgttctg tcttgggggc 13260
 gttgagtcca ctaggtaggc cctgtgagga tcccagctgt cacctagtgg gtacctgcag 13320
 aggtttttgtt ttgttctttt aattgaggcg gggctcttgct ctgttgccca ggcaggagtg 13380
 aagtggcgca atcctagggtc actgcatcct tgaactcctg ggcccaagtg gtcctcagcc 13440

tcccaaagtg	ctaggggttac	ggattagagc	caccggccgc	ggcccccttt	ggcttgtgtt	13500
tttaagtctt	tttttcctcc	atgtttcttg	ctgttaattt	tttataaaag	tttgtatctg	13560
gcatggattg	aagcaggaga	cactggctca	gtgcccagtg	gggagagccc	tgtgtgggta	13620
cctcccgggc	ggtgtgggtg	catgagttgg	ctgctgcggg	tggcaccggg	aggccctgtg	13680
tgggtgcctc	ccgggcagtg	tgggtgcatg	agttggctgc	tgcgggtggc	actgggaggc	13740
atggcagggg	ccatgaggcc	acaggtggca	gctccctgtg	cttgacctgc	atcgggtctg	13800
gagtggcggg	cactcgggaa	ggaagcgcca	ggacgacaca	ttggaaggag	gagctgcagg	13860
gcagggtccag	ccagtgcctg	ggagctcggg	agggcacctg	ggtggacgca	gcctgcagga	13920
gcgcctgctc	tcctgtatga	ggaacagacc	tgaatgccgt	ttcagacact	gactcattcc	13980
tgccctcctg	tgactttgtt	ggttttcttg	gttttagtgc	agattgtagc	tggcagctca	14040
caacagagcc	cttcctgggtg	caggctgaca	ggcaggaggg	attctcctct	acagaaatgg	14100
acattactat	atctggaaac	agtttttttt	tttttcttct	ttttgagatg	gagtcttgct	14160
gtgtcgccca	ggctggagtg	cagtggcgtg	atctacgctc	actgcaacct	ctgccccctg	14220
ggttcaagcg	attctcctgc	ctcagcctcc	cgaatagccg	gaattacagg	tgcacgtcac	14280
cacgccaac	taatttttgt	gttttttagtg	gagatggggg	ttcgtcacgt	tggccaggct	14340
ggtctcgaac	tcctgactga	cctaaggtaa	tccgcctgcc	ttagcctccc	aaagtgttgg	14400
gattacaggc	ataagccact	gtgcctggcc	tagattttgtt	ttaattgaaa	aagctgaagc	14460
aaataaaaaat	tgtgctgggg	aaaactctga	gtgctcagag	catgggctgc	gtttctcact	14520
ggcctcaccg	caccttcctc	tgcactcaga	ctctcctaaa	gatgctggca	tttactgaac	14580
taaaggtagc	atttcgcaga	gcagggccgt	gggccatgaa	ccacagtcac	ctggccattt	14640
agacagctct	gccctgcccc	cggcgacttg	ggtacccaac	catgtcattc	ccacccttag	14700
ctttgtgggg	caacgttcct	ctaggccggg	ttctcttccc	tgaagcacc	caggcatgta	14760
gaagatgttc	tcagcagagg	ggagacacca	cagtgtagct	gatttctgtc	tctcttcact	14820
ttcaccatat	ttacttactt	ttggtagaag	tggggttgaa	aatttatgtg	aagttttctt	14880
tactactggg	gacagcgggg	cagattactg	ccacacataa	tcacctctgc	tccagcctag	14940
gtgctacctg	ctggaggcga	tgggtggcttt	cctgagagca	cctcctgggc	tcgggctggg	15000
ggggagcttg	attttacctt	tttatgtgga	aataacttca	aatttacaaa	atattcacia	15060
aaataaagag	agcaggggact	acccaggtgc	tctacataga	gttgccctgtt	gctggggctg	15120
cgtcctcttt	tgcacaccct	gccccgtcgg	agaggatccc	agatgtggcc	cttgcccatg	15180
cacgtcaggc	atgtgaaaag	agcaggtgtc	ccctggcgat	gggcagcagg	ctttgtgggtg	15240
ccccagcctt	ggtgcaggag	tcccctgagg	tgagattcag	ggctctggcc	ggctgggtctc	15300
atgtgtttcc	tgagcgatgg	atgtgggcac	ggcgtcctcc	acctcgcacc	acctggcatg	15360

gggtctgcgt	gtgtgtcatg	ccaggggttct	gggcctgcac	cgggggctcc	tgccctgcct	15420
gtgcactggt	tgtcgggggc	ttgtgcccgt	gggtgagtgt	atcttttcat	ttttagtgat	15480
gtctttaaaa	tattcttttt	tagtttgga	atctttggag	agagccccct	gttctttcac	15540
tcaacttttag	aagttatacg	cacctggcca	gcctcccctc	tcccattcct	tccattccat	15600
tgttttgaag	cgagtcctgg	ccatggaatt	ccgttcgtgc	acccgtcagt	gtgtaaaggt	15660
tctgaaaagc	agcctgaatt	ttctattgtc	tgatgaaccg	aagtttggtc	atttcatcgt	15720
actcctcaaa	tttatcaccc	tttccccatg	tggcatgtgc	atgctgggtc	tgcttgaaga	15780
aactcgtccc	cactctgcag	ccgtaaagct	gctccctgta	tttcctctag	gagtcctttac	15840
aaattttttt	ttttttttta	aatggagtct	cactctgtca	cccaggctgg	agtgcaatgg	15900
cacagtcttg	gctcactgca	gcctccgcct	cctgggctca	agcgattctg	cctcagcttc	15960
ccgagtaact	ggcattacag	gcgcctgcca	ccacacctgg	ctaatttttg	tatttttagt	16020
agagatgggg	tttcgccgtg	ttggccaggc	tgggtctcgaa	ctcttgacct	caagtgatcg	16080
cccacctcgg	cctctcaaag	tgctgggatt	ataggcatga	gccattaatg	gctcctggcc	16140
aaaaattttt	ttaataaaaa	taaaaggctg	ggtgtggtgg	ctcacgcctg	taatcccagc	16200
actttgagag	gccaaggggg	gcggatcacg	aggtcaagag	attgagatca	tcctggccaa	16260
catggtgaaa	ccccgtctct	actaaaaata	caaaaattag	ctgggtcgtg	tggtgggcac	16320
ctgtagtccc	agctactagg	gaggctgaga	ggcaggagaa	tcgcttgaag	ccaggaggca	16380
gaggttgcag	tgagctgaga	tcatgccact	gcactccagc	ctagcgacag	agcgagactc	16440
catctcaaaa	aaacaaaaaa	tagagatggc	gtctcactgt	gttgcctagg	ttggtcttga	16500
actcctgagc	tcaaattgatc	cactcacatc	agcctccgaa	agtgcctggg	tcataggggt	16560
gagccatggt	gcccggccac	ctcgaggagt	cttgagccct	cctggaattg	ctgtttggtg	16620
tgaggttggg	tcaggcctgg	tgagtcattc	gggttggtgag	ccagcctgct	tggcctcctg	16680
gtctcaggcc	agctgtgacc	tggaccttgg	gccaccagc	agctttgtgc	tgagcctcag	16740
tctgcggttg	ggagggcacc	ccttgctctc	agaatggtcc	tggcgtgtga	ggtccttggt	16800
gcgtccgtgt	acagtgggcg	gcaccttgtc	tccagaacgg	tcctggcgtg	tgaggctcct	16860
ggtgcgtctg	tgtacatggg	cggcaccttg	tctccagaat	ggtcctggcg	tgtgaggtcc	16920
ttggttcgtc	tgtgtacagt	gggcagtagc	tttaaagttc	ctgcttcacc	ttgtcagact	16980
tctggttaga	gtcctcccga	cattagtgtt	cctttaggaa	gcttggactc	tttacaccgt	17040
taattacact	gttaatttca	cttaaacadg	gaggtggccc	agtttgctgc	ccatgtaaga	17100
ctactttttac	ttttttgaga	caaggctctg	ctctgtcgcc	caggctggag	ttcagtgcca	17160
tgatctcagt	tactgcagc	ctccacctcc	agagttcaag	tgagtctcat	gcctcagcct	17220
cctgagtagc	tgggattata	ggcatgtgcc	acgactcctg	gctctctcta	ggacttcttt	17280

aatgtcataa	tagaagctga	tggttttctc	cataaaagtc	tattagattt	atttctaggc	17340
acttgттаат	atagtacttc	ttaaaattag	cattttttgtt	tttagctatg	tgttgtagat	17400
ggatacaaca	tgttgtaaga	ttctttttgga	tcgtctttgt	ggacagtatc	tgtgaagaat	17460
gactcctgcc	tcagtctccc	acttgaaccc	aggaggtgga	gcttgcagtg	agccgagatc	17520
gcgccattgc	actccagcct	gggtgacaga	gcaagactgt	ctcaaaaaaa	aaaaaaaaaa	17580
aaggagcaga	ctgtgctcct	tgccatagga	ttttatccat	gttcttttat	ttctgcctct	17640
caaatactgt	ttaatcttaa	cgtattttta	aattttgtgt	gatttttctg	tttttgagag	17700
gaatttttga	tatgtttaata	ctatagatct	tagtataattt	taaaaactga	tgatgtttga	17760
tacatgccaa	atcatgtgtg	gaccataaat	gtaaatagca	atgcccacaa	accaccccat	17820
agcctaggaa	ttgggagact	cagcaccgcc	ggctgaattc	gcctgtgcgg	tgtgtcccat	17880
tccatccctg	ttctgccgcc	cgtgggggaa	ttttgtcagc	ttcactgaga	cacttagtac	17940
cctcccctta	cctgtggttt	tgctttccat	ggtgtcagtt	gcccattggc	agccatagtc	18000
tgaaagtatt	aaatgggaaa	ttccagcagt	aagcagttca	taagttttca	gttgtgcacc	18060
attctgagca	gcgtgatcac	agtctccctg	tcgtgtcctg	cctgccggga	gcgtgaatct	18120
aacctttgtc	cagtgcaccc	acgccgtgca	gctgcctgca	ccttaacccg	cttagtagca	18180
ggtaccaaca	tccaggctgg	cgctcgggtc	tgttccaagt	ttcaggcagc	cgctggaggt	18240
cttggaacca	tcccccttgg	ttaagggggg	actcttgtca	ttcacacacc	atgaaattta	18300
cccatgtgaa	gtgtacagtt	gagtggattt	tagttcacag	agttgtgcat	ttatcaccac	18360
caccaatttt	agaaattttt	atcacctcca	aaggaaaccc	tgtgcctctt	agccattacc	18420
ccaacttccc	accttctcat	cctccccccag	ccctggcagc	cacccatctg	catcctgtct	18480
ccgtgtattt	gcctgttgtg	ggcatttcat	agaaatgtgt	ttttttgttc	ttttatgaaa	18540
tgtgtttttt	ttgttctttc	actgagcatg	tttccaaggt	ccatccatcg	tgccccctgg	18600
gtcgggtgtgc	cattcctttt	cacggctgcg	taatgttcat	tgtatggata	ggccacgttt	18660
aaaaaaatct	gttcatcttt	tgatggatgt	ttgttttttt	cctacttttg	agctattatg	18720
aacaatgctg	ctctgaacat	ttgtgtaaca	gttttttaggt	agtgtgtttt	tatttctctt	18780
gggtatatatt	taggagtgga	gttgctgggt	catgtggtaa	tgctgttttag	tattttgagg	18840
aagtgccaga	ttgtttttcca	aggcagctgt	acccctgtat	attctcaggc	gctgtccttt	18900
tgatgatacc	cattccatgg	tgcaggggca	gcgatcgctc	gggggttttg	ctttgcagtt	18960
ccacagtggc	tgaagatggt	gcctggccat	ttgcatattt	tcttttgaga	aatgtctggt	19020
cagatctttg	ccatgttttc	attggtttagc	tgacttctta	ctgttggtcat	aagagttctc	19080
tgttctgggt	actagactcc	ctattagata	catgacttgc	aaatactttc	tcttcattct	19140
gtgggtcatc	ttattgtttt	attcagtggt	ctttgaagca	caaacattta	caagttggat	19200

ggagtccagt	ttatctgctt	tctcttcggt	tgcttgctgct	tttggtgtca	catgtaagaa	19260
gcctttgcct	aaccaagggt	cttttaatag	gagttttatt	tattttttata	tgtgcacata	19320
tatatatatatt	ttaagagctg	gggtctcact	gtgttgccca	ggctgttctg	aaactcctgg	19380
gctcaagcaa	ttttctcgcc	ttggcatccc	aaagtgctgg	gattacaagt	gtgagtcacc	19440
gcgtctggcc	actgttttcc	tatatatgat	aatttctacc	tttacctgat	agcagtctag	19500
aagcatgtgg	ggcttgcttg	cttgctcgct	cgctttcttt	ctttctcttt	ctttcttttt	19560
taaccaaatt	aatttttttaa	actcttgata	tccattgctt	gtgctgtggg	caaaggacgt	19620
gacctgcatg	gtacgcattc	attctttgat	gcttgttgag	actttatggc	ccaatagggt	19680
atacatTTTT	tataaatagt	ccatctgtgc	ttgtgaatat	gtgtgttctt	tctttctttg	19740
atgcaatat	ctctatgagc	ttgttgggtc	aaatgtataa	gtcatgctta	cattttttaat	19800
atccttaatt	tattttttgcg	tgaattgggt	tcattaggct	tgtagggtta	acagtcttca	19860
ttgtagctcc	tggagggttt	gcatttgata	ttttcaagat	gtcccgttag	gcagtgtgta	19920
cgcatttgga	atttggtcac	ttccttcgat	catgggtccag	tagccctcat	agccctcttc	19980
actccagtca	tggtgtttct	ctggcctcag	ttcagctgat	cggtcagcag	taccactttt	20040
ctgctcctta	atgtacacct	gaacattttc	cttatgtcca	gcttcagcct	tgctgtggta	20100
ttttaaaact	ggaattacat	atatatatat	atatatatatt	tttttgga	cgaagtctca	20160
ttccgtcacc	caggctggag	tgtagtggcg	cgatctcggc	tcactacaac	ctccgcctcc	20220
cgtgttcaag	cgattctcct	gcctcaagcc	ttccaagtag	ctgggattac	aggcacctgt	20280
caccacgcct	ggctaatttt	tgtattttta	gtagagacga	ggtttcacca	tcttggttag	20340
gctggtcttg	aactcctgac	ctcatgatcc	acctgcttcg	gcctcccaaa	gtgctgggat	20400
tacaggcggtg	agccactgca	cctggcctaa	aactggaatt	ttaaattttt	aaccagattg	20460
acaacctctt	cattagcaac	tttagactgt	tgacttatgt	tggattactg	atctattttg	20520
gatttgtttc	catgatcctt	tttttccctt	tttctttatt	tttcctctga	gactgagtct	20580
tgctatgttg	cccaggctgg	agtgcagtgg	tgtgatctcg	gctcactgta	atgtcttcct	20640
cccgggttca	agcagttctc	ctgcctcagt	ctcctgaata	gctaggatta	cgggcgtgtg	20700
ccatcatgcc	tggctaattt	ttttgtattt	ttagtagaga	tgggggtttca	ccatgttggc	20760
caggctggtc	ttgaactcct	gacctcatga	tctgcctgcc	tcggcttcca	aaagtgctgg	20820
gattacaggc	ggaagccact	gcgcccggct	cactttttct	tctgatgtat	tgactcagat	20880
ggttttgctg	ttttggagtc	gggggcattt	ttcctcctct	actcatttgg	aagggtgcaca	20940
ttccatttaa	ctctctgtat	tcaagttttt	tcagtatttt	actgtgggtca	cgtacacata	21000
acataaaaatg	taccatctta	atcattgtaa	gtgtatagct	caggggcatg	aagcactttg	21060
atattgctgt	tgctgtgcag	ccatcaccac	catctatctt	cagagccttt	ttgtcttccc	21120

agtctgaaac	tctgtcccag	ttaaacagga	actccccatt	cctccctccc	tggctgcggg	21180
cgttcacctg	tctactccct	gtctccatga	atttgactcc	tcttcagtcc	ctcacataag	21240
cgcaatcagg	cagtatttgt	ctttttgtgt	ctggcttatt	gcactgagca	catttttcaag	21300
gcttttccat	gtcattgcat	gcattggaac	ttccttgcct	tttaaggctg	aataaccctc	21360
tgtatgtgga	gataagcttt	tgcttaccca	ttccttcctc	catgggctct	tgggttgctt	21420
ccatgtttta	gccatcgtga	ataatgctgt	gaagaacatg	ggtgataagt	atctgagcct	21480
gttttcagtt	cttttgggtg	tatacccgga	agtgggattg	ccaggtcaca	cggcattttct	21540
gtctttaagt	gtttgaggag	ccacaacact	gttttccaca	gcggctacac	cttttttacgt	21600
tcttgccggc	ggtgcgcaag	ggttcagggg	ttccacatcc	ttaccaacat	tttttttttg	21660
ttttttgtta	gtattttttg	atagccattc	taatagggtg	gaagggtgaca	tcatagtttg	21720
gatttgcatt	tctgtgttta	ttgagcattt	gcacttttga	agaattgtct	atttgagtat	21780
tttaaccagt	tttgaattgg	attatttttt	gtcaagtttt	aggagtttct	ggatattaat	21840
cccatttcag	atacgtgatt	tgaaaatatt	tcctcccatt	cactctgctg	tgtgtctttt	21900
aatgcacaca	ttttaacatt	tttcatgaag	cccagtttat	tttttctttt	tgtttcctgt	21960
gcctttgggtg	tcttattcaa	gaaatcattg	ccaaatccag	tgtcatgaag	cttttccctt	22020
ctgttttaag	agatttatag	tttttaggtct	tatgtctagg	tcttcgattt	tgagttaatt	22080
tttgtgtgtg	aaatgaggca	aggtccaact	tcacattctt	ttgcatggcc	atatccactt	22140
gtcctggtgc	catctgctga	aaagactgtc	ttttcccctc	tgaatggctc	tggcaccctt	22200
gtcaacagtc	tctgctggat	acgggagggg	tcgcttcagg	gttctctgct	attcccatgg	22260
tctatatgtg	tgtcctttcg	tcagcaccgc	actgttttga	tgaccgtagc	tttgtggtaa	22320
gttctgaagt	caggaagcgt	gactttttaca	gctttttcaa	ggtgatttgg	ctgttgggtg	22380
ttccttgagt	ttccatatga	attttagggg	ggatttttct	atttctgcaa	aagacatcat	22440
tggtattttg	ataggaattg	cattggatct	gtagatcact	ttcgggtggta	ttgacatctt	22500
aacaatattg	gtcttcctat	ccatacgttc	cctagtgcac	ttagtgggtc	ccatggtaaa	22560
gtattttatgt	gtcagcatgt	agactgggct	taacaaaaat	ccaaggaata	cttgatgttg	22620
gaactctggg	gacctccctc	atttccgggg	ttctttttccc	tgggtgtgggt	tctttttttt	22680
ttttttttga	gacaaagtct	cgctcttggt	cctcaggctg	gagtgcacatg	gtgcatctt	22740
ggctcattgc	aacgtccgct	tcccgggctc	aagcgattct	cccacctcag	cctcccaggt	22800
agctgggatt	gtaggcacgc	accaccacac	ccggctaatt	tttgtatttt	tagtagcgac	22860
ggggtttcac	catgttggcc	aggctgggtc	cgaactgctg	acctcagggtg	atccgcccgc	22920
ctcggcctcc	caaagtgctg	ggattatagg	cgtgagctac	cgcgcccggc	tcctagtgtg	22980
ggttctgact	tgtctgcagt	catccctgag	gtgggcattg	cggctgctgc	ccccaggcct	23040

ccttcatctt	gcccgtattg	ccggcatctc	agtccacca	gccctcaggg	actgctttct	23100
gtggttgtca	tatgtatgtc	tttgcttagt	tttcttatct	gaggtctggt	tccccttaac	23160
tcttctgtcc	gagttgagtg	ggggtgctca	tcgcctcagc	tgggaagctg	ctcttccatg	23220
ctcttcctca	ctctaatttt	ggggttttca	gcctcttgct	cacatgagct	ccctcagcct	23280
ccaagtagt	tggggtcaca	gacgcacacc	tctgcatctg	gctcttaagt	ttttaagca	23340
ttttcatcaa	agcctttttg	ctcacaactt	gtgggctggt	gacattttcca	gaagccctga	23400
ctctggagaa	gatggtggct	tgttgtgatg	aggtcagaat	tttttctcac	atctcatggg	23460
tcagaagaaa	taaaaacaaa	cagaatcttt	ctttctttct	tttttaagag	atagggctgg	23520
ctgggtgtgg	tggctcatgt	ctgtaatccc	agtacttcgg	gaggctgagg	ccaggagttc	23580
gagaccagcc	tggccaacat	agcaaaaccc	catcttcact	gaaaacataa	aaattagctg	23640
ggtgtgggcc	gggagcgggtg	gcttatgcct	gtcatcccag	cacttgggac	gccgaggtgg	23700
acagaaacct	gaggtcagga	gttttgagac	cagcctggcc	agcacgggtga	aaccttgtct	23760
ctgctaaaaa	tacaaaaatt	agctgggtgt	ggcaacgggc	gcctgtagtc	ccagctacta	23820
gggaggctga	ggcaggagaa	tcgcttgaac	ccgggaggcg	gaggttgagc	tgagctgaaa	23880
ttgcgtcact	gcactccagc	ctgggcaata	gagtgagact	ccatctcaaa	aaaaaaaaaa	23940
aaaaagactc	actgaaccag	aagagttaag	tgttgtgcta	ttgactgcac	atggtttgcc	24000
tcagtcccaa	gtggtaaaaa	cagactgctc	ccaaaacagc	aggccatgag	gtgaccaggt	24060
ccagcccagc	ccagcctgca	gaggaggagc	cctgtggggc	tgtgatcaga	gcccagcgtc	24120
ccagtgcacc	tgccgatgcc	ctctccagcc	tctggcttcc	ttccgttggc	tgtccttgga	24180
gggctgtgct	tcgtgggcag	gcgtttgggtg	tcctgagcat	gcagactgcc	cttgaggtgg	24240
tgggcaggat	tgtttttagcc	taaaggggtt	ccacatctgc	acgggagctc	ttgctgatgg	24300
gtcactctgg	caaagttccc	ccggcgggtt	ccctgtgaga	aggcacctca	cggtcagttc	24360
tgggtctggt	gagtgttctc	ctgcaggctc	cacgcttgag	ccaggttctg	ccggaagagg	24420
cggcgttggg	ttcccttggt	ggatgagagg	cgccgagggt	cttggtggct	ccatgtgccc	24480
ccaagggcag	cactttgcat	ttcttgggag	acttgaagat	gactgatgtc	attttagtga	24540
gcacggtaaa	acagcacctg	ttagcggctt	ataccttgag	ctctgtctct	cttcttgag	24600
gtgacatggg	gaccagtgcc	cacctctccc	atggcgtctg	agtggacagt	gtctaagggt	24660
ttgggcccgc	tgacacagat	gtggttgaag	tcagcttggtg	tgcagatggg	gtcacataga	24720
aagaatgaaa	actgttgaag	gctttgctca	gcatgttttc	tttttgata	ttcaatctga	24780
taaatgttgt	agtgaatgag	aagtggatga	aatctttttt	ttttaatttt	atttttgagt	24840
cagaatctcg	ctctgtcacc	caggctggag	tgcagtgggtg	cgatttcggc	tcacaccaac	24900
ctccacctcc	caggtttaag	cgattctcct	gcctcagcct	ccaagtagc	cgggactaca	24960

cgcatgtacc tccacattca gctaataattt ttgtatTTTT ggtagagacg gggtttcacc 25020
 atgttggcca ggctggctc aaactcctgg cctcaaggga tccccccacc tcggcctccc 25080
 aaagtgctgg gattacaggc aggagccact gtgccaacc taaaaagtgg atgaaatctt 25140
 caaatggttg gagatcgtga attttgTTTT agcagcacta gggatagtag aaatacagggt 25200
 gtcatggctt agggagaaaa agtatgaaca catggTTTTg ttagcagtag tactgtcatt 25260
 tctagtttat gtgatgaaac atcttaattgt actgcttcaa agatacactg agtgtttTgtg 25320
 ttttaggtta aataggattg cgaatcagggt ggccattcag cggaagaagc agtttTgtgga 25380
 gcgagcccac agctactggc tgctcaagcg gctgtccagg aacggggccc ccctgctgcg 25440
 gcggctgcag tccagcctgc agtctcagcg aagctcacag caggtagctg gctcgtggcc 25500
 cccacagcct ttctgctct ctcaccccca ccctccaggg agagggcacc tggttcaggc 25560
 tggtcacgggt gctgccccct gggctcaggc tccccgttg gaagcctccg tctcgtccag 25620
 ccctggagct gtaggaggag tcgggtgcac tgggctctgg atctgcgact cggggggTga 25680
 ttctcatcca tgcctgtcac tgggtgtttc ttccgttttg agttgcgctc ccccgccctg 25740
 ccctgcccgc acacacgctg atgggcagtg gccagctctg gggactgggg atgctcaggg 25800
 atgtgttatg ctgtggctgc atgagcgtcc gggctgggga cagcaggagt gggacccgggt 25860
 gagcacacc gggccatgtg ggtgggggtc tgtgtgctgg gcctggaggg tgggtttTgg 25920
 aaatatgtga ggagtggttt atttcgaggg ccttgctctg aaggttgggt tagctgggac 25980
 tgtggtgggg tggggagggc tcgagagagc tgagagagtg ggagggaagg agcaggactt 26040
 ttctttttta aaagaacaaa ccaaactct tacattttga tagtaggttt gagtttatag 26100
 actcatctga tactgtgaaa cagttctgaa aaaatgggtt tacagagtga agtagcttct 26160
 gtgagatgca caaaagtaat tgaaaacaca agcagctctg cacacggccc aggacctgcg 26220
 gccagcggc tctgcccacg ctgcctagtg aggagccgct ggtctttTgg aagagctgtc 26280
 tgctgctggc gtcgcgtggc ctgagtgcct cttggaattt cctcatgggt cttttcactg 26340
 ccatcctggg cacgaagcga aaccatcctg ctctgctgtg aggtggtgac ttgctcacag 26400
 gaagcttgag tctacagggt tcacggtttg gtccccgtgc ggtgggtgcg tggccgggtgc 26460
 tgtctctgcc tcgcagactc tgctctgca gcgtgagccc tgttcagtca cctgccgtgg 26520
 ctgccttggt ctccaggga ctacagagtc tgggtgctctg gtctggcagg tgggtggttTg 26580
 tgggtctgat ttgggtgtgt ggggtccaaa gttttctgac ggccttggtg cgtcttggt 26640
 gtggaatatt ccaggctctc cgtgccgtcc acctgggttg cccactgtg tgtctgtcca 26700
 tggaacttgg ggctccccgg gcaggctctgt gagccagccc ctctccaagt gagatcctca 26760
 cgtgtccaca gagttgaggc ttcagtgaat gtcttttTggc tgttgaatga gttaaccgca 26820
 tcttccgttc catgatgtga cttggctttc catctctcct ctgtcctcac ctcgtagcct 26880

ccagggctcc tggaaaggac agcacggttt ggggggctct taagaggcca gagtcaggaa 26940
 gtggctgagc ggcgggaagc tctgggggct ggaggcattg gtctcggccc ccgcatgtt 27000
 ggggaaagga agagagtgga ggcattggc tcggcccccg cgatgttggg gaaaggaaga 27060
 gagtggaggc attggtctcg gccccgcga tggtggggaa aggaagagag tggaggcatt 27120
 ggtctcggcc cccgcatgtt tggggaaagg aagagagtgg aggcattggt ctcgccccct 27180
 gcgatgttgg ggaatggaag agagtggagg cattggtctc ggcccccgcg atgttgggga 27240
 aaggaagaga gtggaggcat tggctcggc ccccgcatg ttggggaaag gaagagagtg 27300
 gaggcattgg tctcggcccc cgcatgttg gggaaaggaa gagagtggag gcattggtct 27360
 cggcccccg cgatgttggg aatggaagag agtggaggca ttggtctcgg ccccgcat 27420
 gttggggaaa ggaagagagt ggaggcattg gtctcggccc ccgcatgtt ggggaaagga 27480
 agagagtgga ggcattggc tcggcccccg cgatgttggg gaaaggaaga gagtggaggc 27540
 attggtctcc gccctgcga tggtggggaa tggaagagag gcttttttg gaggaagttg 27600
 aggttcaga tcttgggtct tatctcggg ctggacagaa ctccaggtaa gtcacagagg 27660
 agctgttgaa gggcagattc catggaagt gggcatggat gtggccaaac gggacgtcca 27720
 gcggtgggaa tcatcatttc tactctcata agctgggaca cgagggccag aatccctgca 27780
 gtagcaggac acgtgggcca gtgggcagg gccacgggga cccaggcagc tgggggtggc 27840
 ctggggcgtg gtggagccag gcaagggaag ggggagggtt gcaaagaaga ggtgaccgc 27900
 tggaacctgc cagctgggcc ccctctgctt catcgctgcc tgtaggcct cacgtcctgt 27960
 tgtgtgcaga atacgtttgt aaggagagaa catccccaca gcccggccag agcagctctt 28020
 ctggcagagg gggcagcttc tcgaagtcac ctgggaagga gcagaaggag ctgcctggac 28080
 ttgagccgca gcagagcagt cggcgtggag agcctggaaa cctgcctctg tgtgaacggg 28140
 ctattgggag agactgtgct tctgccttct cctggccagg gatggccacg gcttggtgtg 28200
 gtttcatggt tttgaagctt tggaagtacc tgatgcattt gattcctctg tgtctctgtg 28260
 tacctgcccc agcctgacct tgtggtggg gcagggtgat gaactggacg gggctctctg 28320
 agaggccagc aggcgccgag ggtgctgcag cgatggcggg tctgtgctgg ctgcgcatat 28380
 aggccctggt ttttgctctg gtgtgtgtgg ccgttgaagg ccgcaggag aacgggggcg 28440
 gcagcttagc tttataaagg gtgagcagcg atttgtgtgt atgctggtgg cgtcccactg 28500
 tctcctgctc ccgtggcca cccttcctgt cctgggagcg catcctgcag ctttcccagt 28560
 gtctgtcctg tggcatgctg ggcgctgctg cccacatcgt ccttgctgtg ctgcgtgtgt 28620
 ggcctctgcc agtcccacgc acctgctggc ggcattgggc ccagtggcag gagctgggtg 28680
 ctgttgccag gaggcctggg gtcggccact ctgtaacact gtgggccctc accccaggag 28740
 ctggcctgtc ctcccagcag aggcgatttt aacctgaccg tcgcctagcc ctgccctgtg 28800

tctcccccat	gtgtctggaa	tggacgcgga	ggcgctcagc	ttcttcccgg	cagtggggat	28860
gcctgagtgc	ttcttcactt	ctgtcctttc	tcgtggagcc	tcccactggt	gattttttaca	28920
tgggaaatat	cgagcccttg	agagccgagt	aaggaacctt	gatccttctg	agtggatgag	28980
gctgcatgtg	tcattgggacc	ttggggcccag	gtgctgtcaa	cggggcccct	gccactgccg	29040
atggcgtttg	ttaaaccagg	ttgggggagca	gactagtaag	tttgcttagc	agccagtgcg	29100
tgataaaggg	cttcctgggtg	atgaggagat	ttcttccgct	tcgtaacaga	ttcccaggcc	29160
aaatgtcctt	ttggccgagg	ttttcagaag	tgctgattgc	aagctgaggg	acaggcaata	29220
ctcagatgct	ccggggagct	tggttttttg	agcccacatt	ttcctgcctt	tgcttgagtt	29280
gatgtggatc	ttgttactct	tagccccgat	ttcctgcctg	agggtgttca	gcagcattga	29340
ggaaaaccca	gagatcctca	cttcagggga	catcagcgca	gcaacgctct	gcctcgccat	29400
ggagaggcct	ggccttgggg	gtgggggtgc	ggggctcggc	ctccctcctt	gagcctcctt	29460
ttctgagcag	tttgaggtcc	ggggtccttg	catcttgtag	tccagcagtc	tttgccttca	29520
caccctcggt	cctttgaaaa	gctatgacag	taccctgagg	tgggcagggtg	ggaacccggg	29580
gacagaggcg	ttggggggaga	ggaggggtctc	agggtgtga	agcatgtgcc	agaaactgat	29640
ggccacggct	gtcctgaggc	tggcagggcg	cgctgttcct	gcagctgtgg	aggatgacag	29700
gtgttggggg	cccactcagg	ccaactccct	cgttgacttc	ctctcccctt	tgaggtaggg	29760
tcagggtgcc	ccggcactgt	gacgagcagg	cctcatgcgg	ccaggcctgg	gggacatggt	29820
gaggacaggg	ctgagggcag	accaggtgg	cctactctcc	tgcgaccatc	tgtgtcatgg	29880
cctttgcagg	gctctgattt	tcttctgatc	accctggtgt	ggatgttggc	tgggttgagg	29940
ctggcggcct	gtgctgtctg	gcgcactctt	ctgtccactg	agaagggagc	agggtgctatc	30000
aatggaaatg	ggaccctggc	tgtggggaaa	cggtgtgcgg	gatggaggcg	ctgtgcaggg	30060
cagggcagca	ggtcggctgt	gtgccaggcc	cccgtcatgc	tctgtccaag	ccaggatggg	30120
cccagcatgt	ggttttctgt	gcacctgggt	tctttccagc	cgctccctcc	attcccagct	30180
gtcctggagg	tgggtcctca	agggtccccc	attctgtcct	ggccccgcct	ttgtgggcct	30240
ttctggttct	gtgaaggccc	ccccgcgcgc	cagcttgcg	aggcctctgt	cctctctgga	30300
gggtggaacc	tgtgcaccgg	ggccagcccc	ctgctcacct	tggcctggcg	gtcccctctg	30360
tcccctccat	cccctctgtc	tcctctgtcc	cctctgtccc	gggtgctgtcc	tggggctctc	30420
tagtgagggt	ggtctgttgt	cctccccagg	atgggcctgc	aggaccctg	acgtggatgc	30480
ctcttgggggt	ctcacctgca	tggatatctga	ggccttttgc	agaaggagcc	tctgggtttg	30540
actgacggaa	gtgttcactg	ttctcagaga	gaaaatgatg	aggagatgaa	ggctgccaaa	30600
gagaagctga	agtactggca	gcggctgcgg	cacgacctgg	agcgcgctcg	cctgctgatc	30660
gagctgctgc	gcaagcggga	gaagctcaag	cgtgagcagg	tggggctggg	cctgtcccgc	30720

gtggagtgca ccactggcca cggggcgtgg gtgggacgct ggcagctcct gcagaggggtg 30780
 ggctgcaccc ggtgggtcct ggcccacagc agtgccttgc tcagcccagg ctgggtgcct 30840
 ggctgccatg ggttgcagag ccctaagtgg aggctgcgtt gtcctatgcc tggacggaag 30900
 cacaggagcc tgtgggggct gcttgtggct gtggggcgcg tgggtgggct gagacgcatg 30960
 tgcggctccc taaagctgag gctctgctgg gcccccccag gtgaagggtg agcaggctgc 31020
 catggagctg cggctgaccc cgctgacggc gctgctgcgc tcagtgtctg accagctgca 31080
 agacaaggac cccgccagga tatttgcgca gcccgtagt ctgaaggagg tgggtgtttg 31140
 cggcgctcgt ctgtcctgcc gcatgtgtga ctgtggattt gtgctgccgt ttgaccggga 31200
 aacgtgtgat tggctagagt acatcccttg tggtcttaac tctcgttttt atttaacgta 31260
 tttgttttta aattataaaa gtgatacatt tcattgtaaa atactgaaag tgcagtgggc 31320
 aagaggccct gtgcagacag taccctgggc tgtggcagca tccgcccggc accctccagc 31380
 gtgcttttagc tgaaggagct ggggtcggga aggtccctgc agtacttggg tcttatgtct 31440
 gtaccgtctg cgtgtctgtg tgtggctctc agataggatg ggcttgtggg ttatatggaa 31500
 taataaatat agagtcaact aatatcctgt tttagtttgt aaatgattct ttttcctccc 31560
 taaaattagg taccagatta tttggatcac attaaacatc ccatggactt tgccacaatg 31620
 aggaaacggt tagaagctca agggataaaa aacctccatg agtttgagga ggattttgat 31680
 ctattatag ataactgcat gaagtacaat gccagggaca ccgtgttcta tagagccgcg 31740
 gtgaggctgc gcgatcaggg aggtgttggt ctgaggcagg cccggcgcgga ggtggacagc 31800
 atcggcttgg aagaggcctc ggggatgcac ctgcctgagc ggcttgcctg ggcaccgcgg 31860
 cggcctttct cctgggaaga cggtaagaac tggcgccgtg tccggaggag gtggaagacg 31920
 ctctttctag cagactctgt cctgctctgt gttgccagct ccctatcagc atggtccggg 31980
 agccactagc atgcaaccgc atacatcatt tacattttct agtagcctca ctgaaaaatg 32040
 gaaaagaagc aggtggaatt tactttaata atatgtctta ttttactcag tgtatggaag 32100
 agttgatttc aacatgggat cggtaaacgt tactgaggat atataagctt ttcaccttct 32160
 ttgccatctg atgtgtatth tacgctcagc acctgtcact tcagactggc tgcgttttgg 32220
 gtgcccacca agaggactgt gtggctgtgg ctccctatgg ggtggccggc ctcgacgatg 32280
 cagaccaga aagcagctta gttgggggtt ttcattccaga gtgggcctcg gctaacttct 32340
 gaggggtggcg tgaagccgga ggatctctgg ctttgcctcc agccctttta cttcgtgggt 32400
 tgcgcatgtt gtagaaagg ctctgggagt gtgtccacgg cccacctgga ctccagcctg 32460
 ggcttgtgct tcttggttgc ccaggcacgt gttgcagcca tgtgtagtca cctgcaccct 32520
 tctctagact ggtcctcgga aaccggggct gtcccctgca ccctgcttgt tagtcgcacc 32580
 tctgggattg tctgtcgaag ggtgtgtcct gccacatgg gttgtcaaag cacatagagc 32640

tccagcacct cagctccagc tctgagccac aaactgagca gaaatcacgg tgtgttttggc 32700
 ttatattgtct aaaggtccca tatccctgat ttcttggcct accctttctc ttgatgaatc 32760
 actgacccgt gctaggtcgc ttcttctggg atgaggcctg gccctcttcc tctcttccac 32820
 gttcccggca tcgcggggtc cacgttcccg ccatcgtggg gtccgcttcc cgccatcatg 32880
 ggattcaacc aagctggggc atcgttgctt cgccctctgt tcccctcatg gcagcagtgc 32940
 tgggctccct gcttcccgtg ggggtcccgc gcactcggca cctgtgagga ttggaggatg 33000
 ggctgtactc ttttaaagac tggggatagg gccaggttct gtggctcatg cctgtaatcc 33060
 cagcactttg ggataccgag gcgggtggat cacttgaggt cgggagattg agaccggcct 33120
 ggccaacatg cagaaacccc gtttctactg caaacacaaa aaattagccg ggcatggtgg 33180
 cgcatgcctg taatcccagc tatttcggag gctgaggcag gagaatcgct tgaacctggg 33240
 aggtggaggc tgcagtgagc ggagattgtg ccattgcact gaagcctggg tgacagagtg 33300
 agaccctatc tccaaaaaaa aaaagaaaaa agagagaggc cgaggcgggt ggatcacgaa 33360
 gtcaggaggt cgagaccatc ctggctaata tgggtgaaacc ctgtctgtac taaaaataca 33420
 aaaacaaaat tagccaggca tgggtggcggg cacctgtagt cctagctact cgggaggctg 33480
 aggcaggaga atggcgtgaa tccgggaggc ggtgctctca gtgagccgag attacgccac 33540
 tacactccag cctgggcgac agagcaagac tgtctcaaaa aaaaaaaaaa gaaaacacaa 33600
 tggaaatata tttttattct ttgtctctga ttttctgatt ttaagagtaa cacgtttttc 33660
 ttggagaaat ttagaaaata aacacaggct ccaaacagtg ctgctggctc ttgggtgaacc 33720
 cgggtggagc ttttaaccct ctctgaagct gcgctctcca ggctgatcag accagctatc 33780
 cttttccgaa acattccctt ggggacctga gaggagccag ggctgagccg cctcagagtg 33840
 agcccagggc tccatgggct catgcttgtc ctggaaggct cttctctagc ctctggtgac 33900
 cgtgggaaag cttccccgc ctcccacagc catcacgctc tggcttggat gccttgccctg 33960
 ccatcagctc tcagacgccc acctggtgcc cgctgctggc tgctcaggac gtacgctggt 34020
 tgcctgccct tcaaagcagg aagatgcac actcttggac tccctgtctt cttagcagcc 34080
 cggagcctcc tttgccccta gggccctgct cctgtgctcc cccacatcat gctcccagca 34140
 ccatgccacc tccctcgaac agagcagggt tggaggcctc tgagtgtcag cgcctttcag 34200
 atcatggtca tgcaagttgt taagtgaatt tctacacccg ggttttttttg ctaaggagaa 34260
 aatgaaggat gagaagtcct ggtttctttc tttctcagta tttgggagtg cacagggtgag 34320
 ccagtggcca tgggtgtgcct tgagggtgca cctgctgctc ttctcagcac ggccgacacc 34380
 tcccattgcc tggagcctga gccacagacc tggaggccgt gggctctctgt tttgaaagaa 34440
 acagacgttt cgaggctttc ttcttgggggt ttcaaaggat atgtttttttt gtatgaacat 34500
 ttttgagtat tatctatgac atttgttcag agtgatgtta tggatacgag tcttgaagga 34560

ctgatttggt tctctctact actttattta ttcagccaaa gagacagatg ctttcagtaa 34620
 aacacctggt tctagatgaa gagagttcat gtgttttttg gcagtcttac tttttttttt 34680
 gtcttctctc tccttcttgc tctggtttta tcttcagata gctctgagga gcagtttctt 34740
 ttcttggcca tctcttactt atttaaggta atttggtgctt tataaatttt ggtgtttttg 34800
 aaagcgaagt ttctagggac tttgtagtaa atacgtctga aattcagaat gatctgtgaa 34860
 ttttacaacc cgtttactaa agccacacaa tgatcagcat ctgaagactt cagctaagac 34920
 accacttagg acgtgcaatg gccctaactt aacacatttg ccagtcgtca cctctgtttg 34980
 cagagccagc gcaggtccag ctgagttgtg ctgggttggt ggagcgtgcg cctctctctg 35040
 gggctctgtg gctgtgccct caccagggca gccaggctgt ctgcacaggc agcctcagtg 35100
 ggtcctggtc caggagaggt tgcgatctta cctggacgct gccccttctg tgctgacagt 35160
 gtcattgggag tagaaattct ccatcaggag catcccagtg tgttcctgta gcagcagcta 35220
 ctctgtgggt cagagtgcag gggagggggc tctgggtgagc agggcctggg cctcgttgct 35280
 aacagccact ctcaaggcct tggccgaagg tgatttgaaa gacgggaccg gtgatgtggt 35340
 cccagaaggt gcgtgggacc tgatgactga cagcatgtgt cctttgggtc agtggacagg 35400
 ttgctggacc ccgccaacag agcccacctg ggcctggagg agcagctgag agagctgctg 35460
 gacatgctcg acctcacctg cgctatgaag tccagcggct cccggagcaa gcgggcaaag 35520
 ctgctcaaaa aggaaattgc ctttctccga aacaagctga gccagcagca cagccagccc 35580
 ctgcccacgg ggccaggctt ggaaggcttc gaagaggacg gagctgcgct ggggccggag 35640
 gcgggcgagg aaggtaagcc acggtgaggg ctgccccgct ttcttgccgt cacggctcag 35700
 gctcaggcac aggcacaggc acagaccag acacgggcag cagcgttggt gtgcacggag 35760
 gctcaggtta gcgccagttc cgtaaaattc acaaattccc gaacacttga ttttcccact 35820
 gggctgttgt ttctttttaa ggagacgtta tcatttgtgc tgtctgtgtt ttctcacata 35880
 atcagttgat aaaatttaag gcacagaata aaatttttct ctttaacttg ggcctgagaa 35940
 actatttgaa aagaatgggt tagatcttac actgtgttga agagtaaagg tagcctgaat 36000
 gatgagagtg agaaatgttt ttgttgcat tttgtaaattg gcggggagcc atttaattgt 36060
 gcgctaggcc atcaaattgca gcattttgaa agggctgttt tagattattg gcacttaaac 36120
 actcctccct gatgttgtga agggagggac ggtacctgca gaggccttgt ctgccccgag 36180
 ggtccccggc cagctcctct gcagagttgg taggtagcag ccgtgatcgt tgccaaagaa 36240
 tgtctgggga aagggtccagg cagcaaagat gtttttatga ctaagaaata aatgtaggag 36300
 aagaaatccg cctatctaga aggagaacta gcaaaacca gggcaaatgc aggcttggca 36360
 tggtgaggag ggagtttacg agcgatggtc tgtgtacatc taggttggca gggtagaggaa 36420
 gtagagctgg tgtgtcccat atttggtgaa agaggaattt taaaactaaa tcggatttgg 36480

agacgttgat agaaagtcac tgccactgtg ttttgccgtg tgtctctgca gttgaccttg 36540
 acacgctccg gttacacgaa gcagacagta gggtcagacc ctcttgccac aacattggca 36600
 gctggttctc agccttcttg cccttcccct gcacatccaa gtcagaacag cagacagttg 36660
 aaataagaca cagaattggt gtacaatttt cttttaaaac aaagacgttc atctcttatg 36720
 gacgtctgtg ctcgagtgta cgcagtgtct tcatgatgcc acacaggtca gcgggtcccc 36780
 atggtctcag ggcccttttc atcccctgaa cttcttgtgg ctctgtctg cagagcttcg 36840
 ccccttttat cctgttttac atttgggcaa cttgaatgag agccattaag cttttctttc 36900
 tcttagaaac ttactcctgt gcacacagct gagaatttct tatgcatcct ttggtcatt 36960
 ttgtgcctgt ttgttggtct gttctgggat tcttctgttt agttcaccat agtatagtac 37020
 ccagactcta gaatgtctct tttggggctg gcgtcaatta cgtaggtcc atagaataaa 37080
 ggcttttcta tttcaagggt agcgcttggg tttctcttct ccaaaggct gttattgttg 37140
 gcacgtgatc cctcctctgc tttgggggccc tcgtgttttt tctgaggaaac gtttgcctac 37200
 aggaccctct ggcagccata cctgtccggg ccctgggccc cccattgtg cccaggagtc 37260
 caggctgggt cctggtcccg cagcctttgc ttccgtgcag cccaccgagg catcctctgg 37320
 ctgtgtgggc ctccctctag gtggactgcc cttctgtcct tgtcttctgt gccctgagga 37380
 ctgagctgtg aggaggttgc cagcaggtcc actcaccatt ctcaaaatag cttccagcgc 37440
 ttctcttcgc acacagttgt accattttca ctgtagagca gtaagttacc tttttagctt 37500
 ttttttttta tcagtggctt cctccattgc cattgggtca cagggtgag gaccccgcac 37560
 ggttattggg tcacaggact gaggacctg catggttgag gacctcgcat gggtggagct 37620
 gcttggccct ccgctgtggt gagcgtggct cctcataaca cccaacgctt agcaggagcc 37680
 tttggagacc atttttgttg ctttgtagct ctggaatcat tttattttta tttttttct 37740
 ggggccattg taaggaaaac acaatgtgac tgagacacag gtggcctcga atcacctgtt 37800
 acaccaatg ttcacaatga ctcaattgat gaaaaggcg tttcctgtga gctgatggga 37860
 aggtgggtgg agaagcaagt ggcctgaaag tccactccga ggtgaggcg gtggggctct 37920
 gccctcaga gaagtgggg gtctctggag aggtggggcg ggcggggtcc tgcccctcag 37980
 agaagatggg ggtctgtgga gaggcgagtg ccagactgga atggggaagt cattagggaa 38040
 gacagccaca agcaagcttg ttggttggt tagaaatgtg acttacgata cttataatag 38100
 gatgaaaacc aagaatcaga catttaaaaa cccagcatga acgttctcat cagtaaagcc 38160
 tctccatcat ttttcttaa tataaccaca atattttcac acctgaaaag agtaattcct 38220
 caatggcaaa tacacagtct cacttaaaact gtaatttaac agtctgtgtg ttgtgtttat 38280
 attttgcttc cacttgaagc agcttgaaca aggctgtagc cgttgagtgc caggcccctg 38340
 agactttagt gtctgaagga aggagatgcc tcctgggctg cactgttggc cccggtctgc 38400

cccttgtaca ggccctcccg caggctctgc cgtagccgcg actcagtctg ccgagtgcag 38460
 aggttcgggc tgttgctgtg agctggccct gccagggctg acttccccgg ggctgcgtgt 38520
 cagctgggtg ggggccacgc cacagctggg attccagcct gtctgatgag cctgtgggac 38580
 ggatctgcag ctggcagagg gcttgtctcc accctgtgct gaggttggtg tctgctgaca 38640
 tggtgaggca ggtcagtgc gtcaggagga gcaggggtccg caccgtgagc tggtgaaagg 38700
 ggcttgcatt tcgagtcctt gggatttcca gtcacagtga tgcccttgcc agaggctccg 38760
 gggcgtcccc tttagtccag ggcagctgtg gtcccttgaa gtgcacgtgt ggaagtcccc 38820
 ctgccgtgca cgggcgctga gcacgtggag aactgtggcc ctcaggaaag gaaggatgca 38880
 tgttcctcgg ggtagaaggt gcggagtacg cagctttgtc ccttcaggat tgcagaggac 38940
 ggaggcttga ggcaggggtg ttggagggac agtgtggttg ctgctcgtga gctgaaatgt 39000
 ctgcctgtgt tacgctgtga ttttccaaag cgtcttgtgg cttcgataca gaaacctttg 39060
 attgacatag caggtggcca ggtgggcaaa atattgcgga attgttcttg cagacttaac 39120
 taggttttag tttcttagag tagctggttt ggattagata accatthttgc catttggaat 39180
 atgtgacaag ttccagggac tggaattcta ggatcctgac tgatacattg acgggggtctg 39240
 agaggcttat gcccaggttg tatcgagcat gttctgagct tttcccatgt gcagttcagt 39300
 gacagcctga gatggcccag agtcagtgag ctttgagagg agtcttcctt cttgtatgtc 39360
 aacttccagt cacatttcca ctcgagggac gggttgggct tgcctgttgt gtcgtggagc 39420
 ctggaggctc aggtcctggt tttttctgat tttatggatg aggaactgag gcactgaacg 39480
 gtccgggggc ctctgtggtc ccggctgggg gaggggcctg tgtggtcccg gctgggcagg 39540
 tggctgagca cccagacgac agcgacctct gttcctgacc acggcctcct ttgtgcctct 39600
 ggtcagtgtc gcaggaacgc atccttttagc ctgcgtgctg gcggctggag cccggcctgc 39660
 atgtggagct gccgcagaga ggccaggggc gtgtgggggtg catgcatggc tgagacaggg 39720
 ccacaggggc ccgctttggg atagaccagg ggtagggaa gtccctgcag cgtggagatg 39780
 gttatgagtc ccacatgggg catcctagct gggccgagag ctggagggtgt ggaggagggg 39840
 ttgtggacca tggttgggga gggcctgctg caaggcaggg acgtctgcac gtctcagggc 39900
 tcctgccaga caggagtggg ggttctgagg gttgcaccag aagcaggaga tgggggttta 39960
 cggtgcccgt tagatagctg tggagaaggt gggctgtagg tggttggggg aaggggggtc 40020
 tcccggaggt gctgggagga gactccttca ggctccatct ctgctggccg ggtgctgttg 40080
 caccataga ggccaagact ggggttgacc actggccttg gccctgcaga ggccattggg 40140
 attgcagtga gcagttgaat gtcgtaatca gggcgagagc tcggggcggg ggggggaggc 40200
 taggaggtgg aagaggaggg aggaagtga gtagggcccc aggccctgtg cccagacac 40260
 cagctgcgtc aggacgctgc aggctcagtg ctgaatcccc tcctgctgct gtacaggcac 40320

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 2
CONTENANT LES PAGES 1 À 250

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 2
CONTAINING PAGES 1 TO 250

NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

Claims

1. A method for determining the predisposition for a mental disease, such as schizophrenia (SCH) and/or bipolar disorder (BPD) in a subject comprising
5 determining in a biological sample isolated from said subject one or more polymorphisms in the DNA sequence of chromosome 22q13 containing the NHP2L1, PACSIN2, SERHL, PIPPIN, BRD1, EP300, FAM19A5 and/or GPR24 genes or in a translational or transcriptional product of said genes.
- 10 2. The method according to claim 1, wherein the polymorphism is a single nucleotide polymorphism (SNP).
3. The method according to claim 2, wherein the predisposition is determined by determining the presence of one or more SNPs in the DNA sequence of one
15 individual gene selected from any of the genes of the group consisting of the NHP2L1, PACSIN2, SERHL, PIPPIN, BRD1, EP300 and GPR24 genes.
4. The method according to claims 1-3, wherein at least one of the polymorphisms is located in a non-coding region of the gene DNA sequence.
20
5. The method according to claim 5, wherein the non-coding region is an intron, or a region controlling the gene expression, such as a promoter region.
6. The method according to claims 1-4, wherein the at least one of the polymorphisms
25 is determined in a coding region of the gene DNA sequence.
7. The method according to any of the preceding claims, wherein two or more polymorphisms are determined.
- 30 8. The method according to claim 7, wherein the two or more polymorphisms are SNPs.
9. The method according to claim 8, wherein the two or more SNPs are determined in the DNA sequence of same gene, said gene is selected from any of the genes as
35 defined in claim 1.

10. The method according to claim 8, wherein the two or more SNPs are determined in the DNA sequences of two or more different genes selected from any of the genes as defined in claim 1.
- 5
11. The method according to any of the preceding claims, wherein the SNP is selected from the group consisting of SNPs having refSNP IDs: rs11561, rs5758405 rs8779, rs132806, rs2068943, rs2267487, rs881542, rs926333, rs1060387, rs1006407, rs6002408, rs4468, rs138855, rs2239848, rs138880, rs138881, rs20551, 10 rs2294976, rs2076578, rs1046088, rs133068, rs133069, rs133070, rs133073, rs6002408.
12. The method according to claims 1 to 11, wherein the predisposition to SCH and/or BPD is determined by determining of the presence of the risky allele of an SNP 15 selected from the SNPs as defined in claim 11.
13. The method according to claims 1 to 11, wherein the predisposition to SCH and/or BPD is determined by determining the presence of a specific haplotype comprising two or more of the SNPs selected from the SNPs as defined in claim 11.
- 20
14. The method according to claim 13, wherein the specific haplotype comprises at least one of the SNPs as determined in claim 11 and a polymorphism of the DNA sequence in the region comprising 100 to 10000 base pairs upstream or down stream from said SNP.
- 25
15. The method according to claim 14, wherein the polymorphism is SNP.
16. The method according to claim 15, wherein the SNP is selected from the group consisting of SNPs having refSNP ID No: rs132234, rs3752466, rs6010260, 30 rs137931, rs137932, rs3810971, rs2272843, rs1063900, rs715519, rs916005.
17. The method according to claim 16, wherein the SNP is a part of a haplotype comprising at least one of the SNPs selected from the group consisting of SNPs having refSNP ID NOs: rs4468, rs138855, rs2239848, rs138880, rs138881.
- 35

18. The method according to claim 15, the SNP is selected from the group consisting of SNPs having refSNP ID No: rs909660, rs710193, rs1573745.
19. The method according to claim 18, wherein the SNP is a part of a haplotype comprising at least one of the SNPs selected from the group consisting of SNPs having refSNP ID NOs: rs133068, rs133069, rs133070, rs133073.
20. The method according to claim 13, wherein the specific haplotype comprises a polymorphism in a chromosome region adjacent to the region containing a gene selected from the group consisting the NHP2L1, PACSIN2, SERHL, PIPPIN, BRD1, EP300 and GPR24 genes.
21. The method according to any of the preceding claims 14-20, wherein the polymorphism is in linkage disequilibrium with at least one of the SNPs identified in any of the claims 11, 16, 17, 18 or 19.
22. The method of claim 21, wherein the polymorphism is SNP.
23. The method according to any of the preceding claims, wherein a SNP(s) is(are) determined in
- i) a nucleotide sequence selected from SEQ ID NOs: 1-7 and 94,
 - ii) a sequence having at least 90% sequence identity with SEQ ID NOs: 1-7 and 94, or a fragment thereof, and
 - iii) a sequence being complementary to one of these sequences or a fragment thereof.
24. The method according to any of the preceding claims, wherein the SNP(s) is (are) determined in the transcriptional or translational products of the genes as defined in claim 1.
25. The method according to claim 24, wherein the transcriptional products of the genes being selected from
- (i) nucleic acid sequences identified as SEQ ID NO: 8-14, or fragments thereof,
 - (ii) nucleic acid sequences having at least 90% identity with SEQ ID NO: 8-14 or fragments thereof,

- (iii) nucleic acid sequences being complementary to any of the sequences of (i) or (ii),

said nucleic acid sequences comprising the polymorphism(s) of the corresponding genomic sequences identified as SEQ ID NO: 1-7 and 94.

5

26. The method according to claim 24, wherein the translational products of the genes being selected from

- ii) polypeptide sequences identified as SEQ ID NOs: 87-93 , or fragments thereof,

- 10 iii) polypeptide sequences having at least 90% identity with the polypeptide sequences of (i), or fragments thereof,

said polypeptide sequences comprising a polymorphism(s) corresponding to the polymorphism(s) of the corresponding nucleic acid sequence(s), said nucleic acid sequence(s) being identified as SEQ ID NO: 1-7 and 94 or SEQ ID NO: 8-14.

15

27. The method according to any of the preceding claims, wherein the presence or absence of a polymorphism is detected in a target nucleic acid sequence isolated from a biological sample.

- 20 28. The method according to claim 27, said method comprising amplification of the target nucleotide sequence.

29. The method according to claims 27 or 28, wherein the nucleotide sequence is a genomic DNA sequence, a mRNA sequence, or a cDNA sequence.

25

30. The method according to claim 27, wherein amplification comprises use of a primer pair selected from the oligonucleotide sequences identified as SEQ ID Nos: 15-86.

- 30 31. The method according to any of the preceding claims, wherein the presence or absence of the polymorphism is determined in a variant protein, said variant protein being a translational product of a gene selected from the genes as defined in claim 1, wherein said variant protein comprising a polymorphism corresponding to the polymorphism of the nucleic acid sequence encoding said variant protein.

32. The method according to claim 31, wherein the variant protein is NHP2L1 protein (SEQ ID NO: 87) having the sequence wherein amino acid residue Thr at position 43 is substituted for amino acid residue Ala.
- 5 33. The method according to claim 31, wherein the variant protein is SERHL protein (SEQ ID NO: 89) having the sequence wherein amino acid residue Ala at position 2 is substituted for amino acid residue Val.
- 10 34. The method according to claim 31, wherein the variant protein is SERHL protein (SEQ ID NO: 89) having the sequence wherein amino acid residue Ser at position 46 is substituted for amino acid residue Ala.
- 15 35. The method according to claim 31, wherein the variant protein is EP300 protein (SEQ ID NO: 92) having the sequence wherein amino acid residue Ile at position 997 is substituted for amino acid residue Val.
- 20 36. The method according to claim 31, the variant protein is EP300 protein (SEQ ID NO: 92) having the sequence wherein amino acid residue Gln at position 2223 is substituted for amino acid residue Pro.
- 25 37. The method according to claim 1, wherein the predisposition to a mental disease is determined by determining the presence of a haplotype comprising the SNPs having refSNP No: rs133068, rs133069, rs133070, rs133073 and one or more SNPs as defined in claim 18.
- 30 38. The method according to claim 1, wherein the predisposition to a mental disease is determined by determining the presence of a haplotype comprising the SNPs having refSNP No: rs4468, rs138855, rs2239848, rs138880, rs138881 and one or more SNPs as defined in claim 16.
39. The method according to any of the preceding claims, wherein the predisposition to a mental disease is determined by determining in a biological sample isolated from said subject the risky allele(s) of an SNP(s) as defined in claim 16.

40. A method for determining the absence of predisposition to a mental disease, such as SCH and/or BPD, in a subject comprising determining in a biological sample isolated from said subject a protective allele of an SNP(s) selected from the SNP(s) as defined in claim 11.
- 5
41. A method for determining a protection against a mental disease, such as SCH and/or BPD, in a subject comprising determining in a biological sample isolated from said subject a protective allele of an SNP(s) selected from the SNP(s) as defined in claim 11
- 10
42. An isolated oligonucleotide comprising at least 10 contiguous nucleotides being 100% identical to a subsequence of a gene selected from the NHP2L1, PACSIN2, SERHL, PIPPIN, BRD1, EP300 and GPR24 genes comprising or adjacent to a polymorphism or mutation being correlated to an mental disease.
- 15
43. The isolated oligonucleotide according to claim 42, wherein the polymorphism is located in the centre of the nucleic acid sequence.
44. The isolated oligonucleotide according to claim 42, wherein the polymorphism is located in the 5' end of the nucleic acid sequence.
- 20
45. The isolated oligonucleotide according to claim 42, wherein the mutation/polymorphism is located in the 3' end of the nucleic acid sequence.
- 25
46. The isolated oligonucleotide according to claim 42, wherein the sequence is adjacent to the mutation/polymorphism, either in the 3' or 5' direction.
- 30
47. The isolated oligonucleotide according to any of the claims 42-46, said oligonucleotide being selected from the oligonucleotides identified as SEQ ID NO: 15-86.
48. A diagnostic kit comprising at least two oligonucleotides as defined by any of the preceding claims 42-47.

49. The diagnostic kit according to claim 48, wherein the at least two oligonucleotides are the amplification primers or probes for determining a polymorphism associated with a predisposition to a mental disease, such as SCH and/or BPD, as defined in any of the preceding claims.
- 5
50. The diagnostic kit according to claim 49, wherein the probe is linked to a detectable label.
51. The diagnostic kit of claim 49, wherein the primers or probes are selected from the
- 10 nucleic acid sequences identified as SEQ ID NOs: 15-86.
52. A variant protein as defined in any of the claims 32-36, said protein being indicative of a predisposition to a mental disease, such as SCH and/or BPD.
- 15 53. An antibody capable of selectively binding to a variant protein of claim 52 to an epitope comprising a polymorphism of the variant protein as defined in any of the claims 32-36.
54. A diagnostic kit comprising an antibody selected from the antibody(s) according to
- 20 claim 53.
55. The diagnostic kit according to claim 54, said kit further comprising the diagnostic kit according to any of the claims 48-51.
- 25 56. A gene therapy vector comprising
- (i) a DNA sequence selected from the sequences identified as SEQ ID NO 1-7 and 94, or a fragment thereof, or
 - (ii) a DNA sequence selected from the sequences identified as SEQ ID NOs: 8-14, or a fragment of said DNA sequence.
- 30
57. The gene therapy vector according to claim 56, wherein the DNA sequence or a fragment thereof comprises the protective allele of an SNP selected from the SNPs as defined in any of the claims 11, 16, 17, 18 or 19

58. A method of treatment of a subject having the predisposition to a mental disease such as SCH and/or BPD, said method comprising administering to said subject a therapeutically effective amount of a gene therapy vector as defined in claims 56 or 57.

5

59. A vector comprising a nucleic acid sequence selected from the nucleic acid sequences identified as SEQ ID NOs: 1-14, or a fragment thereof, said sequence, or said fragment comprising a polymorphism associated with a predisposition to an mental disease according to any of the claims 1 to 39, said sequence being operably linked to a promoter sequence capable of directing the expression of a variant protein encoded by said sequence.

10

60. A compound capable of

15

i) inhibiting expression of a gene selected from the genes according to claim 1 said compound being selected from an isolated antisense nucleotide sequence or an nucleotide sequence complementary to the regulatory region of said gene, said nucleotide sequence being capable of forming triple helix structures that prevent transcription of said gene, and/or

20

ii) inhibiting activity of a transcriptional product of a gene selected from the genes according to claim 1, said transcriptional product being as defined in claim 25, said compound is selected from an isolated antisense sequence or a ribozyme molecule, and/or

25

iii) inhibiting activity of a translational product of a gene selected from the genes according to claim 1, said transcriptional product being as defined in claim 26, said compound is selected from an antibody molecule against said transcriptional product, or a molecule capable of interfering with biological activity of said transcriptional product.

30

61. Use of a compound according to claim 60 for the manufacture of a medicament for treatment of a mental disease.

62. The use according to claim 61, wherein the disease is SCH and/or BPD.

35

63. A pharmaceutical composition for the treatment of mental disease, such as SCH and/or BPD, comprising a compound according to claim 59.

64. A method of treatment of a mental disease, such as SCH and/or BPD, comprising administering a compound according to claim 60 or a pharmaceutical composition according to claim 63.
- 5
65. A method of screening for a candidate compound for therapeutic treatment of a mental disease, such as SCH and/or BPD, said method comprising an in vitro or in vivo model system comprising a gene according to claim 1 or a product of said gene, said product being a transcriptional product of the gene as defined in claim 10 25 or a translational product of the gene as defined in claim 26.
66. The method according to claim 65, said method further comprising a cell expressing a gene according to claim 1, a transcriptional product of said gene as defined in claim 25 or a translational product of said gene as defined in claim 26.
- 15
67. A method for prognosis of the likelihood of development of a mental disease comprising determining a polymorphism of a gene selected from the genes according to claim 1, said polymorphism being as defined in any of the claims 2-39.
- 20
68. The method according to claim 67, wherein the mental disease is SCH and/or BPD.
69. A method of predicting the likelihood of a subject to respond to a therapeutic treatment of a mental disease, such as SCH and/or BPD, said method comprising determining the genotype of said subject in the chromosome areas comprising the 25 NHP2L1, PACSIN2, SERHL, PIPPIN, BRD1, EP300, FAM19A5 and/or GPR24 gene.
70. The method according to claim 69, wherein the determining comprises assessing a polymorphism in the DNA sequence of the NHP2L1, PACSIN2, SERHL, PIPPIN, 30 BRD1, EP300, FAM19A5 and/or GPR24 gene, or the corresponding polymorphism in a transcriptional or translational product of said gene, and/or assessing a polymorphism in the DNA sequence of chromosome 22q13, said polymorphism being in linkage disequilibrium with a SNP of the NHP2L1, PACSIN2, SERHL, PIPPIN, BRD1, EP300, FAM19A5 and/or GPR24 gene, said SNP being associated 35 with a predisposition of a subject to a mental disease, such as SCH and/or BPD..

71. The method according to claim 70, wherein the polymorphism comprises or corresponds to an SNP selected from the SNPs having refSNP IDs: rs11561, rs5758405 rs8779, rs132806, rs2068943, rs2267487, rs881542, rs926333, 5 rs1060387, rs1006407, rs6002408, rs4468, rs138855, rs2239848, rs138880, rs138881, rs20551, rs2294976, rs2076578, rs1046088, rs133068, rs133069, rs133070, rs133073, rs6002408.
72. A method for assessing a therapeutic treatment for a mental disease, such as SCH 10 and/or BPD comprising using genotype data.

1/13

GPR24							<i>Haplotype frequency</i>			<i>Empirical p-values</i>		
<i>Haplotype</i>							<i>Controls</i>	<i>BPD</i>	<i>Schiz.</i>	<i>Combined</i>	<i>BPD</i>	<i>Schiz</i>
<i>rs909660</i>	<i>rs133068</i>	<i>rs133069</i>	<i>rs133070</i>	<i>rs133073</i>	<i>rs710193</i>	<i>rs1573745</i>						
C	G	A	G	T			0.0183	<0.0001	<0.0001	0.0005	0.0238	0.0929
	G	A	G	T	*		0.0202	<0.0001	<0.0001	0.0006	0.0257	0.0675
	G	A	G	T	C		0.0176	<0.0001	<0.0001	0.0070	0.0505	0.1390
		A	G	T	C	C	0.0102	<0.0001	<0.0001	0.0372	0.1913	0.1728
		A	G	T			0.0202	<0.0001	<0.0001	0.0006	0.0264	0.0683
			G	T			0.0254	0.0041	<0.0001	0.0050	0.0252	0.0279
C	G	A	A	T			0.0386	0.0086	0.0486	0.2406	0.0141	0.5868
C	G	A	A	T	C		0.0476	0.0102	0.0403		0.0118	
	G	A	A	T	C	**	0.0629	0.0097	0.0927	0.4174	0.0017	0.1858
		A	A	T	C		0.0629	0.0191	0.0991	0.7430	0.016	0.1341
	G	A	A	T	C	C	0.0045	<0.0001	<0.0001		0.0065	
	G	A	A	T			0.0537	0.0082	0.0897	0.4945	0.0021	0.0841
	G	A	A				0.0522	0.0162	0.0977	0.9692	0.0256	0.0442
		A	A	T			0.0537	0.0165	0.0949	0.8646	0.0247	0.0534
T	G	A	G	C		***	0.0129	0.0619	0.0236	0.0263	0.0296	0.4308
T	G	A	G				0.0169	0.0621	0.0380	0.0330	0.0474	0.3964
	G	C	G				<0.0001	0.0121	0.0065	0.0723	0.0468	0.0405
C	G	C	G				<0.0001	0.0128	0.0066	0.0953	0.0459	0.1130
C	G	C	G	C		****	<0.0001	0.0128	0.0070	0.0923	0.0454	0.1244
	G	C	G	C	C		<0.0001	0.0125	0.0058	0.0996	0.0404	0.1800
	G	C	A			****	0.0027	<0.0001	0.0148	0.3363	0.9430	0.0150
	G	C	A	T			0.0028	<0.0001	0.0154	0.3609	0.9003	0.0271

* - **** = Core haplotype 1-5

Figure 1

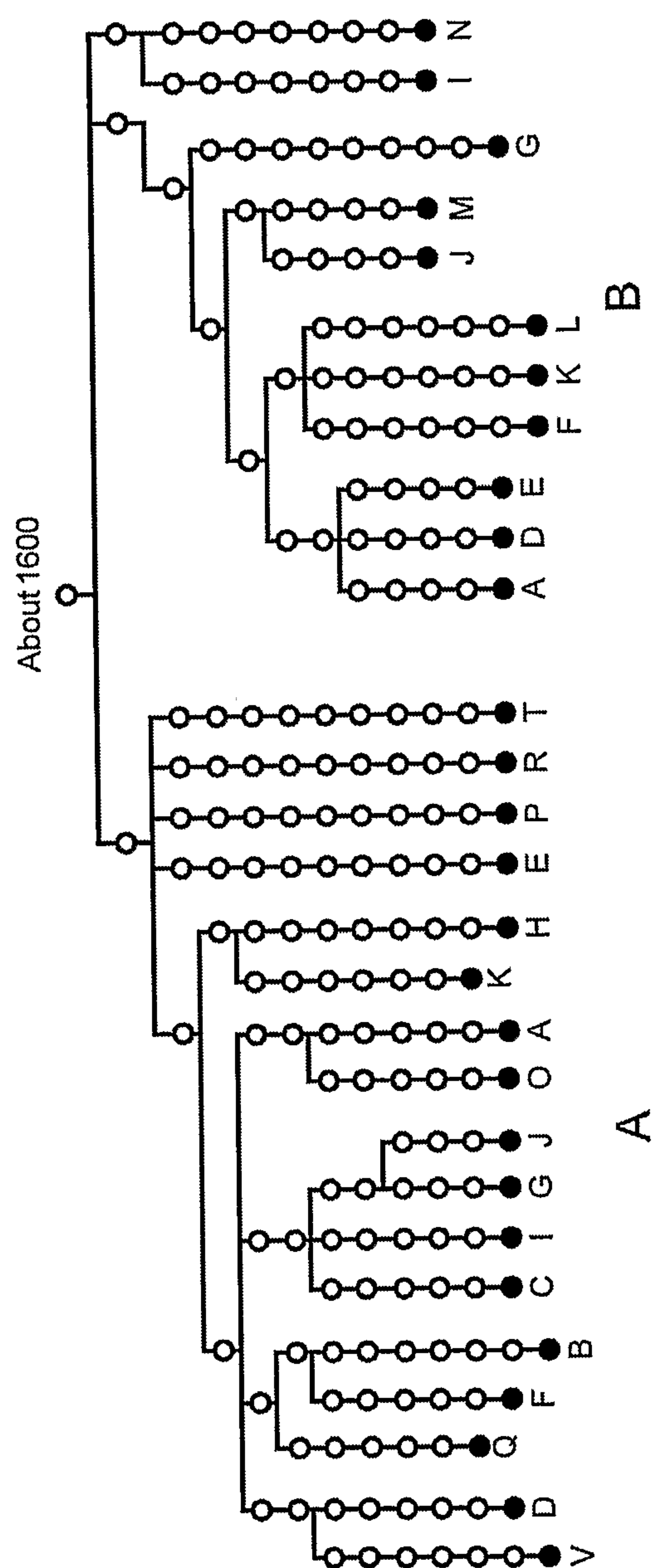
2/13

Haplotype frequency				Empirical p-values			
		Controls	BPD	Schiz.	Combined	BPD	Schiz
BRD1	rs4468						
	rs138855	G					
	rs2239848	G					
	rs138880	A					
	rs138881	C					
		0.831514	0.768511	0.766582	0.0216608	0.041435	0.000001
		0.0104547	0.0936689	0.0992718	0.00000028	0.00000155	0.000156
		0.0080259	0.0591138	0.0667076	0.00004	0.000106	0.001468
		0.00267616	0.025713	0.032581	0.002168	0.001808	0.00983
		0.158177	0.0880647	0.292274	0.13183	0.20842	0.00797
		0.242692	0.163933	0.381749	0.08825	0.24293	0.00463
		0.242612	0.164938	0.39643	0.06459	0.244	0.00717
		0.0249551	0.0205075	0.0848486	0.04292	0.95744	0.00504
		0.239598	0.164974	0.105374	0.05278	0.31824	
		0.024647	0.0205072	0.0800716	0.04051	0.96155	0.01474
		0.0105332	0.0895483	0.0950648	<<0.000001	0.000003	0.000004
		0.002690	0.0191269	0.0287835	0.00635	0.01582	0.00336
		<5e-5	0.0193641	0.0179351	0.00477	0.01626	0.00306
		0.0590265	0.0756265	0.109756	0.11191	0.35797	0.03349
		0.0636609	0.0949446	0.121647	0.07074	0.22774	0.03636

2/13

Figure 2

3/13



A: Bipolar affective disorder (17 affected)
B: Schizophrenia (11 affected)

Figure 3

4/13

Combined vs. controls		Single marker		2-marker		3-marker		4-marker		5-marker	
Gene	Marker	T1	T4	T1	T4	T1	T4	T1	T4	T1	T4
ADSL	rs909669	0,0036	0,0036								
	22S279	0,0705	0,0268	0,0805	0,0501						
GPR24	rs133068	0,1461	0,1461	0,1818	0,127	0,4823	0,3996				
GPR24	rs133069	0,074	0,074	0,0716	0,1107	0,2017	0,2023	0,2672	0,2064		
GPR24	rs133070	0,0116	0,0116	0,0145	0,0828	0,0525	0,1831	0,0064	0,0023	0,0059	0,0031
GPR24	rs133073	0,0546	0,0546	0,004	0,0188	0,0268	0,0626	0,0479	0,1278	0,0063	0,0023
ST13	rs710193	0,7631	0,7631	0,4522	0,4361	0,1158	0,1697	0,1755	0,2417	0,1722	0,239
Schiz. vs. controls		Single marker		2-marker		3-marker		4-marker		5-marker	
Gene	Marker	T1	T4	T1	T4	T1	T4	T1	T4	T1	T4
ADSL	rs909669	0,1178	0,1178								
	22S279	0,0322	0,0174	0,0447	0,008						
GPR24	rs133068	0,0197	0,0197	0,0542	0,0476	0,099	0,0353				
GPR24	rs133069	0,01	0,01	0,0469	0,0353	0,187	0,1727	0,1481	0,0495		
GPR24	rs133070	0,2107	0,2107	0,3299	0,3299	0,4456	0,4456	0,0426	0,0367	0,0156	0,0054
GPR24	rs133073	0,0083	0,0083	0,5104	0,5104	0,3269	0,3269	0,4486	0,4486	0,0432	0,0375
ST13	rs710193	0,7065	0,7065	0,3844	0,3368	0,3917	0,5842	0,3957	0,5862	0,3932	0,5852
BPD vs. Controls		Single marker		2-marker		3-marker		4-marker		5-marker	
Gene	Marker	T1	T4	T1	T4	T1	T4	T1	T4	T1	T4
ADSL	rs909669	0,0373	0,0373								
	22S279	0,3548	0,239	0,4799	0,3873						
GPR24	rs133068	0,8365	0,8365	0,4893	0,5473	0,8745	0,7353				
GPR24	rs133069	0,6815	0,6815	0,6484	0,5936	0,5063	0,4663	0,7976	0,5424		
GPR24	rs133070	0,0195	0,0195	0,003	0,0159	0,0316	0,0972	0,1531	0,1076	0,2957	0,1557
GPR24	rs133073	0,4217	0,4217	0,0009	0,0093	0,0042	0,0303	0,0433	0,0997	0,1521	0,1056
ST13	rs710193	1	1	0,8264	0,861	0,085	0,194	0,1753	0,2821	0,1745	0,2831

Figure 4

[illegible]

6/13

SNP	Gene	1 marker p-value	2 marker p-value	3 marker p-value	4 marker p-value	5 marker p-value
rs909660	ADSL	3,72E-01				
rs133068	GPR24	8,54E-01	6,58E-01			
rs133069	GPR24	7,13E-01	1,79E-01	3,27E-01		
rs133070	GPR24	6,72E-01	4,59E-01	5,95E-01	1,14E-01	
rs133073	GPR24	5,63E-01	2,62E-02	1,55E-02	1,84E-02	7,89E-03
rs710193	ST13	2,32E-01	1,25E-01	7,15E-02	7,90E-02	1,90E-01
rs1573745	ST13	7,99E-01	4,43E-01	6,32E-02	1,35E-01	2,60E-01
rs20551	EP300	9,98E-02	4,58E-01	6,56E-01	1,51E-01	1,70E-01
rs2294976	EP300	3,85E-01	6,32E-01	7,64E-01	5,04E-01	1,74E-01
rs2076578	EP300	7,99E-01	8,87E-01	2,01E-01	3,15E-01	3,52E-01
rs1046088	EP300	5,43E-01	2,59E-01	2,38E-01	5,34E-02	1,68E-01
rs6002408	PIPPIN	6,11E-01	3,04E-01	4,80E-01	7,12E-01	1,05E-01
rs1006407	PIPPIN	5,89E-02	1,35E-01	7,72E-02	1,69E-01	2,54E-01
rs8779	NHP2L1	3,50E-01	1,00E-05	4,30E-04	2,90E-04	3,51E-03
rs132806	NHP2L1	5,78E-01	6,44E-01	4,68E-01	9,96E-01	9,74E-01
rs11561	NHP2L1	2,02E-01	1,99E-01	3,70E-01	2,54E-01	9,78E-01
rs5758405	NHP2L1	4,09E-01	6,74E-01	1,36E-01	2,91E-01	3,35E-01
rs2269665	SEPT3	1,00E+00	5,62E-01	9,24E-01	5,07E-01	6,49E-01
rs1062753	SEPT3	8,94E-01	9,30E-01	8,07E-01	7,87E-01	6,38E-01
rs881542	SEHRL	5,66E-01	9,57E-01	9,43E-01	8,13E-01	7,79E-01
rs926333	SEHRL	5,11E-02	3,17E-01	8,80E-01	4,99E-01	6,75E-01
rs1060387	SEHRL	1,25E-01	7,00E-02	2,31E-01	8,24E-01	7,73E-01
rs2068943	PACSIN2	3,59E-01	1,69E-01	2,34E-01	7,56E-01	7,62E-01
rs2267487	PACSIN2	8,30E-01	3,04E-01	2,99E-01	4,41E-01	6,76E-01
rs135085	C22ORF1	5,45E-01	8,36E-01	4,39E-01	5,55E-01	5,39E-01
rs132234	TAF A5	6,12E-01	5,51E-01	7,98E-01	8,00E-01	9,05E-01
rs3752466	TAF A5		7,97E-01	7,52E-01	9,09E-01	8,82E-01
rs4468	BRD1	2,15E-02	2,85E-02	2,41E-01	3,26E-01	4,56E-01
rs138855	BRD1	4,19E-01	3,18E-01	2,16E-01	7,84E-01	6,11E-01
rs2239848	BRD1	3,98E-01	6,66E-01	2,09E-01	2,54E-01	6,18E-01
rs138880	BRD1	4,61E-03	3,81E-02	5,00E-05	6,07E-02	6,58E-02
rs138881	BRD1	1,73E-01	6,81E-02	1,55E-01	2,70E-04	2,71E-01
rs6010260	MLC1	1,63E-01	4,04E-01	2,10E-01	3,09E-01	3,38E-03
rs137931	MLC1	5,80E-01	1,60E-01	1,93E-01	2,36E-01	4,55E-01
rs137932	MLC1	6,33E-01	9,21E-01	4,05E-01	3,06E-01	3,21E-01
rs3810971	MOV10L1	9,87E-01	9,88E-01	8,35E-01	2,92E-01	7,85E-02
rs2272843	MOV10L1	3,02E-01	1,97E-01	4,69E-01	4,30E-01	3,82E-01
rs715519	MAPK8IP2	7,98E-01	7,52E-01	4,57E-01	4,12E-01	3,36E-01
rs916005	MAPK8IP2	1,00E+00	9,55E-01	9,79E-01	7,15E-01	7,00E-01

Figure 6

SNP	Gene	7/13				
		1 marker p-value	2 marker p-value	3 marker p-value	4 marker p-value	5 marker p-value
rs909660	ADSL	1,14E-01				
rs133068	GPR24	9,36E-01	3,82E-01			
rs133069	GPR24	8,73E-01	1,24E-01	2,12E-01		
rs133070	GPR24	3,64E-01	6,18E-02	6,91E-02	1,38E-01	
rs133073	GPR24	8,47E-01	1,88E-01	1,48E-02	1,35E-02	7,58E-02
rs710193	ST13	2,76E-01	3,27E-01	2,24E-01	1,07E-01	1,46E-01
rs1573745	ST13	2,88E-01	1,39E-01	2,10E-01	2,46E-01	1,87E-01
rs20551	EP300	7,75E-01	6,76E-01	4,51E-01	4,96E-01	3,00E-01
rs2294976	EP300	2,92E-01	5,28E-01	5,46E-01	1,80E-01	1,41E-01
rs2076578	EP300	7,02E-01	6,14E-01	4,38E-01	4,72E-01	1,90E-01
rs1046088	EP300	2,33E-01	2,03E-01	6,03E-01	6,32E-01	8,72E-01
rs6002408	PIPPIN	1,22E-01	5,29E-02	2,22E-01	8,02E-01	5,72E-01
rs1006407	PIPPIN	1,02E-01	5,21E-02	1,77E-02	1,55E-01	5,86E-01
rs8779	NHP2L1	7,18E-01	2,98E-02	3,38E-02	1,76E-02	1,35E-01
rs132806	NHP2L1	3,95E-01	5,97E-01	7,57E-01	8,99E-01	7,72E-01
rs11561	NHP2L1	3,90E-04	2,73E-02	9,21E-02	1,74E-01	5,50E-01
rs5758405	NHP2L1	3,54E-01	5,75E-02	7,05E-02	3,73E-02	9,30E-02
rs2269665	SEPT3	6,32E-01	6,62E-01	6,24E-01	3,13E-01	2,38E-01
rs1062753	SEPT3	7,58E-01	8,07E-01	6,87E-01	4,57E-01	3,60E-01
rs881542	SEHRL	3,29E-01	8,73E-01	8,99E-01	6,27E-01	6,03E-01
rs926333	SEHRL	1,09E-02	6,02E-02	4,09E-01	3,00E-01	1,91E-01
rs1060387	SEHRL	1,52E-02	3,73E-03	3,79E-02	3,17E-01	5,28E-01
rs2068943	PACSIN2	2,03E-02	9,74E-01	2,10E-03	3,45E-02	3,25E-01
rs2267487	PACSIN2	3,45E-01	4,22E-02	1,34E-02	4,49E-01	7,37E-02
rs135085	C22ORF1	1,00E+00	7,26E-01	1,17E-01	1,08E-01	5,24E-01
rs132234	TAF5	7,01E-01	7,76E-01	9,33E-01	6,08E-01	3,17E-01
rs3752466	TAF5	0,00E+00	8,03E-01	9,36E-01	9,57E-01	7,11E-01
rs4468	BRD1	8,80E-03	1,39E-02	1,85E-01	5,08E-01	2,05E-01
rs138855	BRD1	5,01E-01	1,14E-01	5,92E-02	1,52E-01	3,81E-01
rs2239848	BRD1	1,00E+00	5,94E-01	4,70E-02	3,85E-02	1,08E-01
rs138880	BRD1	6,07E-03	1,61E-01	1,00E-05	2,03E-02	1,34E-02
rs138881	BRD1	2,90E-01	5,21E-02	1,36E-01	2,00E-04	5,17E-02
rs6010260	MLC1	8,23E-02	3,10E-01	7,63E-02	1,59E-01	8,30E-04
rs137931	MLC1	8,44E-01	1,84E-01	3,33E-01	1,10E-01	2,52E-01
rs137932	MLC1	7,61E-01	8,97E-01	2,82E-01	4,18E-01	1,51E-01
rs3810971	MOV10L1	3,43E-01	8,97E-01	9,06E-01	1,84E-01	3,10E-01
rs2272843	MOV10L1	2,70E-01	3,41E-01	7,13E-01	7,85E-01	1,39E-01
rs715519	MAPK8IP2	7,25E-01	6,42E-01	5,63E-01	5,24E-01	6,66E-01
rs916005	MAPK8IP2	8,48E-01	9,76E-01	9,19E-01	7,46E-01	7,17E-01

Figure 7

8/13

SNP	Gene	1 marker p-value	2 marker p-value	3 marker p-value	4 marker p-value	5 marker p-value
rs909660	ADSL	8,97E-01				
rs133068	GPR24	6,76E-01	9,38E-01			
rs133069	GPR24	6,75E-01	3,58E-01	3,59E-01		
rs133070	GPR24	9,37E-01	1,85E-01	7,79E-02	3,81E-03	
rs133073	GPR24	3,47E-01	1,47E-01	5,30E-03	2,60E-04	2,10E-04
rs710193	ST13	3,29E-01	1,20E-01	2,25E-01	2,52E-02	1,80E-03
rs1573745	ST13	1,59E-01	5,90E-01	1,23E-01	4,00E-01	1,59E-01
rs20551	EP300	1,64E-02	6,72E-02	4,18E-01	1,42E-01	3,32E-01
rs2294976	EP300	5,76E-01	2,21E-01	2,35E-01	4,61E-01	1,66E-01
rs2076578	EP300	8,72E-01	7,52E-01	4,23E-02	3,70E-02	7,32E-02
rs1046088	EP300	1,92E-02	7,28E-02	6,67E-02	2,76E-03	1,12E-03
rs6002408	PIPPIN	7,29E-01	5,11E-01	6,13E-01	3,88E-01	4,04E-02
rs1006407	PIPPIN	8,03E-02	4,42E-02	2,64E-01	1,02E-01	1,42E-01
rs8779	NHP2L1	3,36E-01	5,70E-04	2,56E-03	9,52E-03	1,86E-02
rs132806	NHP2L1	8,29E-01	8,64E-01	7,44E-01	9,93E-01	9,91E-01
rs11561	NHP2L1	2,54E-01	6,37E-01	7,84E-01	7,07E-01	9,94E-01
rs5758405	NHP2L1	4,93E-01	2,14E-01	2,47E-01	5,24E-01	4,90E-01
rs2269665	SEPT3	8,96E-01	6,43E-01	3,98E-01	4,17E-01	5,03E-01
rs1062753	SEPT3	9,30E-01	1,00E+00	9,45E-01	8,66E-01	8,51E-01
rs881542	SEHRL	1,29E-01	8,74E-01	9,84E-01	8,46E-01	8,06E-01
rs926333	SEHRL	4,11E-01	5,81E-01	8,86E-01	6,05E-01	7,17E-01
rs1060387	SEHRL	4,34E-01	2,62E-01	3,62E-01	8,18E-01	5,89E-01
rs2068943	PACSIN2	8,16E-01	4,26E-01	6,14E-01	8,50E-01	5,27E-01
rs2267487	PACSIN2	2,09E-01	3,89E-01	4,82E-01	7,92E-01	8,23E-01
rs135085	C22ORF1	4,23E-01	7,13E-01	9,49E-01	9,40E-01	9,79E-01
rs132234	TAF5	6,27E-01	5,36E-01	6,48E-01	9,33E-01	9,58E-01
rs3752466	TAF5		8,42E-01	6,74E-01	8,01E-01	9,42E-01
rs4468	BRD1	7,81E-01	6,24E-01	6,38E-01	2,33E-01	2,36E-01
rs138855	BRD1	4,38E-01	7,69E-02	8,64E-02	1,49E-01	1,15E-01
rs2239848	BRD1	2,50E-01	5,00E-01	1,74E-01	1,42E-01	1,86E-01
rs138880	BRD1	2,74E-02	7,97E-02	6,00E-05	1,88E-01	1,45E-01
rs138881	BRD1	2,47E-01	1,82E-01	2,10E-01	1,30E-03	3,80E-01
rs6010260	MLC1	4,01E-01	6,88E-01	6,11E-01	5,60E-01	3,46E-03
rs137931	MLC1	2,84E-01	2,79E-01	1,01E-01	4,43E-01	5,77E-01
rs137932	MLC1	3,23E-01	8,34E-01	3,98E-01	2,01E-01	6,34E-01
rs3810971	MOV10L1	5,37E-01	8,15E-01	8,02E-01	7,72E-01	1,39E-01
rs2272843	MOV10L1	5,06E-01	1,40E-01	3,28E-01	3,26E-01	3,90E-01
rs715519	MAPK8IP2	4,68E-01	6,39E-01	3,06E-01	4,47E-01	4,16E-01
rs916005	MAPK8IP2	8,43E-01	8,89E-01	8,33E-01	3,17E-01	3,34E-01

Figure 8

9/13

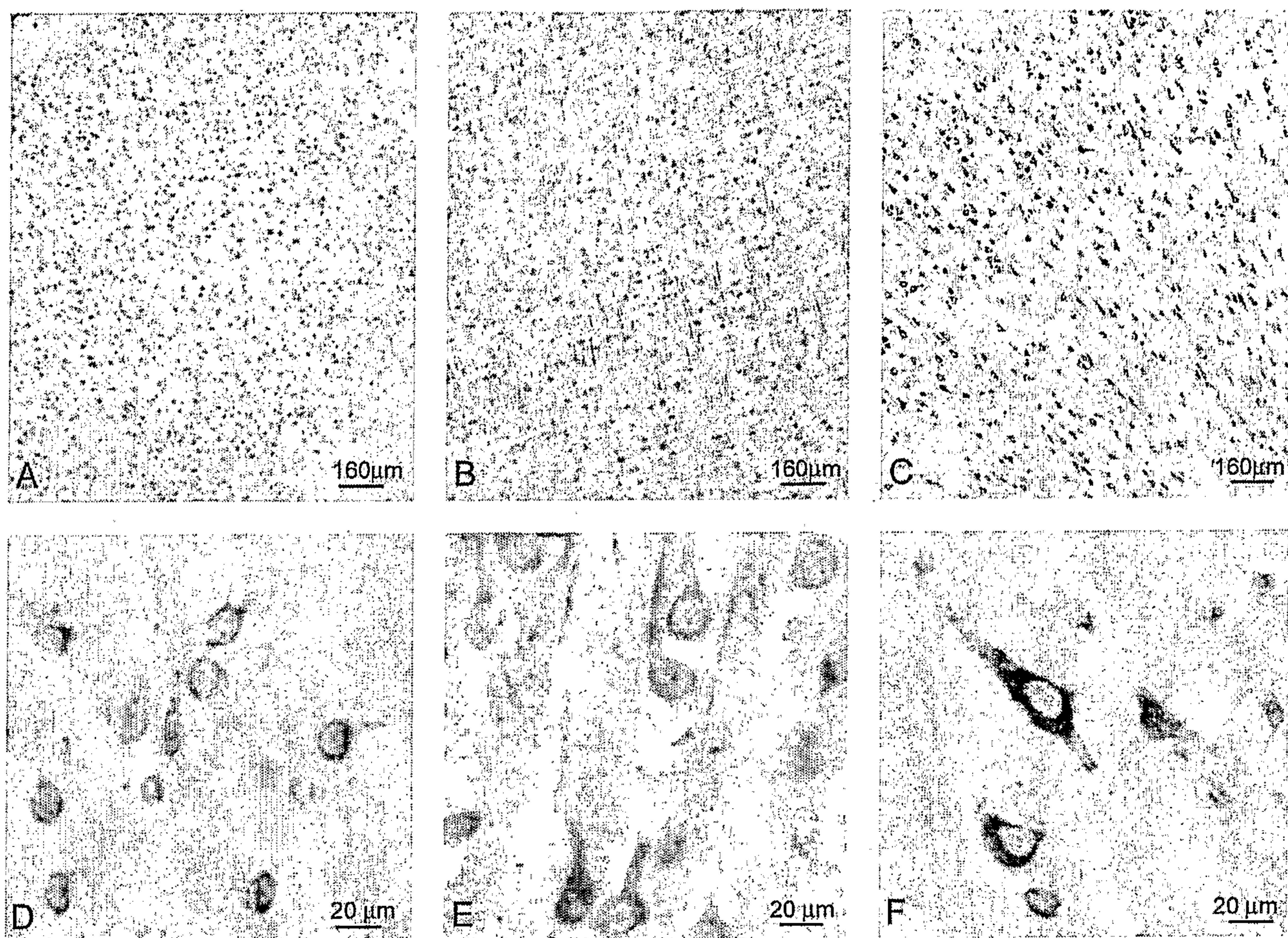


Figure 9

10/13

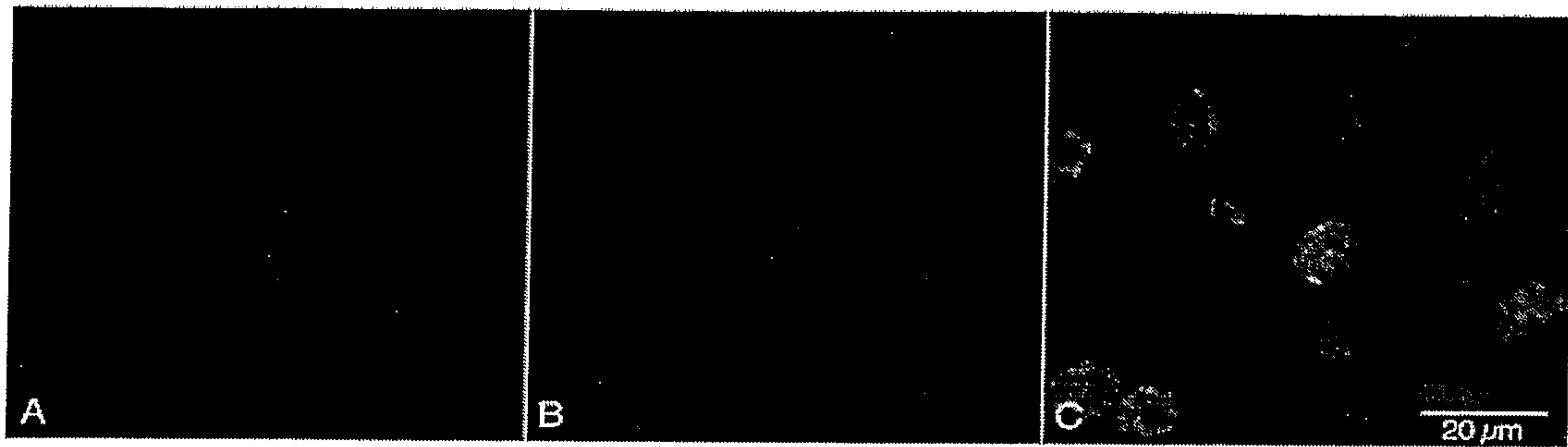


Figure 10

11/13

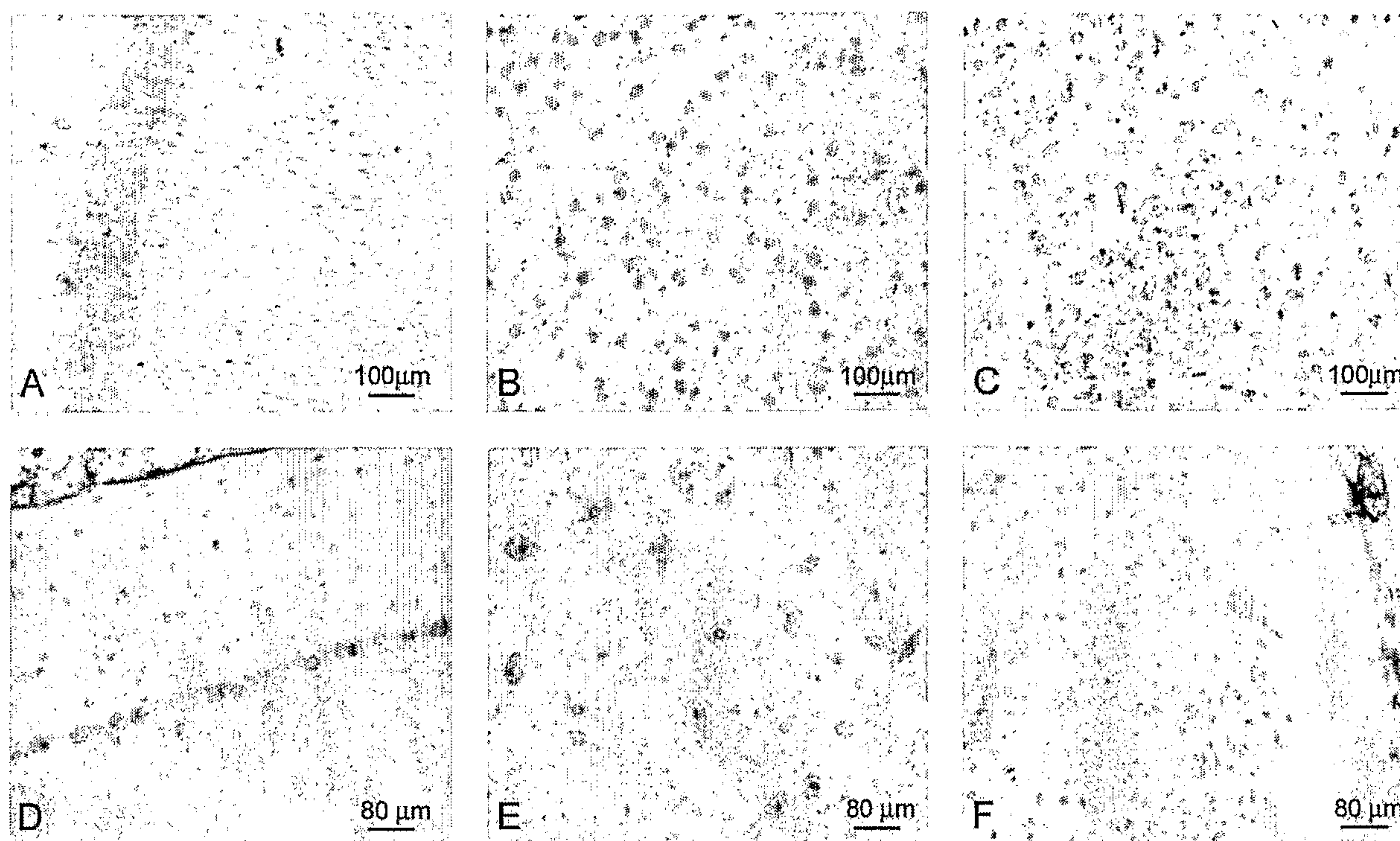


Figure 11

12/13

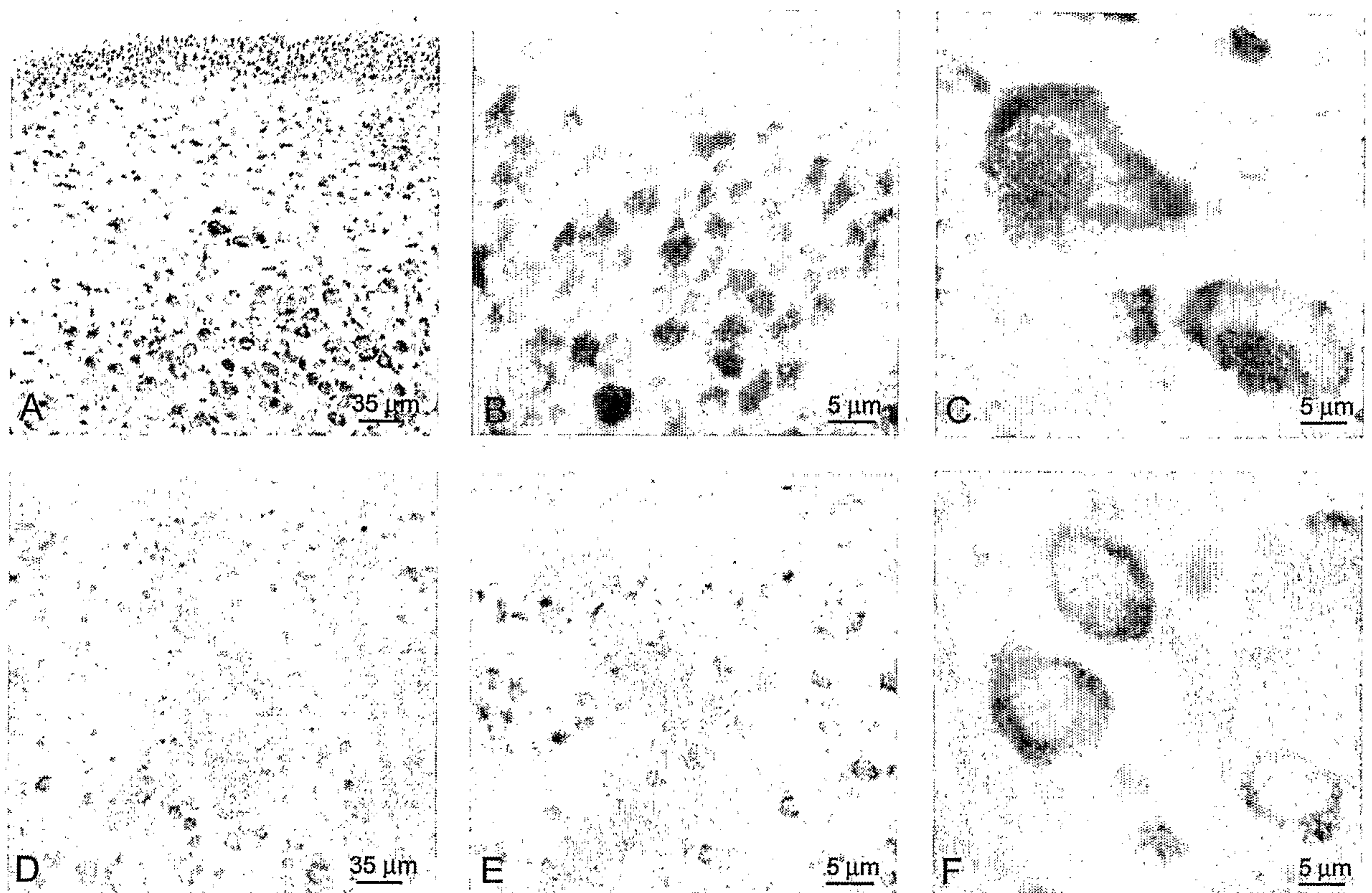


Figure 12

13/13

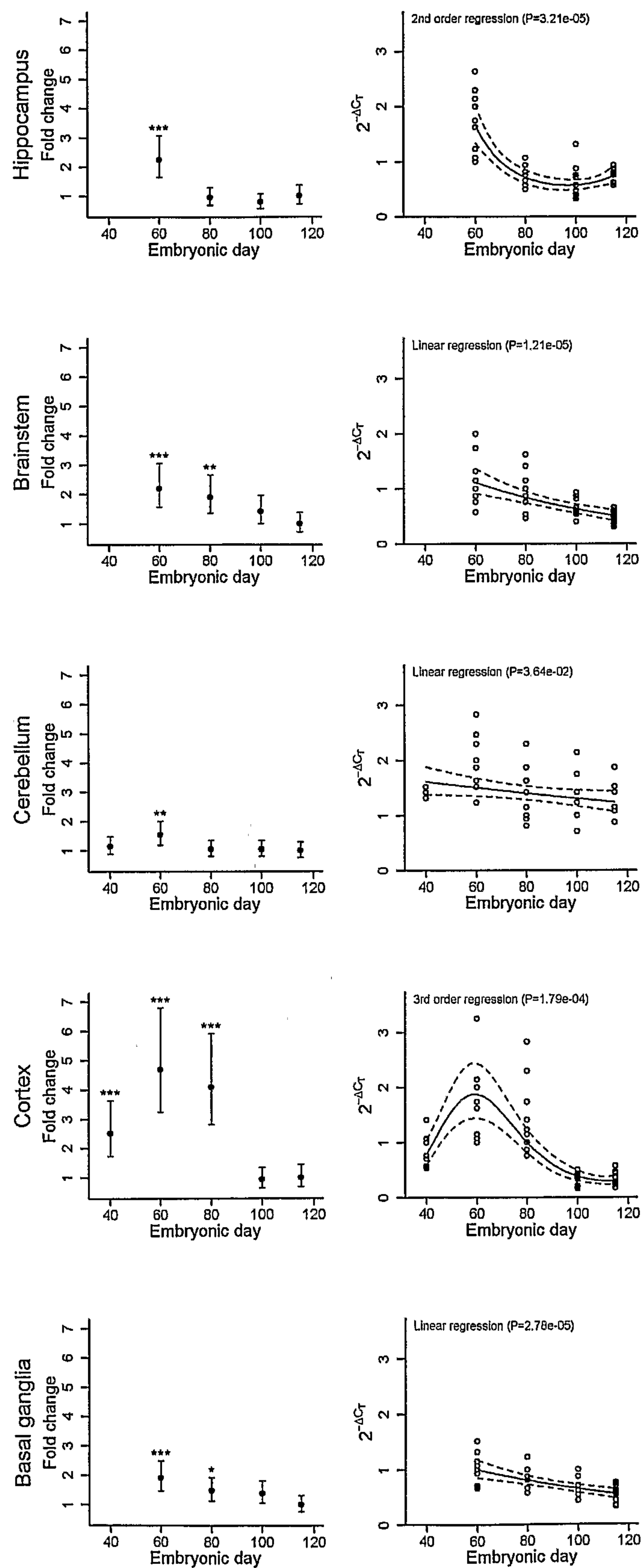


Figure 13