(12) STANDARD PATENT (11) Application No. **AU 2005254805 B2**
(19) AUSTRALIAN PATENT OFFICE

(54) Title
Human obesity susceptibility gene encoding a taste receptor and uses thereof

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

(21) Application No: **2005254805** (22) Date of Filing: **2005.06.14**

(87) WIPO No: **WO05/123948**

(30) Priority Data

(31) Number	(32) Date	(33) Country
60/578,830	**2004.06.14**	**US**

(43) Publication Date: **2005.12.29**
(44) Accepted Journal Date: **2009.02.12**

(71) Applicant(s)
Integragen

(72) Inventor(s)
Philippi, Anne;Roschmann, Elke;Rousseau, Francis

(74) Agent / Attorney
Davies Collison Cave, 1 Nicholson Street, Melbourne, VIC, 3000

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
29 December 2005 (29.12.2005)

PCT

(10) International Publication Number
WO 2005/123948 A1

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

(21) International Application Number:
PCT/IB2005/002309

(22) International Filing Date: 14 June 2005 (14.06.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/578,830 14 June 2004 (14.06.2004) US

(71) Applicant *(for all designated States except US)*: **INTEGRAGEN** [FR/FR]; CCI Essonne - Pépinière Génopole Industries, 4, rue Pierre Fontaine, F-91000 Evry (FR).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **PHILIPPI, Anne** [FR/FR]; 8, Résidence du Prieuré, F-77310 St Fargeau Ponthierry (FR). **ROUSSEAU, Francis** [FR/FR]; 11, rue Théophile Gautier, F-91300 Savigny sur Orge (FR). **ROSCHMANN, Elke** [DE/FR]; 58, Quai Jacques Bourgoin, F-91100 Corbeil Essonnes (FR).

(74) Agents: **GALLOIS, Valérie** et al.; Becker et Associés, 25, rue Louis Le Grand, F-75002 Paris (FR).

(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, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, 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.

(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, 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).

Declaration under Rule 4.17:
— *of inventorship (Rule 4.17(iv)) for US only*

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

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

(54) Title: HUMAN OBESITY SUSCEPTIBILITY GENE ENCODING A TASTE RECEPTOR AND USES THEREOF

(57) Abstract: The present invention discloses the identification of a human obesity susceptibility gene, which can be used for the diagnosis, prevention and treatment of obesity and related disorders, as well as for the screening of therapeutically active drugs. The invention more specifically discloses that the TAS1R1 gene on chromosome 1 and certain alleles thereof are related to susceptibility to obesity and represent novel targets for therapeutic intervention. The present invention relates to particular mutations in the TAS1R1 gene and expression products, as well as to diagnostic tools and kits based on these mutations. The invention can be used in the diagnosis of predisposition to, detection, prevention and/or treatment of coronary heart disease and metabolic disorders, including hypoalphalipoproteinemia, familial combined hyperlipidemia, insulin resistant syndrome X or multiple metabolic disorder, coronary artery disease, diabetes and dyslipidemia.

WO 2005/123948 A1

HUMAN OBESITY SUSCEPTIBILITY GENE ENCODING A TASTE RECEPTOR AND USES THEREOF

FIELD OF THE INVENTION

The present invention relates generally to the fields of genetics and medicine. The present invention more particularly discloses the identification of a human obesity susceptibility gene, which can be used for the diagnosis, prevention and treatment of obesity and related disorders, as well as for the screening of therapeutically active drugs. The invention more specifically discloses certain alleles of the taste receptor, type 1, member 1 (TAS1R1) gene on chromosome 1 related to susceptibility to obesity and representing novel targets for therapeutic intervention. The present invention relates to particular mutations in the TAS1R1 gene and expression products, as well as to diagnostic tools and kits based on these mutations. The invention can be used in the diagnosis of predisposition to, detection, prevention and/or treatment of coronary heart disease and metabolic disorders, including hypoalphalipoproteinemia, familial combined hyperlipidemia, insulin resistant syndrome X or multiple metabolic disorder, coronary artery disease, diabetes and dyslipidemia.

BACKGROUND OF THE INVENTION

Approximately three to eight percent of the total health costs of modern industrialized countries are currently due to the direct costs of obesity (Wolf, 1996). In Germany, the total costs (both direct and indirect) related to obesity and comorbid disorders were estimated at 21 billion German marks (29.4 US Dollar) in 1995 (Schneider, 1996). By 2030 these costs will rise by 50% even if the prevalence of obesity does not increase further.

Obesity is often defined simply as a condition of abnormal or excessive fat accumulation in adipose tissue, to the extent that health may be impaired. The underlying disease is the process of undesirable positive energy balance and weight gain. An abdominal fat distribution is associated with higher health risks than a gynoid fat distribution.

The body mass index (BMI; kg/m^2) provides the most useful, albeit crude, population-level measure of obesity. It can be used to estimate the prevalence of obesity within a population and the risks associated with it. However, BMI does not account for body compositon or body fat distribution (WHO, 1998).

Table 1: Classification of overweight in adults according to BMI (WHO, 1998)

Classification	BMI (kg/m²)	Risk of co-morbidities
Underweight	< 18.5	Low (but risks of other clinical problems increased)
Normal range	18.5 – 24.9	Average
Overweight	≥ 25	
Pre-obese	25 – 29.9	Increased
Obese class I	30 – 34.9	Moderate
Obese class II	35 – 39.9	Severe
Obese class III	≥ 40	Very severe

Obesity has also been defined using the 85th and 95th BMI-percentiles as cutoffs for definition of obesity and severe obesity. BMI-percentiles have been calculated within several populations; centiles for the German population based on the German National Nutrition Survey have been available since 1994 (Hebebrand et al., 1994, 1996). Because the WHO classification of the different weight classes can only be applied to adults, it has become customary to refer to BMI-percentiles for the definition of obesity in children and adolescents.

The recent rise in the prevalence of obesity is an issue of major concern for the health systems of several countries. According to reports of the Center of Disease Control and Prevention (CDC) there has been a dramatic increase in obesity in the United States during the past 20 years. In 1985 only a few states were participating in CDC's Behavioral Risk Factor Surveillance System (BRFSS) and providing obesity data. In 1991, four states were reporting obesity prevalence rates of 15-19 percent and no states reported rates at or above 20 percent. In 2002, 20 states have obesity prevalence rates of 15-19 percent; 29 states have rates of 20-24 percent; and one state reports a rate over 25 percent. Similar trends have been observed in other countries in Europe and South America.

Children and adolescents have not been exempt from this trend. Quite to the contrary, the increase in the USA has been substantial. Thus, between the 1960ies and 1990, overweight and obesity increased dramatically in 6 through to 17 year olds. The increments translate into relative increases of 40% using the 85th BMI-centile (calculated in the 1960ies) as a cutoff and 100% upon use of the 95th centile. In a cross sectional study of German children

and adolescents treated as inpatients for extreme obesity between 1985 and 1995, a significant increase of the mean BMI of almost 2 kg/m^2 over this ten year period has been reported. Within this extreme group, the increments were most pronounced in the uppermost BMI ranges.

The mechanisms underlying this increase in the prevalence of obesity are unknown. Environmental factors have commonly been invoked as the underlying cause. Basically, both an increased caloric intake and a reduced level of physical activity have been discussed. In England the increase in obesity rates has been attributed to the latter mechanism. Thus, in this country, the average caloric intake even decreased somewhat within the last two decades, whereas indirect evidence stemming from the increases in hours spent watching television and from the average number of cars per household points to reduced levels of physical activity as the relevant causative factor.

Genetic factors have previously not been considered as a contributing cause. Quite to the contrary, the fact that the increased rates of obesity have been observed within the last two decades has been viewed as evidence that genetic factors cannot be held responsible. However, it has been proposed that an increase in the rate of assortative mating could very well constitute a genetic contribution to the observed phenomenon. This hypothesis is based on evidence suggesting that stigmatisation of obese individuals represents a rather recent social phenomenon, thus invariably having led to increased rates of assortative mating. As a consequence, the offspring have a higher loading with both additive and non-additive genetic factors underlying obesity. Indeed, an exceedingly high rate of (deduced) assortative mating amongst the parents of extremely obese children and adolescents has been observed.

Potentially life-threatening, chronic health problems associated with obesity fall into four main areas: 1) cardiovascular problems, including hypertension, chronic heart disease and stroke, 2) conditions associated with insulin resistance, namely Non-Insulin Dependent Diabetes Mellitus (NIDDM), 3) certain types of cancers, mainly the hormonally related and large-bowel cancers, and 4) gallbladder disease. Other problems associated with obesity include respiratory difficulties, chronic musculo-skeletal problems, skin problems and infertility (WHO, 1998).

The main currently available strategies for treating these disorders include dietary restriction, increments in physical activity, pharmacological and surgical approaches. In adults, long term weight loss is exceptional using conservative interventions. Present pharmacological interventions typically induce a weight loss of between five and fifteen kilograms; if the medication is discontinued, renewed weight gain ensues. Surgical treatments are comparatively successful and are reserved for patients with extreme obesity and/or with serious medical complications.

Recently, a 10 year old massively obese girl, in whom a leptin deficiency mutation had been detected, was treated successfully with recombinant leptin. This is the first individual who therapeutically profited from the detection of the mutation underlying her morbid obesity.

Several twin studies have been performed to estimate heritability of the BMI, some of which have encompassed over 1000 twin pairs. The results have been very consistent: The intrapair correlations among monozygotic twins were typically between 0.6 and 0.8, independent of age and gender. In one study, the correlations for monozygotic and dizygotic twins were basically the same, independent of whether the twins had been reared apart or together. Heritability of the BMI was estimated at 0.7; non-shared environmental factors explained the remaining 30% of the variance. Surprisingly, shared environmental factors did not explain a substantial proportion of the variance. Both hypercaloric and hypocaloric alimentation lead to similar degrees of weight gain or loss among both members of monozygotic twin pairs, indicating that genetic factors regulate the effect of environmentally induced variation of energy availability on body weight. Metabolic reactions and changes in body fat distribution upon overeating and undereating are also under genetic control (reviewed in Hebebrand et al., 1998).

A large adoption study has revealed that the BMI of adoptees is correlated with that of their biological parents and not with the BMI of the adoptive parents. Depending on the family study, the correlation between the BMI of sibs is between 0.2 and 0.4. Parent-offspring correlations are typically slightly lower. Segregation analyses have repeatedly suggested a major recessive gene effect. Based on these analyses, sample size calculations have been performed based on both concordant and discordant approaches. In contrast to

the expectations, the concordant sib-pair approach was superior; a lower number of families were required to achieve the same power.

Family studies based on extremely obese young index patients, either mother or father or both, have a BMI > 90th decile in the vast majority of the families. Based on index patients with a BMI > 95th centile, approximately 20% of the respective families have a sib with a BMI > 90th centile.

In conclusion, it is apparent that environmental factors interact with specific genotypes rendering an individual more or less susceptible to the development of obesity. Furthermore, despite the fact that major genes have been detected, it is necessary to consider that the spectrum reaches from such major genes to genes with an only minor influence.

The discovery of the leptin gene at the end of 1994 (Zhang et al., 1994) has been followed by a virtual explosion of scientific efforts to uncover the regulatory systems underlying appetite and weight regulation. It is currently the fastest growing biomedical field. This upswing has also resulted in large scaled molecular genetic activities which, due to obvious clinical interest, are basically all related to obesity in humans, rodents and other mammals (Hebebrand et al., 1998).

In this respect, many genes in which mutations lead to the presently known monogenic forms of obesity have been cloned in rodents. Systemic consequences of these mutations are currently being analysed. These models have provided insights into the complex regulatory systems involved in body weight regulation, the best known of which includes leptin and its receptor.

In mice, but also in pigs, over 15 quantitative trait loci (QTL) have been identified that are most likely relevant in weight regulation (Chagnon et al., 2003).

In humans, four exceedingly rare autosomal recessive forms of obesity have been described as of 1997. Mutations in the genes encoding for leptin, leptin receptor, prohormone convertase 1 and pro-opiomelanocortin (POMC) have been shown to cause

massive obesity of an early onset type, associated with hyperphagia. Distinct additional clinical (e.g. red hair, primary amenorrhea) and/or endocrinological abnormalities (e.g. markedly altered serum leptin levels, lack of ACTH secretion) pinpointed to the respective candidate genes. Both the monogenic animal models and the human monogenic forms have led to new insights into the complex system underlying body weight regulation.

Very recently, the first autosomal dominant form of obesity was described in humans. Two different mutations within the melanocortin-4 receptor gene (*MC4R*) were observed to lead to extreme obesity in probands heterozygous for these variants. In contrast to the aforementioned findings, these mutations do not implicate readily obvious phenotypic abnormalities other than extreme obesity (Vaisse et al., 1998; Yeo et al., 1998). Interestingly, both groups detected the mutations by systematic screens in relatively small study groups (n=63 and n=43).

Hinney et al. (1999) screened the *MC4R* in a total of 492 obese children and adolescents. All in all, four individuals with two different mutations leading to haplo-insufficiency were detected. One was identical to that previously observed by Yeo et al. (1998). The other mutation, which was detected in three individuals, induced a stop mutation in the extracellular domain of the receptor. Approximately one percent of extremely obese individuals harbour haplo-insufficiency mutations in the *MC4R*. In addition to the two forms of haplo-insufficiency, Hinney et al. (1999) also detected additional mutations leading to both conservative and non-conservative amino acid exchanges. Interestingly, these mutations were mainly observed in the obese study group. The functional implications of these mutations are currently unknown.

The identification of individuals with *MC4R* mutations is interesting in the light of possible pharmacological interventions. Thus, intranasal application of adrenocorticotropin$_{4-10}$ (ACTH$_{4-10}$), representing a core sequence of all melanocortins, resulted in reduced weight, body fat mass and plasma leptin concentrations in healthy controls. The question arises as to how mutation carriers would react to this treatment, which could theoretically counterbalance their reduced receptor density.

The involvement of specific genes in weight regulation is further substantiated by data obtained from transgenic mice. For example, MC4R deficient mice develop early onset obesity (Huszar et al., 1997).

Different groups are conducting genome scans related to obesity or dependent phenotypes (BMI, leptin levels, fat mass, etc.). This approach appears very promising, because it is both systematic and model free. In addition, it has already been shown to be exceptionally successful. Thus, positive linkage results have been obtained even by analysing comparatively small study groups. More important, some findings have already been replicated. Each of the following regions has been identified by at least two independent groups: chromosome 1p32, chromosome 2p21, chromosome 6p21, chromosome 10 and chromosome 20q13 (Chagnon et al., 2003).

SUMMARY OF THE INVENTION

The present invention now discloses the identification of a human obesity susceptibility gene, which can be used for the diagnosis, prevention and treatment of obesity and related disorders, as well as for the screening of therapeutically active drugs.

The invention can be used in the diagnosis of predisposition to or protection against, detection, prevention and/or treatment of obesity, coronary heart disease and metabolic disorders, including hypoalphalipoproteinemia, familial combined hyperlipidemia, insulin resistant syndrome X or multiple metabolic disorder, coronary artery disease, diabetes and dyslipidemia, the method comprising detecting in a sample from the subject the presence of an alteration in the TAS1R1 gene or polypeptide, the presence of said alteration being indicative of the presence or predisposition to obesity or associated disorders. The presence of said alteration can also be indicative for protecting against obesity.

A particular object of this invention resides in a method of detecting the presence of or predisposition to obesity or an associated disorder in a subject, the method comprising detecting the presence of an alteration in the TAS1R1 gene locus in a sample from the subject, the presence of said alteration being indicative of the presence of or the predisposition to obesity or an associated disorder.

An additional particular object of this invention resides in a method of detecting the protection from obesity or an associated disorder in a subject, the method comprising detecting the presence of an alteration in the TAS1R1 gene locus in a sample from the subject, the presence of said alteration being indicative of the protection from obesity or an associated disorder.

Another particular object of this invention resides in a method of assessing the response of a subject to a treatment of obesity or an associated disorder, the method comprising detecting the presence of an alteration in the TAS1R1 gene locus in a sample from the subject, the presence of said alteration being indicative of a particular response to said treatment.

A further particular object of this invention resides in a method of assessing the adverse effect in a subject to a treatment of obesity or an associated disorder, the method comprising detecting the presence of an alteration in the TAS1R1 gene locus in a sample from the subject, the presence of said alteration being indicative of an adverse effect to said treatment.

This invention also relates to a method for preventing obesity or an associated disorder in a subject, comprising detecting the presence of an alteration in the TAS1R1 gene locus in a sample from the subject, the presence of said alteration being indicative of the predisposition to obesity or an associated disorder; and, administering a prophylactic treatment against obesity or an associated disorder.

In a preferred embodiment, said alteration is one or several SNP(s) or a haplotype of SNPs associated with obesity. More preferably, said haplotype associated with obesity comprises or consists of several SNPs selected in the group consisting of SNP38, SNP47, SNP49, SNP56 and SNP69. Still more preferably, said haplotype is selected from the haplotypes disclosed in Table 5. More preferably, said SNP associated with obesity can be SNP49.

Preferably, the alteration in the TAS1R1 gene locus is determined by performing a hydridization assay, a sequencing assay, a microsequencing assay, or an allele-specific amplification assay.

A particular aspect of this invention resides in compositions of matter comprising primers, probes, and/or oligonucleotides, which are designed to specifically detect at least one SNP or haplotype associated with obesity in the genomic region including the TAS1R1 gene, or a combination thereof. In a preferred embodiment, said alteration is one or several SNP(s) or a haplotype of SNPs associated with obesity. More preferably, said haplotype associated with obesity comprises or consists of several SNPs selected in the group consisting of SNP38, SNP47, SNP49, SNP56 and SNP69. Still more preferably, said haplotype is selected from the haplotypes disclosed in Table 5. More preferably, said SNP associated with obesity can be SNP49.

The invention also resides in methods of treating obesity and/or associated disorders in a subject through a modulation of TAS1R1 expression or activity. Such treatments use, for instance, TAS1R1 polypeptides, TAS1R1 DNA sequences (including antisense sequences and RNAi directed at the TAS1R1 gene locus), anti- TAS1R1 antibodies or drugs that modulate TAS1R1 expression or activity.

The invention also relates to methods of treating individuals who carry deleterious alleles of the TAS1R1 gene, including pre-symptomatic treatment or combined therapy, such as through gene therapy, protein replacement therapy or through the administration of TAS1R1 protein mimetics and/or inhibitors.

A further aspect of this invention resides in the screening of drugs for therapy of obesity or associated disorder, based on the modulation of or binding to an allele of TAS1R1 gene associated with obesity or associated disorder or gene product thereof.

A further aspect of this invention includes antibodies specific of TAS1R1 polypeptide fragments and derivatives of such antibodies, hybridomas secreting such antibodies, and diagnostic kits comprising those antibodies. More preferably, said antibodies are specific to a TAS1R1 polypeptide or a fragment thereof comprising an alteration, said alteration modifying the activity of TAS1R1.

The invention also concerns a TAS1R1 gene or a fragment thereof comprising an alteration, said alteration modifying the activity of TAS1R1. The invention further

concerns a TAS1R1 polypeptide or a fragment thereof comprising an alteration, said alteration modifying the activity of TAS1R1.

LEGEND TO THE FIGURES

Figure 1 : High density mapping using Genomic Hybrid Identity Profiling (GenomeHIP)

A total of 2263 BAC clones with an average spacing of 1.2 Mega base pairs between clones representing the whole human genome were tested for linkage using GenomeHIP. Each point on the x-axis corresponds to a clone. Several clones are indicated by their library name for better orientation (e.g. BACA17ZF07).

Highly significant evidence for linkage was calculated for clones BACA17ZF07 (p-value $8.0x10^{-11}$) and BACA15ZD05 (p-value $3.8x10^{-10}$). Significant evidence for linkage was calculated for clones BACA13ZH10 and BACA13ZH11 (p-value $2.5x10^{6}$ and $2.1x10^{-7}$, respectively). The whole linkage region is encompassing a region starting from 4126987 base pairs to 7007690 base pairs on human chromosome 1.

The p-value $2x10^{-5}$ corresponding to the significance level for significant linkage was used as a significance level for whole genome screens as proposed by Lander and Kruglyak (1995).

DETAILED DESCRIPTION OF THE INVENTION

The present invention discloses the identification of TAS1R1 as a human obesity susceptibility gene. Various nucleic acid samples from 164 families with obesity were submitted to a particular GenomeHIP process. This process led to the identification of particular identical-by-descent fragments in said populations that are altered in obese subjects. By screening of the IBD fragments, we identified the taste receptor, type1, member 1 on chromosome 1p36.3 (TAS1R1) gene as a candidate for obesity and related phenotypes. This gene is indeed present in the critical interval and expresses a functional phenotype consistent with a genetic regulation of obesity ASNP of the TAS1R1 gene was also identified, as being correlated to obesity in human subjects. SNPs of the TAS1R1 gene were also identified, as being correlated to obesity in human subjects. SNP49 located in the TAS1R1 gene was found to be associated with obesity. Haplotypes disclosed in Table 5 comprising several SNPs selected in the group consisting of SNP38, SNP47, SNP49, SNP56 and SNP69 have also been identified as associated with obesity.

The present invention thus proposes to use TAS1R1 gene and corresponding expression products for the diagnosis, prevention and treatment of obesity and associated disorders, as well as for the screening of therapeutically active drugs.

DEFINITIONS

Obesity and metabolic disorders: Obesity shall be construed as any condition of abnormal or excessive fat accumulation in adipose tissue, to the extent that health may be impaired. Associated disorders, diseases or pathologies include, more specifically, any metabolic disorders, including hypo-alphalipoproteinemia, familial combined hyperlipidemia, insulin resistant syndrome X or multiple metabolic disorder, coronary artery disease, diabetes mellitus and dyslipidemia. The invention may be used in various subjects, particularly human, including adults, children and at the prenatal stage.

Within the context of this invention, the TAS1R1 gene locus designates all TAS1R1 sequences or products in a cell or organism, including TAS1R1 coding sequences, TAS1R1 non-coding sequences (e.g., introns), TAS1R1 regulatory sequences controlling transcription and/or translation (e.g., promoter, enhancer, terminator, etc.), as well as all corresponding expression products, such as TAS1R1 RNAs (e.g., mRNAs) and TAS1R1 polypeptides (e.g., a pre-protein and a mature protein). The TAS1R1 gene locus also comprise surrounding sequences of the TAS1R1 gene which include SNPs that are in linkage disequilibrium with SNPs located in the TAS1R1 gene. For example, the TAS1R1 locus comprises surrounding sequences comprising SNP38, SNP47, SNP56 and SNP69.

As used in the present application, the term "TAS1R1 gene" designates the taste receptor, type 1, member 1 gene on human chromosome 1p36.3, as well as variants, analogs and fragments thereof, including alleles thereof (e.g., germline mutations) which are related to susceptibility to obesity and metabolic disorders. The TAS1R1 gene may also be referred to as GPR 70, T1R1, TR1 and gm 148.

The term "gene" shall be construed to include any type of coding nucleic acid, including genomic DNA (gDNA), complementary DNA (cDNA), synthetic or semi-synthetic DNA, as well as any form of corresponding RNA. The term gene particularly includes recombinant nucleic acids encoding TAS1R1, i.e., any non naturally occurring nucleic acid molecule created artificially, e.g., by assembling, cutting, ligating or amplifying sequences.

A TAS1R1 gene is typically double-stranded, although other forms may be contemplated, such as single-stranded. TAS1R1 genes may be obtained from various sources and according to various techniques known in the art, such as by screening DNA libraries or by amplification from various natural sources. Recombinant nucleic acids may be prepared by conventional techniques, including chemical synthesis, genetic engineering, enzymatic techniques, or a combination thereof. Suitable TAS1R1 gene sequences may be found on gene banks, such as Unigene Cluster for TAS1R1 (Hs. 124574) and Unigene Representative Sequences NM_138697, NM_177539, NM_177540, and NM_177541. Particular examples of a TAS1R1 gene comprise SEQ ID Nos: 1, 3, 5, and 7.

The term "TAS1R1 gene" includes any variant, fragment or analog of SEQ ID Nos: 1, 3, 5, and 7 or of any coding sequence as identified above. Such variants include, for instance, naturally-occurring variants due to allelic variations between individuals (e.g., polymorphisms), mutated alleles related to obesity, alternative splicing forms, etc. The term variant also includes TAS1R1 gene sequences from other sources or organisms. Variants are preferably substantially homologous to SEQ ID Nos: 1, 3, 5, and 7, i.e., exhibit a nucleotide sequence identity of at least about 65%, typically at least about 75%, preferably at least about 85%, more preferably at least about 95% with SEQ ID No:1. Variants and analogs of a TAS1R1 gene also include nucleic acid sequences, which hybridize to a sequence as defined above (or a complementary strand thereof) under stringent hybridization conditions.

Typical stringent hybridisation conditions include temperatures above 30° C, preferably above 35°C, more preferably in excess of 42°C, and/or salinity of less than about 500 mM, preferably less than 200 mM. Hybridization conditions may be adjusted by the skilled person by modifying the temperature, salinity and/or the concentration of other reagents such as SDS, SSC, etc.

A fragment of a TAS1R1 gene designates any portion of at least about 8 consecutive nucleotides of a sequence as disclosed above, preferably at least about 15, more preferably at least about 20 nucleotides, further preferably of at least 30 nucleotides. Fragments include all possible nucleotide length between 8 and 100 nucleotides, preferably between 15 and 100, more preferably between 20 and 100.

A TAS1R1 polypeptide designates any protein or polypeptide encoded by a TAS1R1 gene as disclosed above. The term "polypeptide" refers to any molecule comprising a stretch of amino acids. This term includes molecules of various length, such as peptides and proteins. The polypeptide may be modified, such as by glycosylations and/or acetylations and/or chemical reaction or coupling, and may contain one or several non-natural or synthetic amino acids. A specific example of a TAS1R1 polypeptide comprises all or part of SEQ ID Nos:2, 4, 6 and 8 or a variant thereof.

The terms "response to a treatment" refer to treatment efficacy, including but not limited to ability to metabolize a therapeutic compound, to the ability to convert a pro-drug to an active drug, and to the pharmacokinetics (absorption, distribution, elimination) and the pharmacodynamics (receptor-related) of a drug in an individual.

The terms "adverse effects to a treatment" refer to adverse effects of therapy resulting from extensions of the principal pharmacological action of the drug or to idiosyncratic adverse reactions resulting from an interaction of the drug with unique host factors. "Side effects to a treatment" include, but are not limited to, adverse reactions such as dermatologic, hematologic or hepatologic toxicities and further includes gastric and intestinal ulceration, disturbance in platelet function, renal injury, generalized urticaria, bronchoconstriction, hypotension, and shock.

DIAGNOSIS

The invention now provides diagnosis methods based on a monitoring of the TAS1R1 gene locus in a subject. Within the context of the present invention, the term 'diagnosis" includes the detection, monitoring, dosing, comparison, etc., at various stages, including early, pre-symptomatic stages, and late stages, in adults, children and pre-birth. Diagnosis typically includes the prognosis, the assessment of a predisposition or risk of development, the characterization of a subject to define most appropriate treatment (pharmaco-genetics), etc.

The present invention provides diagnostic methods to determine whether an individual is at risk of developing obesity or an associated disorder or suffers from obesity or an associated disorder resulting from a mutation or a polymorphism in the TAS1R1 gene locus. The present invention also provides methods to determine whether an individual is likely to

respond positively to a therapeutic agent or whether an individual is at risk of developing an adverse side effect to a therapeutic agent.

A particular object of this invention resides in a method of detecting the presence of or predisposition to obesity or an associated disorder in a subject, the method comprising detecting in a sample from the subject the presence of an alteration in the TAS1R1 gene locus in said sample. The presence of said alteration is indicative of the presence or predisposition to obesity or an associated disorder. Optionally, said method comprises a previous step of providing a sample from a subject. Preferably, the presence of an alteration in the TAS1R1 gene locus in said sample is detected through the genotyping of a sample.

Another particular object of this invention resides in a method of detecting the protection from obesity or an associated disorder in a subject, the method comprising detecting the presence of an alteration in the TAS1R1 gene locus in a sample from the subject, the presence of said alteration being indicative of the protection from obesity or an associated disorder.

In a preferred embodiment, said alteration is one or several SNP(s) or a haplotype of SNPs associated with obesity. More preferably, said haplotype associated with obesity comprises or consists of several SNPs selected in the group consisting of SNP38, SNP47, SNP49, SNP56 and SNP69. Still more preferably, said haplotype is selected from the haplotypes disclosed in Table 5. More preferably, said SNP associated with obesity can be SNP49.

Another particular object of this invention resides in a method of assessing the response of a subject to a treatment of obesity or an associated disorder, the method comprising (i) providing a sample from the subject and (ii) detecting the presence of an alteration in the TAS1R1 gene locus in said sample.

Another particular object of this invention resides in a method of assessing the response of a subject to a treatment of obesity or an associated disorder, the method comprising detecting in a sample from the subject the presence of an alteration in the TAS1R1 gene locus in said sample. The presence of said alteration is indicative of a particular response to

said treatment. Preferably, the presence of an alteration in the TAS1R1 gene locus in said sample is detected through the genotyping of a sample.

A further particular object of this invention resides in a method of assessing the adverse effects of a subject to a treatment of obesity or an associated disorder, the method comprising detecting in a sample from the subject the presence of an alteration in the TAS1R1 gene locus in said sample. The presence of said alteration is indicative of adverse effects to said treatment. Preferably, the presence of an alteration in the TAS1R1 gene locus in said sample is detected through the genotyping of a sample.

In a preferred embodiment, said alteration is one or several SNP(s) or a haplotype of SNPs associated with obesity. More preferably, said haplotype associated with obesity comprises or consists of several SNPs selected in the group consisting of SNP38, SNP47, SNP49, SNP56 and SNP69. Still more preferably, said haplotype is selected from the haplotypes disclosed in Table 5. More preferably, said SNP associated with obesity can be SNP49.

In an additional embodiment, the invention concerns a method for preventing obesity or an associated disorder in a subject, comprising detecting the presence of an alteration in the TAS1R1 gene locus in a sample from the subject, the presence of said alteration being indicative of the predisposition to obesity or an associated disorder; and, administering a prophylactic treatment against obesity or an associated disorder. Said prophylactic treatment can be an administration of a drug and/or a diet.

Diagnostics, which analyze and predict response to a treatment or drug, or side effects to a treatment or drug, may be used to determine whether an individual should be treated with a particular treatment drug. For example, if the diagnostic indicates a likelihood that an individual will respond positively to treatment with a particular drug, the drug may be administered to the individual. Conversely, if the diagnostic indicates that an individual is likely to respond negatively to treatment with a particular drug, an alternative course of treatment may be prescribed. A negative response may be defined as either the absence of an efficacious response or the presence of toxic side effects.

Clinical drug trials represent another application for the TAS1R1 SNPs. One or more TAS1R1 SNPs indicative of response to a drug or to side effects to a drug may be

identified using the methods described above. Thereafter, potential participants in clinical trials of such an agent may be screened to identify those individuals most likely to respond favorably to the drug and exclude those likely to experience side effects. In that way, the effectiveness of drug treatment may be measured in individuals who respond positively to the drug, without lowering the measurement as a result of the inclusion of individuals who are unlikely to respond positively in the study and without risking undesirable safety problems.

The alteration may be determined at the level of the TAS1R1 gDNA, RNA or polypeptide. Optionally, the detection is performed by sequencing all or part of the TAS1R1 gene or by selective hybridisation or amplification of all or part of the TAS1R1 gene. More preferably a TAS1R1 gene specific amplification is carried out before the alteration identification step.

An alteration in the TAS1R1 gene locus may be any form of mutation(s), deletion(s), rearrangement(s) and/or insertions in the coding and/or non-coding region of the locus, alone or in various combination(s). Mutations more specifically include point mutations. Deletions may encompass any region of two or more residues in a coding or non-coding portion of the gene locus, such as from two residues up to the entire gene or locus. Typical deletions affect smaller regions, such as domains (introns) or repeated sequences or fragments of less than about 50 consecutive base pairs, although larger deletions may occur as well. Insertions may encompass the addition of one or several residues in a coding or non-coding portion of the gene locus. Insertions may typically comprise an addition of between 1 and 50 base pairs in the gene locus. Rearrangement includes inversion of sequences. The TAS1R1 gene locus alteration may result in the creation of stop codons, frameshift mutations, amino acid substitutions, particular RNA splicing or processing, product instability, truncated polypeptide production, etc. The alteration may result in the production of a TAS1R1 polypeptide with altered function, stability, targeting or structure. The alteration may also cause a reduction in protein expression or, alternatively, an increase in said production.

In a particular embodiment of the method according to the present invention, the alteration in the TAS1R1 gene locus is selected from a point mutation, a deletion and an insertion in the TAS1R1 gene or corresponding expression product, more preferably a point mutation

and a deletion. The alteration may be determined at the level of the TAS1R1 gDNA, RNA or polypeptide.

In this regard, the present invention now discloses a SNP in the TAS1R1 gene and haplotypes characterised by this SNP which are associated with obesity. The point mutations (or single nucleotide alterations are reported in the following table 2.

This point mutation has been detected in subjects having obesity. This mutation was tested for association with obesity and obtained results of this test showed that the mutation is correlated with obesity (Tables 4 and 5).

In a first variant, the method of the present invention comprises detecting the presence of an altered TAS1R1 gene sequence. This can be performed by sequencing all or part of the TAS1R1 gene, polypeptide or RNA, by selective hybridisation or by selective amplification, for instance.

A more specific embodiment comprises detecting the presence of a SNP as disclosed in Table 2 in the TAS1R1 gene sequence of a subject, particularly SNP49.

Table 2

Nucleotide position in genomic sequence of chromosome 1 based on NCBI Build 34	SNP identity	dbSNP reference	Polymorphism	Position in locus	Sequence reference
6028263	SNP38	rs3747977	G/A	3' of TAS1R1 locus	SEQ ID No 9
6272613	SNP47	rs3007434	G/A	3' of TAS1R1 locus	SEQ ID No 10
6354269	SNP49	rs731024	A/G	Intron of TAS1R1 gene	SEQ ID No 11
6432730	SNP56	rs149857	G/C	3' of TAS1R1 locus	SEQ ID No 12
7112026	SNP69	rs705681	T/C	3' of TAS1R1 locus	SEQ ID No 13

In another variant, the method comprises detecting the presence of an altered TAS1R1 RNA expression. Altered RNA expression includes the presence of an altered RNA sequence, the presence of an altered RNA splicing or processing, the presence of an altered quantity of RNA, etc. These may be detected by various techniques known in the art, including by sequencing all or part of the TAS1R1 RNA or by selective hybridisation or selective amplification of all or part of said RNA, for instance.

In a further variant, the method comprises detecting the presence of an altered TAS1R1 polypeptide expression. Altered TAS1R1 polypeptide expression includes the presence of an altered polypeptide sequence, the presence of an altered quantity of TAS1R1 polypeptide, the presence of an altered tissue distribution, etc. These may be detected by various techniques known in the art, including by sequencing and/or binding to specific ligands (such as antibodies), for instance.

As indicated above, various techniques known in the art may be used to detect or quantify altered TAS1R1 gene or RNA expression or sequence, including sequencing, hybridisation, amplification and/or binding to specific ligands (such as antibodies). Other suitable methods include allele-specific oligonucleotide (ASO), allele-specific amplification, Southern blot (for DNAs), Northern blot (for RNAs), single-stranded conformation analysis (SSCA), PFGE, fluorescent in situ hybridization (FISH), gel migration, clamped denaturing gel electrophoresis, heteroduplex analysis, RNase protection, chemical mismatch cleavage, ELISA, radio-immunoassays (RIA) and immuno-enzymatic assays (IEMA).

Some of these approaches (e.g., SSCA and CGGE) are based on a change in electrophoretic mobility of the nucleic acids, as a result of the presence of an altered sequence. According to these techniques, the altered sequence is visualized by a shift in mobility on gels. The fragments may then be sequenced to confirm the alteration.

Some others are based on specific hybridisation between nucleic acids from the subject and a probe specific for wild-type or altered TAS1R1 gene or RNA. The probe may be in suspension or immobilized on a substrate. The probe is typically labelled to facilitate detection of hybrids.

Some of these approaches are particularly suited for assessing a polypeptide sequence or expression level, such as Northern blot, ELISA and RIA. These latter require the use of a ligand specific for the polypeptide, more preferably of a specific antibody.

In a particular, preferred, embodiment, the method comprises detecting the presence of an altered TAS1R1 gene expression profile in a sample from the subject. As indicated above, this can be accomplished more preferably by sequencing, selective hybridisation and/or selective amplification of nucleic acids present in said sample.

Sequencing

Sequencing can be carried out using techniques well known in the art, using automatic sequencers. The sequencing may be performed on the complete TAS1R1 gene or, more preferably, on specific domains thereof, typically those known or suspected to carry deleterious mutations or other alterations.

Amplification

Amplification is based on the formation of specific hybrids between complementary nucleic acid sequences that serve to initiate nucleic acid reproduction.

Amplification may be performed according to various techniques known in the art, such as by polymerase chain reaction (PCR), ligase chain reaction (LCR), strand displacement amplification (SDA) and nucleic acid sequence based amplification (NASBA). These techniques can be performed using commercially available reagents and protocols. Preferred techniques use allele-specific PCR or PCR-SSCP. Amplification usually requires the use of specific nucleic acid primers, to initiate the reaction.

Nucleic acid primers useful for amplifying sequences from the TAS1R1 gene or locus are able to specifically hybridize with a portion of the TAS1R1 gene locus that flank a target region of said locus, said target region being altered in certain subjects having obesity or associated disorders. Examples of such target regions are provided in Table 2.

Primers that can be used to amplify TAS1R1 target region comprising other SNPs as identified in Table 2 may be designed based on the sequence of SEQ ID NO: 1, 3, 5 or 7 or on the genomic sequence of TAS1R1.

Another particular object of this invention resides in a nucleic acid primer useful for amplifying sequences from the TAS1R1 gene or locus including surrounding regions. Such primers are preferably complementary to, and hybridize specifically to nucleic acid sequences in the TAS1R1 gene locus. Particular primers are able to specifically hybridise with a portion of the TAS1R1 gene locus that flank a target region of said locus, said target region being altered in certain subjects having obesity or associated disorders.

The invention also relates to a nucleic acid primer, said primer being complementary to and hybridizing specifically to a portion of a TAS1R1 coding sequence (e.g., gene or RNA) altered in certain subjects having obesity or associated disorders. In this regard, particular primers of this invention are specific for altered sequences in a TAS1R1 gene or RNA. By using such primers, the detection of an amplification product indicates the presence of an alteration in the TAS1R1 gene locus. In contrast, the absence of amplification product indicates that the specific alteration is not present in the sample.

Typical primers of this invention are single-stranded nucleic acid molecules of about 5 to 60 nucleotides in length, more preferably of about 8 to about 25 nucleotides in length. The sequence can be derived directly from the sequence of the TAS1R1 gene locus. Perfect complementarity is preferred, to ensure high specificity. However, certain mismatch may be tolerated.

The invention also concerns the use of a nucleic acid primer or a pair of nucleic acid primers as described above in a method of detecting the presence of or predisposition to obesity or an associated disorder in a subject or in a method of assessing the response of a subject to a treatment of obesity or an associated disorder.

Selective hybridization

Hybridization detection methods are based on the formation of specific hybrids between complementary nucleic acid sequences that serve to detect nucleic acid sequence alteration(s).

A particular detection technique involves the use of a nucleic acid probe specific for wild-type or altered TAS1R1 gene or RNA, followed by the detection of the presence of a hybrid. The probe may be in suspension or immobilized on a substrate or support (as in nucleic acid array or chips technologies). The probe is typically labelled to facilitate detection of hybrids.

In this regard, a particular embodiment of this invention comprises contacting the sample from the subject with a nucleic acid probe specific for an altered TAS1R1 gene locus, and assessing the formation of an hybrid. In a particular, preferred embodiment, the method comprises contacting simultaneously the sample with a set of probes that are specific, respectively, for wild type TAS1R1 gene locus and for various altered forms thereof. In this embodiment, it is possible to detect directly the presence of various forms of alterations in the TAS1R1 gene locus in the sample. Also, various samples from various subjects may be treated in parallel.

Within the context of this invention, a probe refers to a polynucleotide sequence which is complementary to and capable of specific hybridisation with a (target portion of a) TAS1R1 gene or RNA, and which is suitable for detecting polynucleotide polymorphisms associated with TAS1R1 alleles which predispose to or are associated with obesity or metabolic disorders. Probes are preferably perfectly complementary to the TAS1R1 gene, RNA, or target portion thereof. Probes typically comprise single-stranded nucleic acids of between 8 to 1000 nucleotides in length, for instance of between 10 and 800, more preferably of between 15 and 700, typically of between 20 and 500. It should be understood that longer probes may be used as well. A preferred probe of this invention is a single stranded nucleic acid molecule of between 8 to 500 nucleotides in length, which can specifically hybridise to a region of a TAS1R1 gene or RNA that carries an alteration.

A specific embodiment of this invention is a nucleic acid probe specific for an altered (e.g., a mutated) TAS1R1 gene or RNA, i.e., a nucleic acid probe that specifically hybridises to said altered TAS1R1 gene or RNA and essentially does not hybridise to a TAS1R1 gene or RNA lacking said alteration. Specificity indicates that hybridisation to the target sequence generates a specific signal which can be distinguished from the signal generated through non-specific hybridisation. Perfectly complementary sequences are preferred to design

probes according to this invention. It should be understood, however, that certain mismatch may be tolerated, as long as the specific signal may be distinguished from non-specific hybridisation.

Particular examples of such probes are nucleic acid sequences complementary to a target portion of the genomic region including the TAS1R1 gene or RNA carrying a point mutation as listed in Table 2 above. More particularly, the probes can comprise a sequence selected from the group consisting of SEQ ID No 9-13 or a fragment thereof comprising the SNP or a complementary sequence thereof.

The sequence of the probes can be derived from the sequences of the TAS1R1 gene and RNA as provided in the present application. Nucleotide substitutions may be performed, as well as chemical modifications of the probe. Such chemical modifications may be accomplished to increase the stability of hybrids (e.g., intercalating groups) or to label the probe. Typical examples of labels include, without limitation, radioactivity, fluorescence, luminescence, enzymatic labelling, etc.

The invention also concerns the use of a nucleic acid probe as described above in a method of detecting the presence of or predisposition to obesity or an associated disorder in a subject or in a method of assessing the response of a subject to a treatment of obesity or an associated disorder.

Specific Ligand Binding

As indicated above, alteration in the TAS1R1 gene locus may also be detected by screening for alteration(s) in TAS1R1 polypeptide sequence or expression levels. In this regard, a specific embodiment of this invention comprises contacting the sample with a ligand specific for a TAS1R1 polypeptide and determining the formation of a complex.

Different types of ligands may be used, such as specific antibodies. In a specific embodiment, the sample is contacted with an antibody specific for a TAS1R1 polypeptide and the formation of an immune complex is determined. Various methods for detecting an immune complex can be used, such as ELISA, radio-immunoassays (RIA) and immuno-enzymatic assays (IEMA).

Within the context of this invention, an antibody designates a polyclonal antibody, a monoclonal antibody, as well as fragments or derivatives thereof having substantially the same antigen specificity. Fragments include Fab, Fab'2, CDR regions, etc. Derivatives include single-chain antibodies, humanized antibodies, poly-functional antibodies, etc.

An antibody specific for a TAS1R1 polypeptide designates an antibody that selectively binds a TAS1R1 polypeptide, i.e., an antibody raised against a TAS1R1 polypeptide or an epitope-containing fragment thereof. Although non-specific binding towards other antigens may occur, binding to the target TAS1R1 polypeptide occurs with a higher affinity and can be reliably discriminated from non-specific binding.

In a specific embodiment, the method comprises contacting a sample from the subject with (a support coated with) an antibody specific for an altered form of a TAS1R1 polypeptide, and determining the presence of an immune complex. In a particular embodiment, the sample may be contacted simultaneously, or in parallel, or sequentially, with various (supports coated with) antibodies specific for different forms of a TAS1R1 polypeptide, such as a wild-type and various altered forms thereof.

The invention also concerns the use of a ligand, preferably an antibody, a fragment or a derivative thereof as described above, in a method of detecting the presence of or predisposition to obesity or associated disorders in a subject or in a method of assessing the response of a subject to a treatment of obesity or associated disorders.

The invention also relates to a diagnostic kit comprising products and reagents for detecting in a sample from a subject the presence of an alteration in the TAS1R1 gene or polypeptide, in the TAS1R1 gene or polypeptide expression, and/or in TAS1R1 activity. Said diagnostic kit according to the present invention comprises any primer, any pair of primers, any nucleic acid probe and/or any ligand, preferably antibody, described in the present invention. Said diagnostic kit according to the present invention can further comprise reagents and/or protocols for performing a hybridization, amplification or antigen-antibody immune reaction.

The diagnosis methods can be performed in vitro, ex vivo or in vivo, preferably in vitro or ex vivo. They use a sample from the subject, to assess the status of the TAS1R1 gene

locus. The sample may be any biological sample derived from a subject, which contains nucleic acids or polypeptides. Examples of such samples include fluids, tissues, cell samples, organs, biopsies, etc. Most preferred samples are blood, plasma, saliva, urine, seminal fluid, etc. Pre-natal diagnosis may also be performed by testing foetal cells or placental cells, for instance The sample may be collected according to conventional techniques and used directly for diagnosis or stored. The sample may be treated prior to performing the method, in order to render or improve availability of nucleic acids or polypeptides for testing. Treatments include, for instant, lysis (e.g., mechanical, physical, chemical, etc.), centrifugation, etc. Also, the nucleic acids and/or polypeptides may be pre-purified or enriched by conventional techniques, and/or reduced in complexity. Nucleic acids and polypeptides may also be treated with enzymes or other chemical or physical treatments to produce fragments thereof. Considering the high sensitivity of the claimed methods, very few amounts of sample are sufficient to perform the assay.

As indicated, the sample is preferably contacted with reagents such as probes, primers or ligands in order to assess the presence of an altered TAS1R1 gene locus. Contacting may be performed in any suitable device, such as a plate, tube, well, glass, etc. In specific embodiments, the contacting is performed on a substrate coated with the reagent, such as a nucleic acid array or a specific ligand array. The substrate may be a solid or semi-solid substrate such as any support comprising glass, plastic, nylon, paper, metal, polymers and the like. The substrate may be of various forms and sizes, such as a slide, a membrane, a bead, a column, a gel, etc. The contacting may be made under any condition suitable for a complex to be formed between the reagent and the nucleic acids or polypeptides of the sample.

The finding of an altered TAS1R1 polypeptide, RNA or DNA in the sample is indicative of the presence of an altered TAS1R1 gene locus in the subject, which can be correlated to the presence, predisposition or stage of progression of obesity or metabolic disorders. For example, an individual having a germline TAS1R1 mutation has an increased risk of developing obesity or metabolic disorders. The determination of the presence of an altered TAS1R1 gene locus in a subject also allows the design of appropriate therapeutic intervention, which is more effective and customized. Also, this determination at the pre-symptomatic level allows a preventive regimen to be applied.

Linkage Disequilibirum

Once a first SNP has been identified in a genomic region of interest, more particularly in TAS1R1 gene locus, the practitioner of ordinary skill in the art can easily identify additional SNPs in linkage disequilibrium with this first SNP. Indeed, any SNP in linkage disequilibrium with a first SNP associated with obesity or an associated metabolic disorder will be associated with this trait. Therefore, once the association has been demonstrated between a given SNP and obesity or an associated metabolic disorder, the discovery of additional SNPs associated with this trait can be of great interest in order to increase the density of SNPs in this particular region.

Identification of additional SNPs in linkage disequilibrium with a given SNP involves: (a) amplifying a fragment from the genomic region comprising or surrounding a first SNP from a plurality of individuals; (b) identifying of second SNPs in the genomic region harboring or surrounding said first SNP; (c) conducting a linkage disequilibrium analysis between said first SNP and second SNPs; and (d) selecting said second SNPs as being in linkage disequilibrium with said first marker. Subcombinations comprising steps (b) and (c) are also contemplated.

Methods to identify SNPs and to conduct linkage disequilibrium analysis can be carried out by the skilled person without undue experimentation by using well-known methods.

These SNPs in linkage disequilibrium can also be used in the methods according to the present invention, and more particularly in the diagnosic methods according to the present invention.

For example, a linkage locus of Crohn's disease has been mapped to a large region spanning 18cM on chromosome 5q31 (Rioux et al., 2000 and 2001). Using dense maps of microsatellite markers and SNPs across the entire region, strong evidence of linkage disequilibrium (LD) was found. Having found evidence of LD, the authors developed an ultra-high-density SNP map and studied a denser collection of markers selected from this map. Multilocus analyses defined a single common risk haplotype characterised by multiple SNPs that were each independently associated using TDT. These SNPs were unique to the risk haplotype and essentially identical in their information content by virtue of being in nearly complete LD with one another. The equivalent properties of these SNPs

make it impossible to identify the causal mutation within this region on the basis of genetic evidence alone.

Causal Mutation

Mutations in the TAS1R1 gene which are responsible for obesity or an associated metabolic disorder may be identified by comparing the sequences of the TAS1R1 gene from patients presenting obesity or an associated metabolic disorder and control individuals. Based on the identified association of SNPs of TAS1R1 and obesity or an associated metabolic disorder, the identified locus can be scanned for mutations. In a preferred embodiment, functional regions such as exons and splice sites, promoters and other regulatory regions of the TAS1R1 gene are scanned for mutations. Preferably, patients presenting obesity or an associated metabolic disorder carry the mutation shown to be associated with obesity or an associated metabolic disorder and controls individuals do not carry the mutation or allele associated with obesity or an associated metabolic disorder. It might also be possible that patients presenting obesity or an associated metabolic disorder carry the mutation shown to be associated with obesity or an associated metabolic disorder with a higher frequency than controls individuals.

The method used to detect such mutations generally comprises the following steps: amplification of a region of the TAS1R1 gene comprising a SNP or a group of SNPs associated with obesity or an associated metabolic disorder from DNA samples of the TAS1R1 gene from patients presenting obesity or an associated metabolic disorder and control individuals; sequencing of the amplified region; comparison of DNA sequences of the TAS1R1 gene from patients presenting obesity or an associated metabolic disorder and control individuals; determination of mutations specific to patients presenting obesity or an associated metabolic disorder.

Therefore, identification of a causal mutation in the TAS1R1 gene can be carried out by the skilled person without undue experimentation by using well-known methods.

For example, the causal mutations have been identified in the following examples by using routine methods.

Hugot et al. (2001) applied a positional cloning strategy to identify gene variants with susceptibly to Crohn's disease in a region of chromosome 16 previously found to be linked to susceptibility to Crohn's disease. To refine the location of the potential sucecptibility locus 26 microsatellite markers were genotyped and tested for association to Crohn's disease using the transmission disequilibrium test. A borderline significant association was found between one allele of the microsatellite marker D16S136. Eleven additional SNPs were selected from surrounding regions and several SNPs showed significant association. SNP5-8 from this region were found to be present in a single exon of the NOD2/CARD15 gene and shown to be non-synonymous variants. This prompted the authors to sequence the complete coding sequence of this gene in 50 CD patients. Two additional non-synonymous mutations (SNP12 and SNP13) were found. SNP13 was most significant associated (p=6x10-6) using the pedigree transmission disequilibrium test. In another independent study, the same variant was found also by sequencing the coding region of this gene from 12 affected individuals compared to 4 controls (Ogura et al., 2001). The rare allele of SNP13 corresponded to a 1-bp insertion predicted to truncate the NOD2/CARD15 protein. This allele was also present in normal healthy individuals, albeit with significantly lower frequency as compared to the controls.

Similarly, Lesage et al. (2002) performed a mutational analyses of CARD15 in 453 patients with CD, including 166 sporadic and 287 familial cases, 159 patients with ulcerative colitis (UC), and 103 healthy control subjects by systematic sequencing of the coding region. Of 67 sequence variations identified, 9 had an allele frequency >5% in patients with CD. Six of them were considered to be polymorphisms, and three (SNP12-R702W, SNP8-G908R, and SNP13-1007fs) were confirmed to be independently associated with susceptibility to CD. Also considered as potential disease-causing mutations (DCMs) were 27 rare additional mutations. The three main variants (R702W, G908R, and 1007fs) represented 32%, 18%, and 31%, respectively, of the total CD mutations, whereas the total of the 27 rare mutations represented 19% of DCMs. Altogether, 93% of the mutations were located in the distal third of the gene. No mutations were found to be associated with UC. In contrast, 50% of patients with CD carried at least one DCM, including 17% who had a double mutation.

DRUG SCREENING

The present invention also provides novel targets and methods for the screening of drug candidates or leads. The methods include binding assays and/or functional assays, and may be performed in vitro, in cell systems, in animals, etc.

A particular object of this invention resides in a method of selecting biologically active compounds, said method comprising contacting in vitro a test compound with a TAS1R1 gene or polypeptide according to the present invention and determining the ability of said test compound to bind said TAS1R1 gene or polypeptide. Binding to said gene or polypeptide provides an indication as to the ability of the compound to modulate the activity of said target, and thus to affect a pathway leading to obesity or metabolic disorders in a subject. In a preferred embodiment, the method comprises contacting in vitro a test compound with a TAS1R1 polypeptide or a fragment thereof according to the present invention and determining the ability of said test compound to bind said TAS1R1 polypeptide or fragment. The fragment preferably comprises a binding site of the TAS1R1 polypeptide. Preferably, said TAS1R1 gene or polypeptide or a fragment thereof is an altered or mutated TAS1R1 gene or polypeptide or a fragment thereof comprising the alteration or mutation.

A particular object of this invention resides in a method of selecting compounds active on obesity and associated disorders, said method comprising contacting in vitro a test compound with a TAS1R1 polypeptide according to the present invention or binding site-containing fragment thereof and determining the ability of said test compound to bind said TAS1R1 polypeptide or fragment thereof. Preferably, said TAS1R1 polypeptide or a fragment thereof is an altered or mutated TAS1R1 polypeptide or a fragment thereof comprising the alteration or mutation.

In a further particular embodiment, the method comprises contacting a recombinant host cell expressing a TAS1R1 polypeptide according to the present invention with a test compound, and determining the ability of said test compound to bind said TAS1R1 and to modulate the activity of TAS1R1 polypeptide. Preferably, said TAS1R1 polypeptide or a fragment thereof is an altered or mutated TAS1R1 polypeptide or a fragment thereof comprising the alteration or mutation.

The determination of binding may be performed by various techniques, such as by labelling of the test compound, by competition with a labelled reference ligand, etc.

A further object of this invention resides in a method of selecting biologically active compounds, said method comprising contacting in vitro a test compound with a TAS1R1 polypeptide according to the present invention and determining the ability of said test compound to modulate the activity of said TAS1R1 polypeptide. Preferably, said TAS1R1 polypeptide or a fragment thereof is an altered or mutated TAS1R1 polypeptide or a fragment thereof comprising the alteration or mutation.

A further object of this invention resides in a method of selecting biologically active compounds, said method comprising contacting in vitro a test compound with a TAS1R1 gene according to the present invention and determining the ability of said test compound to modulate the expression of said TAS1R1 gene. Preferably, said TAS1R1 gene or a fragment thereof is an altered or mutated TAS1R1 gene or a fragment thereof comprising the alteration or mutation.

In an other embodiment, this invention relates to a method of screening, selecting or identifying active compounds, particularly compounds active on obesity or metabolic disorders, the method comprising contacting a test compound with a recombinant host cell comprising a reporter construct, said reporter construct comprising a reporter gene under the control of a TAS1R1 gene promoter, and selecting the test compounds that modulate (e.g. stimulate or reduce) expression of the reporter gene. Preferably, said TAS1R1 gene promoter or a fragment thereof is an altered or mutated TAS1R1 gene promoter or a fragment thereof comprising the alteration or mutation.

In a particular embodiment of the methods of screening, the modulation is an inhibition. In an other particular embodiment of the methods of screening, the modulation is an activation.

The above screening assays may be performed in any suitable device, such as plates, tubes, dishes, flasks, etc. Typically, the assay is performed in multi-wells plates. Several test compounds can be assayed in parallel. Furthermore, the test compound may be of various origin, nature and composition. It may be any organic or inorganic substance, such as a

lipid, peptide, polypeptide, nucleic acid, small molecule, etc., in isolated or in mixture with other substances. The compounds may be all or part of a combinatorial library of products, for instance.

PHARMACEUTICAL COMPOSITIONS, THERAPY

A further object of this invention is a pharmaceutical composition comprising (i) a TAS1R1 polypeptide or a fragment thereof, a nucleic acid encoding a TAS1R1 polypeptide or a fragment thereof, a vector or a recombinant host cell as described above and (ii) a pharmaceutically acceptable carrier or vehicle.

The invention also relates to a method of treating or preventing obesity or an associated disorder in a subject, the method comprising administering to said subject a functional (e.g., wild-type) TAS1R1 polypeptide or a nucleic acid encoding the same.

An other embodiment of this invention resides in a method of treating or preventing obesity or an associated disorder in a subject, the method comprising administering to said subject a compound that modulates, preferably that activates or mimics, expression or activity of a TAS1R1 gene or protein according to the present invention. Said compound can be an agonist or an antagonist of TAS1R1, an antisense or a RNAi of TAS1R1, an antibody or a fragment or a derivative thereof specific to a TAS1R1 polypeptide according to the present invention. In a particular embodiment of the method, the modulation is an inhibition. In an other particular embodiment of the method, the modulation is an activation.

The invention also relates, generally, to the use of a functional TAS1R1 polypeptide, a nucleic acid encoding the same, or a compound that modulates expression or activity of a TAS1R1 gene or protein according to the present invention, in the manufacture of a pharmaceutical composition for treating or preventing obesity or an associated metabolic disorder in a subject. Said compound can be an agonist or an antagonist of TAS1R1, an antisense or a RNAi of TAS1R1, an antibody or a fragment or a derivative thereof specific to a TAS1R1 polypeptide according to the present invention. In a particular embodiment of the method, the modulation is an inhibition. In an other particular embodiment of the method, the modulation is an activation.

The present invention demonstrates the correlation between obesity (and related disorders) and the TAS1R1 gene locus. The invention thus provides a novel target of therapeutic intervention. Various approaches can be contemplated to restore or modulate the TAS1R1 activity or function in a subject, particularly those carrying an altered TAS1R1 gene locus. Supplying wild-type function to such subjects is expected to suppress phenotypic expression of obesity and associated disorders in a pathological cell or organism. The supply of such function can be accomplished through gene or protein therapy, or by administering compounds that modulate or mimic TAS1R1 polypeptide activity (e.g., agonists as identified in the above screening assays).

The wild-type TAS1R1 gene or a functional part thereof may be introduced into the cells of the subject in need thereof using a vector as described above. The vector may be a viral vector or a plasmid. The gene may also be introduced as naked DNA. The gene may be provided so as to integrate into the genome of the recipient host' cells, or to remain extra-chromosomal. Integration may occur randomly or at precisely defined sites, such as through homologous recombination. In particular, a functional copy of the TAS1R1 gene may be inserted in replacement of an altered version in a cell, through homologous recombination. Further techniques include gene gun, liposome-mediated transfection, cationic lipid-mediated transfection, etc. Gene therapy may be accomplished by direct gene injection, or by administering ex vivo prepared genetically modified cells expressing a functional TAS1R1 polypeptide.

Other molecules with TAS1R1 activity (e.g., peptides, drugs, TAS1R1 agonists, or organic compounds) may also be used to restore functional TAS1R1 activity in a subject or to suppress the deleterious phenotype in a cell.

Restoration of functional TAS1R1 gene function in a cell may be used to prevent the development of obesity or metabolic disorders or to reduce progression of said diseases. Such a treatment may suppress the obese phenotype of a cell, particularly those cells carrying a deleterious allele.

Further aspects and advantages of the present invention will be disclosed in the following experimental section, which should be regarded as illustrative and not limiting the scope of the present application.

GENE, VECTORS, RECOMBINANT CELLS AND POLYPEPTIDES

A further aspect of this invention resides in novel products for use in diagnosis, therapy or screening. These products comprise nucleic acid molecules encoding a TAS1R1 polypeptide or a fragment thereof, vectors comprising the same, recombinant host cells and expressed polypeptides.

More particularly, the invention concerns an altered or mutated TAS1R1 gene or a fragment thereof comprising said alteration or mutation. The invention also concerns nucleic acid molecules encoding an altered or mutated TAS1R1 polypeptide or a fragment thereof comprising said alteration or mutation. Said alteration or mutation modifies the TAS1R1 activity. The modified activity can be increased or decreased. The invention further concerns a vector comprising an altered or mutated TAS1R1 gene or a fragment thereof comprising said alteration or mutation or a nucleic acid molecule encoding an altered or mutated TAS1R1 polypeptide or a fragment thereof comprising said alteration or mutation, recombinant host cells and expressed polypeptides.

A further object of this invention is a vector comprising a nucleic acid encoding a TAS1R1 polypeptide according to the present invention. The vector may be a cloning vector or, more preferably, an expression vector, i.e., a vector comprising regulatory sequences causing expression of a TAS1R1 polypeptide from said vector in a competent host cell.

These vectors can be used to express a TAS1R1 polypeptide in vitro, ex vivo or in vivo, to create transgenic or "Knock Out" non-human animals, to amplify the nucleic acids, to express antisense RNAs, etc.

The vectors of this invention typically comprise a TAS1R1 coding sequence according to the present invention operably linked to regulatory sequences, e.g., a promoter, a polyA, etc. The term "operably linked" indicates that the coding and regulatory sequences are functionally associated so that the regulatory sequences cause expression (e.g., transcription) of the coding sequences. The vectors may further comprise one or several origins of replication and/or selectable markers. The promoter region may be homologous or heterologous with respect to the coding sequence, and provide for ubiquitous, constitutive, regulated and/or tissue specific expression, in any appropriate host cell,

including for in vivo use. Examples of promoters include bacterial promoters (T7, pTAC, Trp promoter, etc.), viral promoters (LTR, TK, CMV-IE, etc.), mammalian gene promoters (albumin, PGK, etc), and the like.

The vector may be a plasmid, a virus, a cosmid, a phage, a BAC, a YAC, etc. Plasmid vectors may be prepared from commercially available vectors such as pBluescript, pUC, pBR, etc. Viral vectors may be produced from baculoviruses, retroviruses, adenoviruses, AAVs, etc., according to recombinant DNA techniques known in the art.

In this regard, a particular object of this invention resides in a recombinant virus encoding a TAS1R1 polypeptide as defined above. The recombinant virus is preferably replication-defective, even more preferably selected from E1- and/or E4-defective adenoviruses, Gag-, pol- and/or env-defective retroviruses and Rep- and/or Cap-defective AAVs. Such recombinant viruses may be produced by techniques known in the art, such as by transfecting packaging cells or by transient transfection with helper plasmids or viruses. Typical examples of virus packaging cells include PA317 cells, PsiCRIP cells, GPenv+ cells, 293 cells, etc. Detailed protocols for producing such replication-defective recombinant viruses may be found for instance in WO95/14785, WO96/22378, US5,882,877, US6,013,516, US4,861,719, US5,278,056 and WO94/19478.

A further object of the present invention resides in a recombinant host cell comprising a recombinant TAS1R1 gene or a vector as defined above. Suitable host cells include, without limitation, prokaryotic cells (such as bacteria) and eukaryotic cells (such as yeast cells, mammalian cells, insect cells, plant cells, etc.). Specific examples include E.coli, Kluyveromyces or Saccharomyces yeasts, mammalian cell lines (e.g., Vero cells, CHO cells, 3T3 cells, COS cells, etc.) as well as primary or established mammalian cell cultures (e.g., produced from fibroblasts, embryonic cells, epithelial cells, nervous cells, adipocytes, etc.).

The present invention also relates to a method for producing a recombinant host cell expressing a TAS1R1 polypeptide according to the present invention, said method comprising (i) introducing in vitro or ex vivo into a competent host cell a recombinant nucleic acid or a vector as described above, (ii) culturing in vitro or ex vivo the

recombinant host cells obtained and (iii), optionally, selecting the cells which express the TAS1R1 polypeptide.

Such recombinant host cells can be used for the production of TAS1R1 polypeptides, as well as for screening of active molecules, as described below. Such cells may also be used as a model system to study obesity and metabolic disorders. These cells can be maintained in suitable culture media, such as DMEM, RPMI, HAM, etc., in any appropriate culture device (plate, flask, dish, tube, pouch, etc.).

EXAMPLES

1. Identification of an Obesity susceptibility locus on human chromosome 1

A. GenomeHIP platform to identify the chromosome 1 susceptibility gene

The GenomeHIP platform was applied to allow rapid identification of an obesity susceptibility gene.

Briefly, the technology consists of forming pairs from the DNA of related individuals. Each DNA is marked with a specific label allowing its identification. Hybrids are then formed between the two DNAs. A particular process (WO00/53802) is then applied that selects all fragments identical-by-descent (IBD) from the two DNAs in a multi step procedure. The remaining IBD enriched DNA is then scored against a BAC clone derived DNA microarray that allows the positioning of the IBD fraction on a chromosome.

The application of this process over many different families results in a matrix of IBD fractions for each pair from each family. Statistical analyses then calculate the minimal IBD regions that are shared between all families tested. Significant results (p-values) are evidence for linkage of the positive region with the trait of interest (here obesity). The linked interval can be delimited by the two most distant clones showing significant p-values.

In the present study, 164 families of German origin (178 independent sib-pairs) concordant for massive obesity (as defined by a body mass index > 90th%ile) were submitted to the GenomeHIP process. The resulting IBD enriched DNA fractions were then labelled with Cy5 fluorescent dyes and hybridised against a DNA array consisting of 2263 BAC clones covering the whole human genome with an average spacing of 1.2 Mega base pairs. Non-

selected DNA labelled with Cy3 was used to normalize the signal values and compute ratios for each clone. Clustering of the ratio results was then performed to determine the IBD status for each clone and pair.

By applying this procedure, several BAC clones (BACA13ZH10, BACA17ZF07 and BACA15ZD05) spanning approximately 3 Mega bases in the region on chromosome 1 (bases 4126987 to 7007690) were identified, that showed significant evidence for linkage to obesity ($p< 2.5x10^{-5}$).

Table 3: Linkage results for chromosome 1 in the TAS1R1 gene locus: Indicated is the region correspondent to 3 BAC clones with evidence for linkage. The start and stop positions of the clones correspond to their genomic location based on NCBI Build34 sequence respective to the start of the chromosome (p-ter).

Table 3

Human chromosome	Clone	Start	End	Proportion of informative pairs	p-value
1	BACA13ZH10		4 126 987	0.92	2.50E-06
1	BACA17ZF07	6 589 325	6 686 208	0.94	8.00E-11
1	BACA15ZD05	6 817 039	7 007 690	0.93	3.80E-10

B. Identification of an obesity susceptibility gene on chromosome 1

By screening the aforementioned 3 Mega bases in the linked chromosomal region, we identified the taste receptor, type 1, member 1 (TAS1R1) gene as a candidate for obesity and related phenotypes. This gene is indeed present in the critical interval, with evidence for linkage delimited by the clones outlined above.

TAS1R1 gene encodes a predicted 841 -amino acid polypeptide for NP_619642 (mRNA NM_138697 2.7 kb) and spreads over 8.84 kb of genomic sequence. The protein encoded by the gene is a G protein-coupled receptor and is a component of the heterodimeric amino acid taste receptor T1R1+3. The T1R1+3 receptor responds to L-amino acids but not to D-enantiomers or other compounds. Most amino acids that are perceived as sweet activate T1R1+3, and this activation is strictly dependent on an intact T1R1+3 heterodimer.

Multiple transcript variants encoding several different isoforms have been found for this gene.

Nelson et al. (2002) showed that cells expressing human TAS1R1 are more than an order of magnitude more sensitive to glutamate (umami) than to other amino acids. Glutamate is besides sweet the main attractive taste modality in humans. Characteristic taste-enhancing effects arise from the presene of purine 5-ribonucleotides such as IMP and GMP.

Kurihara and Kashiwayanagi (2000) compared the taste of agonists for brain glutamate receptors in humans. The order of intensity of umami taste induced by a mixture of 0.5 mmol/L GMP and 1.5 mmol/L of various agonists was glutamate > ibotenate > L(+)-2-amino-4-phosphonobutyric acid (L-AP4) = (+/-)1-aminocyclopentane-trans-1,3-dicarboxylic acid (ACPD). Kainate, N-methyl-D-aspartic acid (NMDA) and (RS)-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA), which are agonists for ionotropic receptors, had no umami taste..

Ozeck et al. (2004) showed that taste receptors can functionally couple to Galpha(i/o) proteins to transmit intracellular signals. Sweeteners and monosodium glutamate induce phosphorylation of ERK1/2 and inhibit cAMP accumulation in HEK293 cells expressing the human sweet T1R(2)/T1R(3) receptor and the human umami T1R(1)/T1R(3) receptor, respectively. The effects of these taste modalities are also prevented by treatment with pertussis toxin.

Taken together, the linkage results provided in the present application, identifying the human TAS1R1 gene in the critical interval of genetic alterations linked to obesity on chromosome 1, with its involvement in the perception of tasting umami and other amino acids, we conclude that alterations (e.g., mutations and/or polymorphisms) in the TAS1R1 gene or its regulatory sequences may contribute to the development of human obesity and represent a novel target for diagnosis or therapeutic intervention.

2. Association study

The same families that have been used for the linkage study were also used to test for association between a sepcific phenotype (here obesity) in question and the genetic marker allele using transmission disequilibrium test (TDT). The TDT is a powerful association test

as it is insensitive to population stratification problems in the tested sample. Briefly, the segregation of alleles from heterozygous parents to their affected off-spring is tested. The portion of alleles transmitted to the affected offspring compared to the non-transmitted alleles is compared to the ratio expected under random distribution. A significant excess of allele transmission over the expected value is evidence for an association of the respective allele with the studied obesity phenotype.

The results of this analysis show that a certain allele of the TAS1R1 gene is positively associated with obesity and might therefore increase the susceptibility to disease. In the tested population, the allele G is correlated to obesity (Chi2 = 4.74, p = 0.029456) as determined by TDT. On the other hand, the allele A is significantly under-transmitted to obese individuals indicating that this allele might help protecting from the disease

An example of the transmission of the alleles to obese individuals is given in Table 4.

Table 4

Allele	Allele transmitted to obese individuals (n)	Allele not transmitted to obese individuals (n)	p- value
G	124	92	0.029456
A	92	124	0.029456

In addition, haplotypes were constructed for SNP34, SNP47, SNP49, SNP56 and SNP69 to identify the phase for all SNPs.

The results of this analysis show that certain haplotypes, all characterized by the presence of allele G at SNP49 are strongly associated with obesity. In the tested population, haplotypes having allele G at SNP49 are strongly correlated to obesity (p=0.003045 for G-A-A of SNP34-SNP47-SNP49, p=0.00588 for A-A-C of SNP34-SNP42-SNP51, p=0.003045 for A-G-G of SNP34- SNP47-SNP49 and p=0.005235 for G-C-C of SNP34-SNP53-SNP69, respectively, as determined by TDT), while certain haplotypes devoid of allele G are preferentially not transmitted to obese subjects (p=0.02026 of G-A-A for SNP34-SNP47-SNP49 and p=0.001073 for A-G-C of SNP49-SNP56-SNP69, respectively). Haplotypes that carry allele A instead of allele G at SNP49 show significant evidence to be under-represented in obese subjects.

Examples of haplotypes with preferential transmission and non-transmission of SNP49 to obese individuals are given in Table 5.

Table 5

SNPs used to construct haplotype	Haplotype	Frequency of haplotype transmitted to obese individuals	Frequency of haplotype not transmitted to obese individuals	Odds Ratio	P-value
SNP34-SNP47-**SNP49**	G-A-**A**	0.01247	0.0341	0.4484	0.02026
SNP34-SNP47-**SNP49**	A-G-**G**	0.2869	0.1773	1.985	0.003045
SNP49-SNP56-SNP69	**A**-G-C	0.06779	0.1664	0.2536	0.001073

REFERENCES

Chagnon YC, Rankinen T, Snyder EE et al. (2003) The human obesity gene map: The 2002 update. Obes Res. 11(3):313-367.

Hebebrand J, Hescker H, Himmelmann GW, Schäfer H, Remschmidt H (1994) Percentiles for the body mass index based on data of the German national nutrition survey and a review of relevant factors with an influence on body weight. Aktuelle Ernahrungsmedizin 19: 259-265.

Hebebrand J, Himmelmann GW, Hescker H, Schäfer H, Remschmidt H (1996) Use of percentiles for the body mass index in anorexia nervosa : diagnostic, epidemiological and therapeutic considerations. Int J Eat Dis 19: 359-369.

Hebebrand J, Hinney A, Roth H et al. (1998) Genetische Aspekte der Adipositas. In:J Wechsler (ed) *Adipositas*. Ex Libris Roche. Blackwell Verlag: 105-118.

Hinney A, Schmidt A, Nottebom K, et al. (1999) Several mutations in the melanocortin-4 receptor gene including a nonsense and a frameshift mutation associated with dominantly inherited obesity in humans. J. Clin Endocrinol Metab 84(4):1483-1486.

Hugot JP, Chamaillard M, Zouali H et al. (2001) Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature 411(6837):599-603.

2005254805 09 Jan 2009

Huszar D, Lynch CA, Fairchild-Huntress V, et al. (1997) Targeted disruption of the melanocortin-4 receptor results in obesity in mice. Cell 88:131-141.

Kurihara K, and Kashiwayanagi M (2000) Physiological studies on umami taste. J Nutr 130(4S Suppl):931S-4S.

Lander E and Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet, 11(3):241-247.

Lesage S5 Zouali H, Cezard Jpet al. (2002) CARD15/NOD2 mutational analysis and genotype-phenotype correlation in 612 patients with inflammatory bowel disease. Am J Hum Genet. 70(4):845-857.

Nelson G, Chandrashekar J5 Hoon MA et al (2002) An amino-acid taste receptor. Nature 416(6877): 199-202.

Ozeck M5 Brust P, Xu H5 et al. (2004) Receptors for bitter, sweet and umami taste couple to inhibitory G protein signaling pathways. Eur J Pharmacol 489(3):139-49.

Rioux JD, Daly MJ5 Silverberg MS et al. (2001) Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. Nat Genet 29(2): 223-228.

Rioux JD5 Silverberg MS5 Daly MJ (2000) Genomewide search in Canadian families with inflammatory bowel disease reveals two novel susceptibility loci. Am J Hum Genet 66(6): 1863-1870.

Schneider R (1996) Relevanz und Kosten der Adipositas in Deutschland. Ernahrungs-Umschau 43:369-374.

Vaisse C5 Clement K5 Guy-Grand B et al (1998) A frameshift mutation in human MC4R is associated with a dominant form of obesity. Nat Genet 20:113-114.

WHO (1998) Preventing and managing the global epidemic. WHO5 Geneva.

Wolf AM, Colditz GA (1996) Social and economic effects of body weight in the United States. Am J Clin Nutr 63(3 Suppl):466S-469S.

Yeo GS5 Farooqi IS, Aminian S5 et al (1998) A frameshift mutation in MC4R associated with dominantly inherited human obesity. Nat Genet 20:111-112.

Zhang Y, Proenca R5 Maffei M5 Barone M, Leopold L5 Friedman JM. (1994) Positional cloning of the mouse obese gene and its human homologue. Nature 372: 425-432.

Throughout this specification and the claims which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" and "comprising", will

2005254805 09 Jan 2009

be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any integer or step or group of integers or steps.

The reference in this specification to any prior publication (or information derived from it), or to any matter which is known, is not, and should not be taken as acknowledgement or admission or any form of suggestion that that prior to publication (or information derived from it) or known matter forms part of the common general knowledge in the field of endeavour to which this specification relates.

2005254805 09 Jan 2009

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. An *in vitro* method of detecting the presence of or predisposition to obesity in a subject, the method comprising detecting the presence of an alteration in the TAS1R1 gene locus in a sample of said subject.

2. The method of claim 1, wherein the presence of an alteration in the TAS1R1 gene locus is detected by sequencing, selective hybridisation and/or selective amplification.

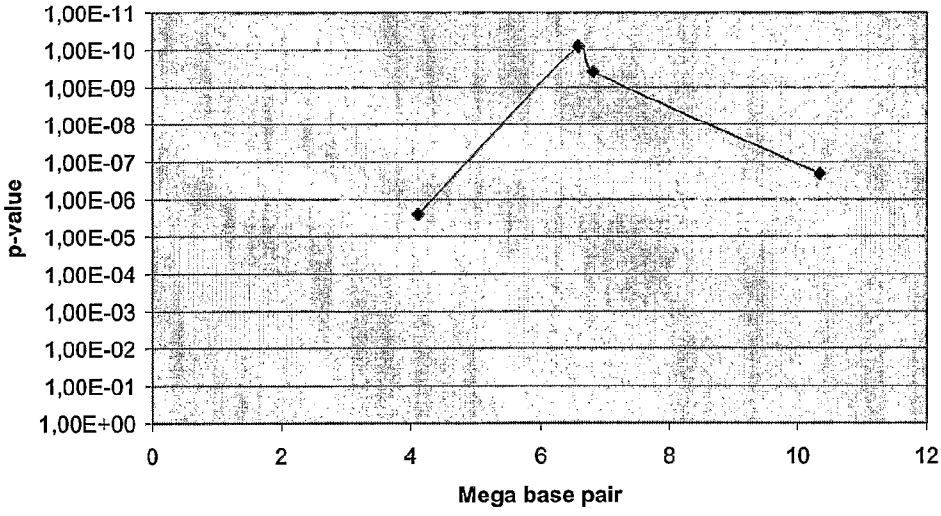
3. The method of claim 1 or claim 2, wherein said alteration is one or several SNP(s) or a haplotype of SNPs associated with obesity.

4. The method of claim 3, wherein said haplotype associated with obesity comprises several SNPs selected in the group consisting of SNP38, SNP47, SNP49, SNP56 and SNP69.

5. The method of claim 3 or claim 4, wherein said SNP associated with obesity is SNP49.

6. The method of claim 1, substantially as before herein described with reference to the Examples.

Figure 1


1,00E-11
1,00E-10
1,00E-09
1,00E-08
1,00E-07
1,00E-06
1,00E-05
1,00E-04
1,00E-03
1,00E-02
1,00E-01
1,00E+00
p-value
0
2
4
6
8
10
12
Mega base pair

SEQUENCE LISTING

<110> INTEGRAGEN

<120> HUMAN OBESITY SUSCEPTIBILITY GENE ENCODING A TASTE RECEPTOR AND USES THEREOF

<130> B0293WO

<150> US 60/578,830
<151> 2004-06-14

<160> 13

<170> PatentIn version 3.1

<210> 1

<211> 2707

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(2526)

<223>

<400> 1

```
atg ctg ctc tgc acg gct cgc ctg gtc ggc ctg cag ctt ctc att tcc       48
Met Leu Leu Cys Thr Ala Arg Leu Val Gly Leu Gln Leu Leu Ile Ser
1               5                   10                  15

tgc tgc tgg gcc ttt gcc tgc cat agc acg gag tct tct cct gac ttc       96
Cys Cys Trp Ala Phe Ala Cys His Ser Thr Glu Ser Ser Pro Asp Phe
            20                  25                  30

acc ctc ccc gga gat tac ctc ctg gca ggc ctg ttc cct ctc cat tct      144
Thr Leu Pro Gly Asp Tyr Leu Leu Ala Gly Leu Phe Pro Leu His Ser
        35                  40                  45

ggc tgt ctg cag gtg agg cac aga ccc gag gtg acc ctg tgt gac agg      192
Gly Cys Leu Gln Val Arg His Arg Pro Glu Val Thr Leu Cys Asp Arg
    50                  55                  60

tct tgt agc ttc aat gag cat ggc tac cac ctc ttc cag gct atg cgg      240
Ser Cys Ser Phe Asn Glu His Gly Tyr His Leu Phe Gln Ala Met Arg
65                  70                  75                  80
```

```
ctt ggg gtt gag gag ata aac aac tcc acg gcc ctg ctg ccc aac atc      288
Leu Gly Val Glu Glu Ile Asn Asn Ser Thr Ala Leu Leu Pro Asn Ile
                85                  90                  95

acc ctg ggg tac cag ctg tat gat gtg tgt tct gac tct gcc aat gtg      336
Thr Leu Gly Tyr Gln Leu Tyr Asp Val Cys Ser Asp Ser Ala Asn Val
            100                 105                 110

tat gcc acg ctg aga gtg ctc tcc ctg cca ggg caa cac cac ata gag      384
Tyr Ala Thr Leu Arg Val Leu Ser Leu Pro Gly Gln His His Ile Glu
        115                 120                 125

ctc caa gga gac ctt ctc cac tat tcc cct acg gtg ctg gca gtg att      432
Leu Gln Gly Asp Leu Leu His Tyr Ser Pro Thr Val Leu Ala Val Ile
    130                 135                 140

ggg cct gac agc acc aac cgt gct gcc acc aca gcc gcc ctg ctg agc      480
Gly Pro Asp Ser Thr Asn Arg Ala Ala Thr Thr Ala Ala Leu Leu Ser
145                 150                 155                 160

cct ttc ctg gtg ccc atg att agc tat gcg gcc agc agc gag acg ctc      528
Pro Phe Leu Val Pro Met Ile Ser Tyr Ala Ala Ser Ser Glu Thr Leu
                165                 170                 175

agc gtg aag cgg cag tat ccc tct ttc ctg cgc acc atc ccc aat gac      576
Ser Val Lys Arg Gln Tyr Pro Ser Phe Leu Arg Thr Ile Pro Asn Asp
            180                 185                 190

aag tac cag gtg gag acc atg gtg ctg ctg ctg cag aag ttc ggg tgg      624
Lys Tyr Gln Val Glu Thr Met Val Leu Leu Leu Gln Lys Phe Gly Trp
        195                 200                 205

acc tgg atc tct ctg gtt ggc agc agt gac gac tat ggg cag cta ggg      672
Thr Trp Ile Ser Leu Val Gly Ser Ser Asp Asp Tyr Gly Gln Leu Gly
    210                 215                 220

gtg cag gca ctg gag aac cag gcc act ggt cag ggg atc tgc att gct      720
Val Gln Ala Leu Glu Asn Gln Ala Thr Gly Gln Gly Ile Cys Ile Ala
225                 230                 235                 240

ttc aag gac atc atg ccc ttc tct gcc cag gtg ggc gat gag agg atg      768
Phe Lys Asp Ile Met Pro Phe Ser Ala Gln Val Gly Asp Glu Arg Met
                245                 250                 255

cag tgc ctc atg cgc cac ctg gcc cag gcc ggg gcc acc gtc gtg gtt      816
Gln Cys Leu Met Arg His Leu Ala Gln Ala Gly Ala Thr Val Val Val
            260                 265                 270

gtt ttt tcc agc cgg cag ttg gcc agg gtg ttt ttc gag tcc gtg gtg      864
Val Phe Ser Ser Arg Gln Leu Ala Arg Val Phe Phe Glu Ser Val Val
        275                 280                 285

ctg acc aac ctg act ggc aag gtg tgg gtc gcc tca gaa gcc tgg gcc      912
Leu Thr Asn Leu Thr Gly Lys Val Trp Val Ala Ser Glu Ala Trp Ala
    290                 295                 300

ctc tcc agg cac atc act ggg gtg ccc ggg atc cag cgc att ggg atg      960
Leu Ser Arg His Ile Thr Gly Val Pro Gly Ile Gln Arg Ile Gly Met
305                 310                 315                 320

gtg ctg ggc gtg gcc atc cag aag agg gct gtc cct ggc ctg aag gcg     1008
Val Leu Gly Val Ala Ile Gln Lys Arg Ala Val Pro Gly Leu Lys Ala
                325                 330                 335

ttt gaa gaa gcc tat gcc cgg gca gac aag aag gcc cct agg cct tgc     1056
Phe Glu Glu Ala Tyr Ala Arg Ala Asp Lys Lys Ala Pro Arg Pro Cys
            340                 345                 350

cac aag ggc tcc tgg tgc agc agc aat cag ctc tgc aga gaa tgc caa     1104
```

```
His Lys Gly Ser Trp Cys Ser Ser Asn Gln Leu Cys Arg Glu Cys Gln
        355                 360                 365

gct ttc atg gca cac acg atg ccc aag ctc aaa gcc ttc tcc atg agt      1152
Ala Phe Met Ala His Thr Met Pro Lys Leu Lys Ala Phe Ser Met Ser
    370                 375                 380

tct gcc tac aac gca tac cgg gct gtg tat gcg gtg gcc cat ggc ctc      1200
Ser Ala Tyr Asn Ala Tyr Arg Ala Val Tyr Ala Val Ala His Gly Leu
385                 390                 395                 400

cac cag ctc ctg ggc tgt gcc tct gga gct tgt tcc agg ggc cga gtc      1248
His Gln Leu Leu Gly Cys Ala Ser Gly Ala Cys Ser Arg Gly Arg Val
                405                 410                 415

tac ccc tgg cag ctt ttg gag cag atc cac aag gtg cat ttc ctt cta      1296
Tyr Pro Trp Gln Leu Leu Glu Gln Ile His Lys Val His Phe Leu Leu
            420                 425                 430

cac aag gac act gtg gcg ttt aat gac aac aga gat ccc ctc agt agc      1344
His Lys Asp Thr Val Ala Phe Asn Asp Asn Arg Asp Pro Leu Ser Ser
        435                 440                 445

tat aac ata att gcc tgg gac tgg aat gga ccc aag tgg acc ttc acg      1392
Tyr Asn Ile Ile Ala Trp Asp Trp Asn Gly Pro Lys Trp Thr Phe Thr
    450                 455                 460

gtc ctc ggt tcc tcc aca tgg tct cca gtt cag cta aac ata aat gag      1440
Val Leu Gly Ser Ser Thr Trp Ser Pro Val Gln Leu Asn Ile Asn Glu
465                 470                 475                 480

acc aaa atc cag tgg cac gga aag gac aac cag gtg cct aag tct gtg      1488
Thr Lys Ile Gln Trp His Gly Lys Asp Asn Gln Val Pro Lys Ser Val
                485                 490                 495

tgt tcc agc gac tgt ctt gaa ggg cac cag cga gtg gtt acg ggt ttc      1536
Cys Ser Ser Asp Cys Leu Glu Gly His Gln Arg Val Val Thr Gly Phe
            500                 505                 510

cat cac tgc tgc ttt gag tgt gtg ccc tgt ggg gct ggg acc ttc ctc      1584
His His Cys Cys Phe Glu Cys Val Pro Cys Gly Ala Gly Thr Phe Leu
        515                 520                 525

aac aag agt gac ctc tac aga tgc cag cct tgt ggg aaa gaa gag tgg      1632
Asn Lys Ser Asp Leu Tyr Arg Cys Gln Pro Cys Gly Lys Glu Glu Trp
    530                 535                 540

gca cct gag gga agc cag acc tgc ttc ccg cgc act gtg gtg ttt ttg      1680
Ala Pro Glu Gly Ser Gln Thr Cys Phe Pro Arg Thr Val Val Phe Leu
545                 550                 555                 560

gct ttg cgt gag cac acc tct tgg gtg ctg ctg gca gct aac acg ctg      1728
Ala Leu Arg Glu His Thr Ser Trp Val Leu Leu Ala Ala Asn Thr Leu
                565                 570                 575

ctg ctg ctg ctg ctg ctt ggg act gct ggc ctg ttt gcc tgg cac cta      1776
Leu Leu Leu Leu Leu Leu Gly Thr Ala Gly Leu Phe Ala Trp His Leu
            580                 585                 590

gac acc cct gtg gtg agg tca gca ggg ggc cgc ctg tgc ttt ctt atg      1824
Asp Thr Pro Val Val Arg Ser Ala Gly Gly Arg Leu Cys Phe Leu Met
        595                 600                 605

ctg ggc tcc ctg gca gca ggt agt ggc agc ctc tat ggc ttc ttt ggg      1872
Leu Gly Ser Leu Ala Ala Gly Ser Gly Ser Leu Tyr Gly Phe Phe Gly
    610                 615                 620

gaa ccc aca agg cct gcg tgc ttg cta cgc cag gcc ctc ttt gcc ctt      1920
Glu Pro Thr Arg Pro Ala Cys Leu Leu Arg Gln Ala Leu Phe Ala Leu
```

625 630 635 640

ggt ttc acc atc ttc ctg tcc tgc ctg aca gtt cgc tca ttc caa cta 1968
Gly Phe Thr Ile Phe Leu Ser Cys Leu Thr Val Arg Ser Phe Gln Leu
645 650 655

atc atc atc ttc aag ttt tcc acc aag gta cct aca ttc tac cac gcc 2016
Ile Ile Ile Phe Lys Phe Ser Thr Lys Val Pro Thr Phe Tyr His Ala
660 665 670

tgg gtc caa aac cac ggt gct ggc ctg ttt gtg atg atc agc tca gcg 2064
Trp Val Gln Asn His Gly Ala Gly Leu Phe Val Met Ile Ser Ser Ala
675 680 685

gcc cag ctg ctt atc tgt cta act tgg ctg gtg gtg tgg acc cca ctg 2112
Ala Gln Leu Leu Ile Cys Leu Thr Trp Leu Val Val Trp Thr Pro Leu
690 695 700

cct gct agg gaa tac cag cgc ttc ccc cat ctg gtg atg ctt gag tgc 2160
Pro Ala Arg Glu Tyr Gln Arg Phe Pro His Leu Val Met Leu Glu Cys
705 710 715 720

aca gag acc aac tcc ctg ggc ttc ata ctg gcc ttc ctc tac aat ggc 2208
Thr Glu Thr Asn Ser Leu Gly Phe Ile Leu Ala Phe Leu Tyr Asn Gly
725 730 735

ctc ctc tcc atc agt gcc ttt gcc tgc agc tac ctg ggt aag gac ttg 2256
Leu Leu Ser Ile Ser Ala Phe Ala Cys Ser Tyr Leu Gly Lys Asp Leu
740 745 750

cca gag aac tac aac gag gcc aaa tgt gtc acc ttc agc ctg ctc ttc 2304
Pro Glu Asn Tyr Asn Glu Ala Lys Cys Val Thr Phe Ser Leu Leu Phe
755 760 765

aac ttc gtg tcc tgg atc gcc ttc ttc acc acg gcc agc gtc tac gac 2352
Asn Phe Val Ser Trp Ile Ala Phe Phe Thr Thr Ala Ser Val Tyr Asp
770 775 780

ggc aag tac ctg cct gcg gcc aac atg atg gct ggg ctg agc agc ctg 2400
Gly Lys Tyr Leu Pro Ala Ala Asn Met Met Ala Gly Leu Ser Ser Leu
785 790 795 800

agc agc ggc ttc ggt ggg tat ttt ctg cct aag tgc tac gtg atc ctc 2448
Ser Ser Gly Phe Gly Gly Tyr Phe Leu Pro Lys Cys Tyr Val Ile Leu
805 810 815

tgc cgc cca gac ctc aac agc aca gag cac ttc cag gcc tcc att cag 2496
Cys Arg Pro Asp Leu Asn Ser Thr Glu His Phe Gln Ala Ser Ile Gln
820 825 830

gac tac acg agg cgc tgc ggc tcc acc tga ccagtgggtc agcaggcacg 2546
Asp Tyr Thr Arg Arg Cys Gly Ser Thr
835 840

gctggcagcc ttctctgccc tgagggtcga aggtcgagca ggccgggggt gtccgggagg 2606

tctttgggca tcgcggtctg gggttgggac gtgtaagcgc ctgggagagc ctagaccagg 2666

ctccgggctg ccaataaaga agtgaaatgc gtaaaaaaaa a 2707

<210> 2

<211> 841

<212> PRT

<213> Homo sapiens

<400> 2

Met Leu Leu Cys Thr Ala Arg Leu Val Gly Leu Gln Leu Leu Ile Ser
1 5 10 15

Cys Cys Trp Ala Phe Ala Cys His Ser Thr Glu Ser Ser Pro Asp Phe
20 25 30

Thr Leu Pro Gly Asp Tyr Leu Leu Ala Gly Leu Phe Pro Leu His Ser
35 40 45

Gly Cys Leu Gln Val Arg His Arg Pro Glu Val Thr Leu Cys Asp Arg
50 55 60

Ser Cys Ser Phe Asn Glu His Gly Tyr His Leu Phe Gln Ala Met Arg
65 70 75 80

Leu Gly Val Glu Glu Ile Asn Asn Ser Thr Ala Leu Leu Pro Asn Ile
85 90 95

Thr Leu Gly Tyr Gln Leu Tyr Asp Val Cys Ser Asp Ser Ala Asn Val
100 105 110

Tyr Ala Thr Leu Arg Val Leu Ser Leu Pro Gly Gln His His Ile Glu
115 120 125

Leu Gln Gly Asp Leu Leu His Tyr Ser Pro Thr Val Leu Ala Val Ile
130 135 140

Gly Pro Asp Ser Thr Asn Arg Ala Ala Thr Thr Ala Ala Leu Leu Ser
145 150 155 160

Pro Phe Leu Val Pro Met Ile Ser Tyr Ala Ala Ser Ser Glu Thr Leu
165 170 175

Ser Val Lys Arg Gln Tyr Pro Ser Phe Leu Arg Thr Ile Pro Asn Asp
180 185 190

Lys Tyr Gln Val Glu Thr Met Val Leu Leu Leu Gln Lys Phe Gly Trp
195 200 205

Thr Trp Ile Ser Leu Val Gly Ser Ser Asp Asp Tyr Gly Gln Leu Gly
210 215 220

Val Gln Ala Leu Glu Asn Gln Ala Thr Gly Gln Gly Ile Cys Ile Ala
225 230 235 240

Phe Lys Asp Ile Met Pro Phe Ser Ala Gln Val Gly Asp Glu Arg Met
245 250 255

Gln Cys Leu Met Arg His Leu Ala Gln Ala Gly Ala Thr Val Val Val

260 265 270

Val Phe Ser Ser Arg Gln Leu Ala Arg Val Phe Phe Glu Ser Val Val
275 280 285

Leu Thr Asn Leu Thr Gly Lys Val Trp Val Ala Ser Glu Ala Trp Ala
290 295 300

Leu Ser Arg His Ile Thr Gly Val Pro Gly Ile Gln Arg Ile Gly Met
305 310 315 320

Val Leu Gly Val Ala Ile Gln Lys Arg Ala Val Pro Gly Leu Lys Ala
325 330 335

Phe Glu Glu Ala Tyr Ala Arg Ala Asp Lys Lys Ala Pro Arg Pro Cys
340 345 350

His Lys Gly Ser Trp Cys Ser Ser Asn Gln Leu Cys Arg Glu Cys Gln
355 360 365

Ala Phe Met Ala His Thr Met Pro Lys Leu Lys Ala Phe Ser Met Ser
370 375 380

Ser Ala Tyr Asn Ala Tyr Arg Ala Val Tyr Ala Val Ala His Gly Leu
385 390 395 400

His Gln Leu Leu Gly Cys Ala Ser Gly Ala Cys Ser Arg Gly Arg Val
405 410 415

Tyr Pro Trp Gln Leu Leu Glu Gln Ile His Lys Val His Phe Leu Leu
420 425 430

His Lys Asp Thr Val Ala Phe Asn Asp Asn Arg Asp Pro Leu Ser Ser
435 440 445

Tyr Asn Ile Ile Ala Trp Asp Trp Asn Gly Pro Lys Trp Thr Phe Thr
450 455 460

Val Leu Gly Ser Ser Thr Trp Ser Pro Val Gln Leu Asn Ile Asn Glu
465 470 475 480

Thr Lys Ile Gln Trp His Gly Lys Asp Asn Gln Val Pro Lys Ser Val
485 490 495

Cys Ser Ser Asp Cys Leu Glu Gly His Gln Arg Val Val Thr Gly Phe
500 505 510

His His Cys Cys Phe Glu Cys Val Pro Cys Gly Ala Gly Thr Phe Leu
515 520 525

Asn Lys Ser Asp Leu Tyr Arg Cys Gln Pro Cys Gly Lys Glu Glu Trp
530 535 540

Ala Pro Glu Gly Ser Gln Thr Cys Phe Pro Arg Thr Val Val Phe Leu
545 550 555 560

Ala Leu Arg Glu His Thr Ser Trp Val Leu Leu Ala Ala Asn Thr Leu
565 570 575

Leu Leu Leu Leu Leu Leu Gly Thr Ala Gly Leu Phe Ala Trp His Leu
580 585 590

Asp Thr Pro Val Val Arg Ser Ala Gly Gly Arg Leu Cys Phe Leu Met
595 600 605

Leu Gly Ser Leu Ala Ala Gly Ser Gly Ser Leu Tyr Gly Phe Phe Gly
610 615 620

Glu Pro Thr Arg Pro Ala Cys Leu Leu Arg Gln Ala Leu Phe Ala Leu
625 630 635 640

Gly Phe Thr Ile Phe Leu Ser Cys Leu Thr Val Arg Ser Phe Gln Leu
645 650 655

Ile Ile Ile Phe Lys Phe Ser Thr Lys Val Pro Thr Phe Tyr His Ala
660 665 670

Trp Val Gln Asn His Gly Ala Gly Leu Phe Val Met Ile Ser Ser Ala
675 680 685

Ala Gln Leu Leu Ile Cys Leu Thr Trp Leu Val Val Trp Thr Pro Leu
690 695 700

Pro Ala Arg Glu Tyr Gln Arg Phe Pro His Leu Val Met Leu Glu Cys
705 710 715 720

Thr Glu Thr Asn Ser Leu Gly Phe Ile Leu Ala Phe Leu Tyr Asn Gly
725 730 735

Leu Leu Ser Ile Ser Ala Phe Ala Cys Ser Tyr Leu Gly Lys Asp Leu
740 745 750

Pro Glu Asn Tyr Asn Glu Ala Lys Cys Val Thr Phe Ser Leu Leu Phe
755 760 765

Asn Phe Val Ser Trp Ile Ala Phe Phe Thr Thr Ala Ser Val Tyr Asp
770 775 780

Gly Lys Tyr Leu Pro Ala Ala Asn Met Met Ala Gly Leu Ser Ser Leu
785 790 795 800

Ser Ser Gly Phe Gly Gly Tyr Phe Leu Pro Lys Cys Tyr Val Ile Leu
805 810 815

Cys Arg Pro Asp Leu Asn Ser Thr Glu His Phe Gln Ala Ser Ile Gln
820 825 830

Asp Tyr Thr Arg Arg Cys Gly Ser Thr
835 840

<210> 3

<211> 2373

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(1443)

<223>

<400> 3

```
atg ctg ctc tgc acg gct cgc ctg gtc ggc ctg cag ctt ctc att tcc       48
Met Leu Leu Cys Thr Ala Arg Leu Val Gly Leu Gln Leu Leu Ile Ser
1               5                   10                  15

tgc tgc tgg gcc ttt gcc tgc cat agc acg gag tct tct cct gac ttc       96
Cys Cys Trp Ala Phe Ala Cys His Ser Thr Glu Ser Ser Pro Asp Phe
            20                  25                  30

acc ctc ccc gga gat tac ctc ctg gca ggc ctg ttc cct ctc cat tct      144
Thr Leu Pro Gly Asp Tyr Leu Leu Ala Gly Leu Phe Pro Leu His Ser
        35                  40                  45

ggc tgt ctg cag gtg agg cac aga ccc gag gtg acc ctg tgt gac agg      192
Gly Cys Leu Gln Val Arg His Arg Pro Glu Val Thr Leu Cys Asp Arg
    50                  55                  60

tct tgt agc ttc aat gag cat ggc tac cac ctc ttc cag gct atg cgg      240
Ser Cys Ser Phe Asn Glu His Gly Tyr His Leu Phe Gln Ala Met Arg
65                  70                  75                  80

ctt ggg gtt gag gag ata aac aac tcc acg gcc ctg ctg ccc aac atc      288
Leu Gly Val Glu Glu Ile Asn Asn Ser Thr Ala Leu Leu Pro Asn Ile
                85                  90                  95

acc ctg ggg tac cag ctg tat gat gtg tgt tct gac tct gcc aat gtg      336
Thr Leu Gly Tyr Gln Leu Tyr Asp Val Cys Ser Asp Ser Ala Asn Val
            100                 105                 110

tat gcc acg ctg aga gtg ctc tcc ctg cca ggg caa cac cac ata gag      384
Tyr Ala Thr Leu Arg Val Leu Ser Leu Pro Gly Gln His His Ile Glu
        115                 120                 125

ctc caa gga gac ctt ctc cac tat tcc cct acg gtg ctg gca gtg att      432
Leu Gln Gly Asp Leu Leu His Tyr Ser Pro Thr Val Leu Ala Val Ile
    130                 135                 140

ggg cct gac agc acc aac cgt gct gcc acc aca gcc gcc ctg ctg agc      480
Gly Pro Asp Ser Thr Asn Arg Ala Ala Thr Thr Ala Ala Leu Leu Ser
145                 150                 155                 160
```

```
cct ttc ctg gtg ccc atg att agc tat gcg gcc agc agc gag acg ctc      528
Pro Phe Leu Val Pro Met Ile Ser Tyr Ala Ala Ser Ser Glu Thr Leu
                165                 170                 175

agc gtg aag cgg cag tat ccc tct ttc ctg cgc acc atc ccc aat gac      576
Ser Val Lys Arg Gln Tyr Pro Ser Phe Leu Arg Thr Ile Pro Asn Asp
            180                 185                 190

aag tac cag gtg gag acc atg gtg ctg ctg ctg cag aag ttc ggg tgg      624
Lys Tyr Gln Val Glu Thr Met Val Leu Leu Leu Gln Lys Phe Gly Trp
        195                 200                 205

acc tgg atc tct ctg gtt ggc agc agt gac gac tat ggg cag cta ggg      672
Thr Trp Ile Ser Leu Val Gly Ser Ser Asp Asp Tyr Gly Gln Leu Gly
    210                 215                 220

gtg cag gca ctg gag aac cag gcc act ggt cag ggg atc tgc att gct      720
Val Gln Ala Leu Glu Asn Gln Ala Thr Gly Gln Gly Ile Cys Ile Ala
225                 230                 235                 240

ttc aag gac atc atg ccc ttc tct gcc cag gtg ggc gat gag agg atg      768
Phe Lys Asp Ile Met Pro Phe Ser Ala Gln Val Gly Asp Glu Arg Met
                245                 250                 255

cag tgc ctc atg cgc cac ctg gcc cag gcc ggg gcc acc gtc gtg gtt      816
Gln Cys Leu Met Arg His Leu Ala Gln Ala Gly Ala Thr Val Val Val
            260                 265                 270

gtt ttt tcc agc cgg cag ttg gcc agg gtg ttt ttc gag tcc gtg gtg      864
Val Phe Ser Ser Arg Gln Leu Ala Arg Val Phe Phe Glu Ser Val Val
        275                 280                 285

ctg acc aac ctg act ggc aag gtg tgg gtc gcc tca gaa gcc tgg gcc      912
Leu Thr Asn Leu Thr Gly Lys Val Trp Val Ala Ser Glu Ala Trp Ala
    290                 295                 300

ctc tcc agg cac atc act ggg gtg ccc ggg atc cag cgc att ggg atg      960
Leu Ser Arg His Ile Thr Gly Val Pro Gly Ile Gln Arg Ile Gly Met
305                 310                 315                 320

gtg ctg ggc gtg gcc atc cag aag agg gct gtc cct ggc ctg aag gcg     1008
Val Leu Gly Val Ala Ile Gln Lys Arg Ala Val Pro Gly Leu Lys Ala
                325                 330                 335

ttt gaa gaa gcc tat gcc cgg gca gac aag aag gcc cct agg cct tgc     1056
Phe Glu Glu Ala Tyr Ala Arg Ala Asp Lys Lys Ala Pro Arg Pro Cys
            340                 345                 350

cac aag ggc tcc tgg tgc agc agc aat cag ctc tgc aga gaa tgc caa     1104
His Lys Gly Ser Trp Cys Ser Ser Asn Gln Leu Cys Arg Glu Cys Gln
        355                 360                 365

gct ttc atg gca cac acg atg ccc aag ctc aaa gcc ttc tcc atg agt     1152
Ala Phe Met Ala His Thr Met Pro Lys Leu Lys Ala Phe Ser Met Ser
    370                 375                 380

tct gcc tac aac gca tac cgg gct gtg tat gcg gtg gcc cat ggc ctc     1200
Ser Ala Tyr Asn Ala Tyr Arg Ala Val Tyr Ala Val Ala His Gly Leu
385                 390                 395                 400

cac cag ctc ctg ggc tgt gcc tct gga gct tgt tcc agg ggc cga gtc     1248
His Gln Leu Leu Gly Cys Ala Ser Gly Ala Cys Ser Arg Gly Arg Val
                405                 410                 415

tac ccc tgg cag acc tct aca gat gcc agc ctt gtg gga aag aag agt     1296
Tyr Pro Trp Gln Thr Ser Thr Asp Ala Ser Leu Val Gly Lys Lys Ser
            420                 425                 430
```

ggg cac ctg agg gaa gcc aga cct gct tcc cgc gca ctg tgg tgt ttt 1344
Gly His Leu Arg Glu Ala Arg Pro Ala Ser Arg Ala Leu Trp Cys Phe
435 440 445

tgg ctt tgc gtg agc aca cct ctt ggg tgc tgc tgg cag cta aca cgc 1392
Trp Leu Cys Val Ser Thr Pro Leu Gly Cys Cys Trp Gln Leu Thr Arg
450 455 460

tgc tgc tgc tgc tgc tgc ttg gga ctg ctg gcc tgt ttg cct ggc acc 1440
Cys Cys Cys Cys Cys Cys Leu Gly Leu Leu Ala Cys Leu Pro Gly Thr
465 470 475 480

tag acacccctgt ggtgaggtca gcaggggcc gcctgtgctt tcttatgctg 1493

ggctccctgg cagcaggtag tggcagcctc tatggcttct ttggggaacc cacaaggcct 1553

gcgtgcttgc tacgccaggc cctctttgcc cttggtttca ccatcttcct gtcctgcctg 1613

acagttcgct cattccaact aatcatcatc ttcaagtttt ccaccaaggt acctacattc 1673

taccacgcct gggtccaaaa ccacggtgct ggcctgtttg tgatgatcag ctcagcggcc 1733

cagctgctta tctgtctaac ttggctggtg gtgtggaccc cactgcctgc tagggaatac 1793

cagcgcttcc cccatctggt gatgcttgag tgcacagaga ccaactccct gggcttcata 1853

ctggccttcc tctacaatgg cctcctctcc atcagtgcct ttgcctgcag ctacctgggt 1913

aaggacttgc cagagaacta caacgaggcc aaatgtgtca ccttcagcct gctcttcaac 1973

ttcgtgtcct ggatcgcctt cttcaccacg gccagcgtct acgacggcaa gtacctgcct 2033

gcggccaaca tgatggctgg gctgagcagc ctgagcagcg gcttcggtgg gtattttctg 2093

cctaagtgct acgtgatcct ctgccgccca gacctcaaca gcacagagca cttccaggcc 2153

tccattcagg actacacgag gcgctgcggc tccacctgac cagtgggtca gcaggcacgg 2213

ctggcagcct tctctgccct gagggtcgaa ggtcgagcag gccgggggtg tccgggaggt 2273

ctttgggcat cgcggtctgg ggttgggacg tgtaagcgcc tgggagagcc tagaccaggc 2333

tccgggctgc caataaagaa gtgaaatgcg taaaaaaaaa 2373

<210> 4

<211> 480

<212> PRT

<213> Homo sapiens

<400> 4

Met Leu Leu Cys Thr Ala Arg Leu Val Gly Leu Gln Leu Leu Ile Ser
1 5 10 15

Cys Cys Trp Ala Phe Ala Cys His Ser Thr Glu Ser Ser Pro Asp Phe
20 25 30

Thr Leu Pro Gly Asp Tyr Leu Leu Ala Gly Leu Phe Pro Leu His Ser
35 40 45

Gly Cys Leu Gln Val Arg His Arg Pro Glu Val Thr Leu Cys Asp Arg
50 55 60

Ser Cys Ser Phe Asn Glu His Gly Tyr His Leu Phe Gln Ala Met Arg
65 70 75 80

Leu Gly Val Glu Glu Ile Asn Asn Ser Thr Ala Leu Leu Pro Asn Ile
85 90 95

Thr Leu Gly Tyr Gln Leu Tyr Asp Val Cys Ser Asp Ser Ala Asn Val
100 105 110

Tyr Ala Thr Leu Arg Val Leu Ser Leu Pro Gly Gln His His Ile Glu
115 120 125

Leu Gln Gly Asp Leu Leu His Tyr Ser Pro Thr Val Leu Ala Val Ile
130 135 140

Gly Pro Asp Ser Thr Asn Arg Ala Ala Thr Thr Ala Ala Leu Leu Ser
145 150 155 160

Pro Phe Leu Val Pro Met Ile Ser Tyr Ala Ala Ser Ser Glu Thr Leu
165 170 175

Ser Val Lys Arg Gln Tyr Pro Ser Phe Leu Arg Thr Ile Pro Asn Asp
180 185 190

Lys Tyr Gln Val Glu Thr Met Val Leu Leu Leu Gln Lys Phe Gly Trp
195 200 205

Thr Trp Ile Ser Leu Val Gly Ser Ser Asp Asp Tyr Gly Gln Leu Gly
210 215 220

Val Gln Ala Leu Glu Asn Gln Ala Thr Gly Gln Gly Ile Cys Ile Ala
225 230 235 240

Phe Lys Asp Ile Met Pro Phe Ser Ala Gln Val Gly Asp Glu Arg Met
245 250 255

Gln Cys Leu Met Arg His Leu Ala Gln Ala Gly Ala Thr Val Val Val
260 265 270

Val Phe Ser Ser Arg Gln Leu Ala Arg Val Phe Phe Glu Ser Val Val
275 280 285

Leu Thr Asn Leu Thr Gly Lys Val Trp Val Ala Ser Glu Ala Trp Ala
290 295 300

Leu Ser Arg His Ile Thr Gly Val Pro Gly Ile Gln Arg Ile Gly Met
305 310 315 320

Val Leu Gly Val Ala Ile Gln Lys Arg Ala Val Pro Gly Leu Lys Ala

325 330 335

Phe Glu Glu Ala Tyr Ala Arg Ala Asp Lys Lys Ala Pro Arg Pro Cys
340 345 350

His Lys Gly Ser Trp Cys Ser Ser Asn Gln Leu Cys Arg Glu Cys Gln
355 360 365

Ala Phe Met Ala His Thr Met Pro Lys Leu Lys Ala Phe Ser Met Ser
370 375 380

Ser Ala Tyr Asn Ala Tyr Arg Ala Val Tyr Ala Val Ala His Gly Leu
385 390 395 400

His Gln Leu Leu Gly Cys Ala Ser Gly Ala Cys Ser Arg Gly Arg Val
405 410 415

Tyr Pro Trp Gln Thr Ser Thr Asp Ala Ser Leu Val Gly Lys Lys Ser
420 425 430

Gly His Leu Arg Glu Ala Arg Pro Ala Ser Arg Ala Leu Trp Cys Phe
435 440 445

Trp Leu Cys Val Ser Thr Pro Leu Gly Cys Cys Trp Gln Leu Thr Arg
450 455 460

Cys Cys Cys Cys Cys Cys Leu Gly Leu Leu Ala Cys Leu Pro Gly Thr
465 470 475 480

<210> 5

<211> 1945

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(1764)

<223>

<400> 5

atg ctg ctc tgc acg gct cgc ctg gtc ggc ctg cag ctt ctc att tcc 48
Met Leu Leu Cys Thr Ala Arg Leu Val Gly Leu Gln Leu Leu Ile Ser
1 5 10 15

tgc tgc tgg gcc ttt gcc tgc cat agc acg gag tct tct cct gac ttc 96
Cys Cys Trp Ala Phe Ala Cys His Ser Thr Glu Ser Ser Pro Asp Phe
20 25 30

acc ctc ccc gga gat tac ctc ctg gca ggc ctg ttc cct ctc cat tct 144

```
Thr Leu Pro Gly Asp Tyr Leu Leu Ala Gly Leu Phe Pro Leu His Ser
        35                  40                  45

ggc tgt ctg cag gtg agg cac aga ccc gag gtg acc ctg tgt gac agg      192
Gly Cys Leu Gln Val Arg His Arg Pro Glu Val Thr Leu Cys Asp Arg
    50                  55                  60

tct tgt agc ttc aat gag cat ggc tac cac ctc ttc cag gct atg cgg      240
Ser Cys Ser Phe Asn Glu His Gly Tyr His Leu Phe Gln Ala Met Arg
65                  70                  75                  80

ctt ggg gtt gag gag ata aac aac tcc acg gcc ctg ctg ccc aac atc      288
Leu Gly Val Glu Glu Ile Asn Asn Ser Thr Ala Leu Leu Pro Asn Ile
                85                  90                  95

acc ctg ggg tac cag ctg tat gat gtg tgt tct gac tct gcc aat gtg      336
Thr Leu Gly Tyr Gln Leu Tyr Asp Val Cys Ser Asp Ser Ala Asn Val
            100                 105                 110

tat gcc acg ctg aga gtg ctc tcc ctg cca ggg caa cac cac ata gag      384
Tyr Ala Thr Leu Arg Val Leu Ser Leu Pro Gly Gln His His Ile Glu
        115                 120                 125

ctc caa gga gac ctt ctc cac tat tcc cct acg gtg ctg gca gtg att      432
Leu Gln Gly Asp Leu Leu His Tyr Ser Pro Thr Val Leu Ala Val Ile
    130                 135                 140

ggg cct gac agc acc aac cgt gct gcc acc aca gcc gcc ctg ctg agc      480
Gly Pro Asp Ser Thr Asn Arg Ala Ala Thr Thr Ala Ala Leu Leu Ser
145                 150                 155                 160

cct ttc ctg gtg ccc atg ctt ttg gag cag atc cac aag gtg cat ttc      528
Pro Phe Leu Val Pro Met Leu Leu Glu Gln Ile His Lys Val His Phe
                165                 170                 175

ctt cta cac aag gac act gtg gcg ttt aat gac aac aga gat ccc ctc      576
Leu Leu His Lys Asp Thr Val Ala Phe Asn Asp Asn Arg Asp Pro Leu
            180                 185                 190

agt agc tat aac ata att gcc tgg gac tgg aat gga ccc aag tgg acc      624
Ser Ser Tyr Asn Ile Ile Ala Trp Asp Trp Asn Gly Pro Lys Trp Thr
        195                 200                 205

ttc acg gtc ctc ggt tcc tcc aca tgg tct cca gtt cag cta aac ata      672
Phe Thr Val Leu Gly Ser Ser Thr Trp Ser Pro Val Gln Leu Asn Ile
    210                 215                 220

aat gag acc aaa atc cag tgg cac gga aag gac aac cag gtg cct aag      720
Asn Glu Thr Lys Ile Gln Trp His Gly Lys Asp Asn Gln Val Pro Lys
225                 230                 235                 240

tct gtg tgt tcc agc gac tgt ctt gaa ggg cac cag cga gtg gtt acg      768
Ser Val Cys Ser Ser Asp Cys Leu Glu Gly His Gln Arg Val Val Thr
                245                 250                 255

ggt ttc cat cac tgc tgc ttt gag tgt gtg ccc tgt ggg gct ggg acc      816
Gly Phe His His Cys Cys Phe Glu Cys Val Pro Cys Gly Ala Gly Thr
            260                 265                 270

ttc ctc aac aag agt gac ctc tac aga tgc cag cct tgt ggg aaa gaa      864
Phe Leu Asn Lys Ser Asp Leu Tyr Arg Cys Gln Pro Cys Gly Lys Glu
        275                 280                 285

gag tgg gca cct gag gga agc cag acc tgc ttc ccg cgc act gtg gtg      912
Glu Trp Ala Pro Glu Gly Ser Gln Thr Cys Phe Pro Arg Thr Val Val
    290                 295                 300

ttt ttg gct ttg cgt gag cac acc tct tgg gtg ctg ctg gca gct aac      960
Phe Leu Ala Leu Arg Glu His Thr Ser Trp Val Leu Leu Ala Ala Asn
```

```
305                 310                 315                 320

acg ctg ctg ctg ctg ctg ctg ctt ggg act gct ggc ctg ttt gcc tgg      1008
Thr Leu Leu Leu Leu Leu Leu Leu Gly Thr Ala Gly Leu Phe Ala Trp
                325                 330                 335

cac cta gac acc cct gtg gtg agg tca gca ggg ggc cgc ctg tgc ttt      1056
His Leu Asp Thr Pro Val Val Arg Ser Ala Gly Gly Arg Leu Cys Phe
            340                 345                 350

ctt atg ctg ggc tcc ctg gca gca ggt agt ggc agc ctc tat ggc ttc      1104
Leu Met Leu Gly Ser Leu Ala Ala Gly Ser Gly Ser Leu Tyr Gly Phe
        355                 360                 365

ttt ggg gaa ccc aca agg cct gcg tgc ttg cta cgc cag gcc ctc ttt      1152
Phe Gly Glu Pro Thr Arg Pro Ala Cys Leu Leu Arg Gln Ala Leu Phe
    370                 375                 380

gcc ctt ggt ttc acc atc ttc ctg tcc tgc ctg aca gtt cgc tca ttc      1200
Ala Leu Gly Phe Thr Ile Phe Leu Ser Cys Leu Thr Val Arg Ser Phe
385                 390                 395                 400

caa cta atc atc atc ttc aag ttt tcc acc aag gta cct aca ttc tac      1248
Gln Leu Ile Ile Ile Phe Lys Phe Ser Thr Lys Val Pro Thr Phe Tyr
                405                 410                 415

cac gcc tgg gtc caa aac cac ggt gct ggc ctg ttt gtg atg atc agc      1296
His Ala Trp Val Gln Asn His Gly Ala Gly Leu Phe Val Met Ile Ser
            420                 425                 430

tca gcg gcc cag ctg ctt atc tgt cta act tgg ctg gtg gtg tgg acc      1344
Ser Ala Ala Gln Leu Leu Ile Cys Leu Thr Trp Leu Val Val Trp Thr
        435                 440                 445

cca ctg cct gct agg gaa tac cag cgc ttc ccc cat ctg gtg atg ctt      1392
Pro Leu Pro Ala Arg Glu Tyr Gln Arg Phe Pro His Leu Val Met Leu
    450                 455                 460

gag tgc aca gag acc aac tcc ctg ggc ttc ata ctg gcc ttc ctc tac      1440
Glu Cys Thr Glu Thr Asn Ser Leu Gly Phe Ile Leu Ala Phe Leu Tyr
465                 470                 475                 480

aat ggc ctc ctc tcc atc agt gcc ttt gcc tgc agc tac ctg ggt aag      1488
Asn Gly Leu Leu Ser Ile Ser Ala Phe Ala Cys Ser Tyr Leu Gly Lys
                485                 490                 495

gac ttg cca gag aac tac aac gag gcc aaa tgt gtc acc ttc agc ctg      1536
Asp Leu Pro Glu Asn Tyr Asn Glu Ala Lys Cys Val Thr Phe Ser Leu
            500                 505                 510

ctc ttc aac ttc gtg tcc tgg atc gcc ttc ttc acc acg gcc agc gtc      1584
Leu Phe Asn Phe Val Ser Trp Ile Ala Phe Phe Thr Thr Ala Ser Val
        515                 520                 525

tac gac ggc aag tac ctg cct gcg gcc aac atg atg gct ggg ctg agc      1632
Tyr Asp Gly Lys Tyr Leu Pro Ala Ala Asn Met Met Ala Gly Leu Ser
    530                 535                 540

agc ctg agc agc ggc ttc ggt ggg tat ttt ctg cct aag tgc tac gtg      1680
Ser Leu Ser Ser Gly Phe Gly Gly Tyr Phe Leu Pro Lys Cys Tyr Val
545                 550                 555                 560

atc ctc tgc cgc cca gac ctc aac agc aca gag cac ttc cag gcc tcc      1728
Ile Leu Cys Arg Pro Asp Leu Asn Ser Thr Glu His Phe Gln Ala Ser
                565                 570                 575

att cag gac tac acg agg cgc tgc ggc tcc acc tga ccagtgggtc          1774
Ile Gln Asp Tyr Thr Arg Arg Cys Gly Ser Thr
            580                 585
```

agcaggcacg gctggcagcc ttctctgccc tgagggtcga aggtcgagca ggccgggggt 1834

gtccgggagg tctttgggca tcgcggtctg gggttgggac gtgtaagcgc ctgggagagc 1894

ctagaccagg ctccgggctg ccaataaaga agtgaaatgc gtaaaaaaaa a 1945

<210> 6

<211> 587

<212> PRT

<213> Homo sapiens

<400> 6

Met Leu Leu Cys Thr Ala Arg Leu Val Gly Leu Gln Leu Leu Ile Ser
1 5 10 15

Cys Cys Trp Ala Phe Ala Cys His Ser Thr Glu Ser Ser Pro Asp Phe
20 25 30

Thr Leu Pro Gly Asp Tyr Leu Leu Ala Gly Leu Phe Pro Leu His Ser
35 40 45

Gly Cys Leu Gln Val Arg His Arg Pro Glu Val Thr Leu Cys Asp Arg
50 55 60

Ser Cys Ser Phe Asn Glu His Gly Tyr His Leu Phe Gln Ala Met Arg
65 70 75 80

Leu Gly Val Glu Glu Ile Asn Asn Ser Thr Ala Leu Leu Pro Asn Ile
85 90 95

Thr Leu Gly Tyr Gln Leu Tyr Asp Val Cys Ser Asp Ser Ala Asn Val
100 105 110

Tyr Ala Thr Leu Arg Val Leu Ser Leu Pro Gly Gln His His Ile Glu
115 120 125

Leu Gln Gly Asp Leu Leu His Tyr Ser Pro Thr Val Leu Ala Val Ile
130 135 140

Gly Pro Asp Ser Thr Asn Arg Ala Ala Thr Thr Ala Ala Leu Leu Ser
145 150 155 160

Pro Phe Leu Val Pro Met Leu Leu Glu Gln Ile His Lys Val His Phe
165 170 175

Leu Leu His Lys Asp Thr Val Ala Phe Asn Asp Asn Arg Asp Pro Leu
180 185 190

Ser Ser Tyr Asn Ile Ile Ala Trp Asp Trp Asn Gly Pro Lys Trp Thr

195 200 205

Phe Thr Val Leu Gly Ser Ser Thr Trp Ser Pro Val Gln Leu Asn Ile
210 215 220

Asn Glu Thr Lys Ile Gln Trp His Gly Lys Asp Asn Gln Val Pro Lys
225 230 235 240

Ser Val Cys Ser Ser Asp Cys Leu Glu Gly His Gln Arg Val Val Thr
245 250 255

Gly Phe His His Cys Cys Phe Glu Cys Val Pro Cys Gly Ala Gly Thr
260 265 270

Phe Leu Asn Lys Ser Asp Leu Tyr Arg Cys Gln Pro Cys Gly Lys Glu
275 280 285

Glu Trp Ala Pro Glu Gly Ser Gln Thr Cys Phe Pro Arg Thr Val Val
290 295 300

Phe Leu Ala Leu Arg Glu His Thr Ser Trp Val Leu Leu Ala Ala Asn
305 310 315 320

Thr Leu Leu Leu Leu Leu Leu Leu Gly Thr Ala Gly Leu Phe Ala Trp
325 330 335

His Leu Asp Thr Pro Val Val Arg Ser Ala Gly Gly Arg Leu Cys Phe
340 345 350

Leu Met Leu Gly Ser Leu Ala Ala Gly Ser Gly Ser Leu Tyr Gly Phe
355 360 365

Phe Gly Glu Pro Thr Arg Pro Ala Cys Leu Leu Arg Gln Ala Leu Phe
370 375 380

Ala Leu Gly Phe Thr Ile Phe Leu Ser Cys Leu Thr Val Arg Ser Phe
385 390 395 400

Gln Leu Ile Ile Ile Phe Lys Phe Ser Thr Lys Val Pro Thr Phe Tyr
405 410 415

His Ala Trp Val Gln Asn His Gly Ala Gly Leu Phe Val Met Ile Ser
420 425 430

Ser Ala Ala Gln Leu Leu Ile Cys Leu Thr Trp Leu Val Val Trp Thr
435 440 445

Pro Leu Pro Ala Arg Glu Tyr Gln Arg Phe Pro His Leu Val Met Leu
450 455 460

Glu Cys Thr Glu Thr Asn Ser Leu Gly Phe Ile Leu Ala Phe Leu Tyr
465 470 475 480

Asn Gly Leu Leu Ser Ile Ser Ala Phe Ala Cys Ser Tyr Leu Gly Lys
485 490 495

Asp Leu Pro Glu Asn Tyr Asn Glu Ala Lys Cys Val Thr Phe Ser Leu
500 505 510

Leu Phe Asn Phe Val Ser Trp Ile Ala Phe Phe Thr Thr Ala Ser Val
515 520 525

Tyr Asp Gly Lys Tyr Leu Pro Ala Ala Asn Met Met Ala Gly Leu Ser
530 535 540

Ser Leu Ser Ser Gly Phe Gly Gly Tyr Phe Leu Pro Lys Cys Tyr Val
545 550 555 560

Ile Leu Cys Arg Pro Asp Leu Asn Ser Thr Glu His Phe Gln Ala Ser
565 570 575

Ile Gln Asp Tyr Thr Arg Arg Cys Gly Ser Thr
580 585

<210> 7

<211> 1302

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(975)

<223>

<400> 7

atg ctg ctc tgc acg gct cgc ctg gtc ggc ctg cag ctt ctc att tcc 48
Met Leu Leu Cys Thr Ala Arg Leu Val Gly Leu Gln Leu Leu Ile Ser
1 5 10 15

tgc tgc tgg gcc ttt gcc tgc cat agc acg gag tct tct cct gac ttc 96
Cys Cys Trp Ala Phe Ala Cys His Ser Thr Glu Ser Ser Pro Asp Phe
20 25 30

acc ctc ccc gga gat tac ctc ctg gca ggc ctg ttc cct ctc cat tct 144
Thr Leu Pro Gly Asp Tyr Leu Leu Ala Gly Leu Phe Pro Leu His Ser
35 40 45

ggc tgt ctg cag gtg agg cac aga ccc gag gtg acc ctg tgt gac agg 192
Gly Cys Leu Gln Val Arg His Arg Pro Glu Val Thr Leu Cys Asp Arg
50 55 60

tct tgt agc ttc aat gag cat ggc tac cac ctc ttc cag gct atg cgg 240
Ser Cys Ser Phe Asn Glu His Gly Tyr His Leu Phe Gln Ala Met Arg

```
65                  70                  75                  80

ctt ggg gtt gag gag ata aac aac tcc acg gcc ctg ctg ccc aac atc      288
Leu Gly Val Glu Glu Ile Asn Asn Ser Thr Ala Leu Leu Pro Asn Ile
                85                  90                  95

acc ctg ggg tac cag ctg tat gat gtg tgt tct gac tct gcc aat gtg      336
Thr Leu Gly Tyr Gln Leu Tyr Asp Val Cys Ser Asp Ser Ala Asn Val
            100                 105                 110

tat gcc acg ctg aga gtg ctc tcc ctg cca ggg caa cac cac ata gag      384
Tyr Ala Thr Leu Arg Val Leu Ser Leu Pro Gly Gln His His Ile Glu
        115                 120                 125

ctc caa gga gac ctt ctc cac tat tcc cct acg gtg ctg gca gtg att      432
Leu Gln Gly Asp Leu Leu His Tyr Ser Pro Thr Val Leu Ala Val Ile
    130                 135                 140

ggg cct gac agc acc aac cgt gct gcc acc aca gcc gcc ctg ctg agc      480
Gly Pro Asp Ser Thr Asn Arg Ala Ala Thr Thr Ala Ala Leu Leu Ser
145                 150                 155                 160

cct ttc ctg gtg ccc atg ctt ttg gag cag atc cac aag gtg cat ttc      528
Pro Phe Leu Val Pro Met Leu Leu Glu Gln Ile His Lys Val His Phe
                165                 170                 175

ctt cta cac aag gac act gtg gcg ttt aat gac aac aga gat ccc ctc      576
Leu Leu His Lys Asp Thr Val Ala Phe Asn Asp Asn Arg Asp Pro Leu
            180                 185                 190

agt agc tat aac ata att gcc tgg gac tgg aat gga ccc aag tgg acc      624
Ser Ser Tyr Asn Ile Ile Ala Trp Asp Trp Asn Gly Pro Lys Trp Thr
        195                 200                 205

ttc acg gtc ctc ggt tcc tcc aca tgg tct cca gtt cag cta aac ata      672
Phe Thr Val Leu Gly Ser Ser Thr Trp Ser Pro Val Gln Leu Asn Ile
    210                 215                 220

aat gag acc aaa atc cag tgg cac gga aag gac aac cag gtg cct aag      720
Asn Glu Thr Lys Ile Gln Trp His Gly Lys Asp Asn Gln Val Pro Lys
225                 230                 235                 240

tct gtg tgt tcc agc gac tgt ctt gaa ggg cac cag cga gtg gtt acg      768
Ser Val Cys Ser Ser Asp Cys Leu Glu Gly His Gln Arg Val Val Thr
                245                 250                 255

ggt ttc cat cac tgc tgc ttt gag tgt gtg ccc tgt ggg gct ggg acc      816
Gly Phe His His Cys Cys Phe Glu Cys Val Pro Cys Gly Ala Gly Thr
            260                 265                 270

ttc ctc aac aag agt gct acc tgg gta agg act tgc cag aga act aca      864
Phe Leu Asn Lys Ser Ala Thr Trp Val Arg Thr Cys Gln Arg Thr Thr
        275                 280                 285

acg agg cca aat gtg tca cct tca gcc tgc tct tca act tcg tgt cct      912
Thr Arg Pro Asn Val Ser Pro Ser Ala Cys Ser Ser Thr Ser Cys Pro
    290                 295                 300

gga tcg cct tct tca cca cgg cca gcg tct acg acg gca agt acc tgc      960
Gly Ser Pro Ser Ser Pro Arg Pro Ala Ser Thr Thr Ala Ser Thr Cys
305                 310                 315                 320

ctg cgg cca aca tga tggctgggct gagcagcctg agcagcggct tcggtgggta     1015
Leu Arg Pro Thr

ttttctgcct aagtgctacg tgatcctctg ccgcccagac ctcaacagca cagagcactt   1075

ccaggcctcc attcaggact acacgaggcg ctgcggctcc acctgaccag tgggtcagca   1135
```

ggcacggctg gcagccttct ctgccctgag ggtcgaaggt cgagcaggcc gggggtgtcc 1195

gggaggtctt tgggcatcgc ggtctggggt tgggacgtgt aagcgcctgg gagagcctag 1255

accaggctcc gggctgccaa taaagaagtg aaatgcgtaa aaaaaaa 1302

<210> 8

<211> 324

<212> PRT

<213> Homo sapiens

<400> 8

Met Leu Leu Cys Thr Ala Arg Leu Val Gly Leu Gln Leu Leu Ile Ser
1 5 10 15

Cys Cys Trp Ala Phe Ala Cys His Ser Thr Glu Ser Ser Pro Asp Phe
20 25 30

Thr Leu Pro Gly Asp Tyr Leu Leu Ala Gly Leu Phe Pro Leu His Ser
35 40 45

Gly Cys Leu Gln Val Arg His Arg Pro Glu Val Thr Leu Cys Asp Arg
50 55 60

Ser Cys Ser Phe Asn Glu His Gly Tyr His Leu Phe Gln Ala Met Arg
65 70 75 80

Leu Gly Val Glu Glu Ile Asn Asn Ser Thr Ala Leu Leu Pro Asn Ile
85 90 95

Thr Leu Gly Tyr Gln Leu Tyr Asp Val Cys Ser Asp Ser Ala Asn Val
100 105 110

Tyr Ala Thr Leu Arg Val Leu Ser Leu Pro Gly Gln His His Ile Glu
115 120 125

Leu Gln Gly Asp Leu Leu His Tyr Ser Pro Thr Val Leu Ala Val Ile
130 135 140

Gly Pro Asp Ser Thr Asn Arg Ala Ala Thr Thr Ala Ala Leu Leu Ser
145 150 155 160

Pro Phe Leu Val Pro Met Leu Leu Glu Gln Ile His Lys Val His Phe
165 170 175

Leu Leu His Lys Asp Thr Val Ala Phe Asn Asp Asn Arg Asp Pro Leu
180 185 190

Ser Ser Tyr Asn Ile Ile Ala Trp Asp Trp Asn Gly Pro Lys Trp Thr

195 200 205

Phe Thr Val Leu Gly Ser Ser Thr Trp Ser Pro Val Gln Leu Asn Ile
210 215 220

Asn Glu Thr Lys Ile Gln Trp His Gly Lys Asp Asn Gln Val Pro Lys
225 230 235 240

Ser Val Cys Ser Ser Asp Cys Leu Glu Gly His Gln Arg Val Val Thr
245 250 255

Gly Phe His His Cys Cys Phe Glu Cys Val Pro Cys Gly Ala Gly Thr
260 265 270

Phe Leu Asn Lys Ser Ala Thr Trp Val Arg Thr Cys Gln Arg Thr Thr
275 280 285

Thr Arg Pro Asn Val Ser Pro Ser Ala Cys Ser Ser Thr Ser Cys Pro
290 295 300

Gly Ser Pro Ser Ser Pro Arg Pro Ala Ser Thr Thr Ala Ser Thr Cys
305 310 315 320

Leu Arg Pro Thr

<210> 9

<211> 401

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (201)..(201)

<223> SNP38 y = t or c

<400> 9

tgacatcaca gtgtggtgtt ttctcttgga tgcaaaccac actcttcact tggcttcaaa 60

acagcttaag cgagcgaaca gagactagga gggaacggaa actgtccgag caggcgggcg 120

ggggcggctt aggatctgta gtccttttg ctgtagggtt tattgcccaa aatactatca 180

atctcgtgga taatggaaga ygacagtttc ggaaggacct gcgtgcaaga acaacaacaa 240

atcagaatcc caccaacaag acgggcctgg cctggggagg aggtggggcc aggctcttcc 300

agatggtcgg ctggccaggc ccgggttgcc ccatcacatc tgggctccat ctgaggcaga 360

tggaggccac gctgcttttg tgtttcagaa ggtttggggc a 401

<210> 10

<211> 701

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (201)..(201)

<223> SNP47 r = a or g

<400> 10

cacagacagc cgtgtgccca ggctcaccca cccgaccgcc tcagctcatg catgagacct 60

gttcacgtgc tgtgacccag ccgctgtgtg cagcccccg gcgcctctct gccggccccg 120

gggagagctg gggcaggcac acacgggtga acaggtctca gttctggtgc tcaggctgag 180

ggtgtgcatg gtgcatctgg rtgccaccac ttgggatctc aaaggcctcg ccacacctca 240

ccaggataaa ggggattcca agaatggcct tgactcagag agtgggcatg gatattccag 300

gagcctagtt ttaaagagct ttgaggctgg acatggtggc ttatacccat aatcccaaca 360

ctttgggaag ccaaggaggc agaaggatca ttttagtcca gatgctcaag accagcctgg 420

gcaacataga gagactccat ctctacaaaa acattaaaaa accagcaggg cctggtggtg 480

tgcacctgca gtcccagctg cttgggaggc tgaggcagga gtgctgcttc aggtcagatg 540

cagtgagctg tgatctcgcc actgcactcc agcctgggta acagagcaag acaaccctgt 600

ctttttttt ttttttttc agatggagtt ttgctcttgt cacccaggat ggagtgcaat 660

ggtatgatct cggctcactg aacctctgcc tcccgggttc a 701

<210> 11

<211> 668

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (81)..(81)

<223> SNP49 r = g or a

<400> 11

tgctgccagt ggaatgctgg gataacccca gtgtcggggg tcctggggca gcctgtgtac 60

tgctggatgc ttgctttcaa rtactctgct ggggtaagac tctggggcag tttccatccc 120

ttccttgggc ttcagcttgc tggactggcc ctgagaactg ggaggggatg gcttcatgaa 180

gtcattttga tgggggatga agatagttgc tgtagaggaa gaggatgaga gagctgcttg 240

gcctctggaa ggtgggagga gctgtcgcag gctgcacagg ggctctgagg tcgagttcgt 300

ttagttcatt gattaagcag tctaccaggt ggccagcaag gctttgtggg tgcttgggat 360

gcatttaagg gaactaaaca aaggtgcttg tgcagccgac ctgcttgcca ggcaatgcag 420

ataacagata agtccatctg cctgcagtag acggtgagga agtgctgccg aagaagaaaa 480

agcagaggcg ggccaagggg ttggggagtt gtgagtgggg ttcgattgcg attttaagta 540

gccaggtctc tagagtctta ctgagaagtg agatataagc aaagactcaa agaggtgaga 600

agttgaccaa gtgggtgact ggggagcatc tgggcagaaa gagaaccagg gcgctgggta 660

ggaggatc 668

<210> 12

<211> 1496

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (201)..(201)

<223> SNP56 s = g or c

<400> 12

tgattaagtt aggtttgttt gggggtgtgg gggcaaagga gtttgactgt ctctgcgggc 60

atatattgct ttagtttttt tgaccctatg acaaaatagt ttgatggaga tttacattat 120

ttcctggcca tttttgctat tacagaggtt gatatacact ctttggcaag acctctctga 180

tctagaaaac tccatcatca scatggtttg tgtaatagct gagagacaca attcagcttt 240

attagaaaaa aatgtaaatg atgtatcggc caggcatggg ggctcacacc tgtaatccca 300

atatgtccag agttcattcc ttttggtggg ttcatggtct cactgacttc aacaatgaag 360

ccatggacct tcgcggtgag tgttaccttа aaggtggtgc ggacccagag tgagcagcag 420

caagatttat tgtggagagc aaaataacaa agcttccaca gtgtggaagg ggacccgagc 480

aggttgttgc tggctggggt ggccagcttt tattccctta tttgtccctg cctatgtcct 540

gctgattggt ccattttaca gagtgccggt tggtccattt acagagtgct gattagtgca 600

cttacaatac tctagctagc cacagagcac taattggtgc tttttacagc gagctgattg 660

gtgcatttac aatcctctag ctagccacag agcactgatt ggtgcgtttt tacagagtgc 720

tgattggtac atttacaacc ctctagcgac agagcgctga ttggtgcgtt tttacagagt 780
gctgattggt aacatttgta agacagaaaa gttctccaag tccccacctg acccaggaag 840
tccagctggc ttcacctctc accaacactt tgggaggctg aggtgggtgg atcatctgag 900
gtcaggagtt tgagaccagc ctgaccaaca tggtgaaacc ccatctctac taaaaataca 960
aaaattagcc aggcgtggtg gcaggcgcct gtaatccaag ctgcttggga ggctgaagca 1020
ggagaatcac ttgaacctgg gaggcggggg ttgcagtgag ccgagatcgc gccactgcac 1080
tccagcctga gcaacaagct caagactccg tttctaaata aataaataaa aataatgtag 1140
caagagcctg ggcaacatgg tgaaatccta tctctacaaa aaaaaaaaac aaaaaacaaa 1200
aaacaattaa ctgggcatgg tggcgtgctc ctgtagtccc agctactcgg gaggctgagg 1260
caggaggaat gcttgagccc tggaggtgga ggttgcagta agccaggatc gtgccagcct 1320
ggctccagcc ttcaagatgt actccagcct gggtgacaga gagagacccc atctgaataa 1380
cgtatcaagg aaggcatgta accataaaaa tatttttaaa tgccaaaatg ttttcataaa 1440
gatgtaagcg tgcatcactt gattattttt aaaggatgtt ggggttgccc tggcta 1496

<210> 13
<211> 572
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (211)..(211)
<223> SNP69 y = t or c

<400> 13
cacccacatg acctcagaag agctccaggc tccacgacct gtccagaggc agcaaggtgg 60
atgcagagcc cgggggtccc ctgggcttcc tggattgaga ggcagtcccc ctccaccctt 120
ttagcactga ctcagctggc gattcctctc atggttccag gattgccctg atgagccatc 180
atgggggcac agcagcagga ttccagtcct ytccctgcta aatggtttcc tgtgtgggaa 240
ttgaggctca ggttgtgcct ggaggagtga gactgggact gaagcagggc tagaggaccc 300
aggggtctgg gatgggaaga aaccaccccc atcatcatgg cagggtgccc atggtggccc 360
ttgaagagtg gtccccacat cttaaaatcc tacaacactt agaatcctag aacatagggc 420
tggagccatg ttagagagca agtaatttga ccctaatatt ggaaataagg ggagaatagg 480
gaagaaggaa atagaacaga agaccttgtc taagagaccc aggaggtcat tggagagctg 540
tgaccagaaa ctcaggtagc tgaagatgga gc 572