



(11) **EP 1 891 237 B9**

(12) **CORRECTED EUROPEAN PATENT SPECIFICATION**

- (15) Correction information:
Corrected version no 1 (W1 B1)
Corrections, see
Description Paragraph(s) 64, 76, 80, 84, 87
- (51) Int Cl.:
C12Q 1/68 (2006.01)
- (86) International application number:
PCT/EP2006/063256
- (48) Corrigendum issued on:
21.07.2010 Bulletin 2010/29
- (87) International publication number:
WO 2006/134154 (21.12.2006 Gazette 2006/51)
- (45) Date of publication and mention of the grant of the patent:
18.11.2009 Bulletin 2009/47
- (21) Application number: **06763749.6**
- (22) Date of filing: **15.06.2006**

(54) **ENPP1 (PC-1) GENE HAPLOTYPE ASSOCIATED WITH THE RISK OF OBESITY AND TYPE 2 DIABETES AND THEIR APPLICATIONS**

MIT DEM RISIKO VON OBESITÄT UND TYP-2-DIABETES ASSOZIIERTER HAPLOTYP DES GENS ENPP1 (PC1) UND ANWENDUNGEN DAVON

HAPLOTYPE DE GENE ENPP1 (PC-1) ASSOCIE AU RISQUE DE DEVELOPPER UNE OBESITE ET UN DIABETE DE TYPE 2 ET APPLICATIONS CORRESPONDANTES

- (84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LI LT LU LV MC NL PL PT RO SE SI SK TR
- (30) Priority: **17.06.2005 EP 05291298**
- (43) Date of publication of application:
27.02.2008 Bulletin 2008/09
- (73) Proprietors:
 - **CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS)**
75016 Paris (FR)
 - **UNIVERSITE DU DROIT ET DE LA SANTE DE LILLE 2**
59800 Lille (FR)
- (72) Inventors:
 - **FROGUEL, Philippe**
F-93170 Bagnole (FR)
 - **MEYRE, David**
F-59164 Marpent (FR)
 - **DINA, Christian**
F-75009 Paris (FR)
- **BOUTIN, Philippe**
F-59200 Tourcoing (FR)
- (74) Representative: **Warcoing, Jacques et al**
Cabinet Régimbeau
20, rue de Chazelles
75847 Paris Cedex 17 (FR)
- (56) References cited:
US-B1- 6 465 185
- **FRITTITTA LUCIA ET AL: "The Q121 PC-1 variant and obesity have additive and independent effects in causing insulin resistance" JOURNAL OF CLINICAL ENDOCRINOLOGY AND METABOLISM, vol. 86, no. 12, December 2001 (2001-12), pages 5888-5891, XP002354157 ISSN: 0021-972X**
 - **HAMAGUCHI KAZUYUKI ET AL: "The PC-1 Q121 allele is exceptionally prevalent in the Dominican Republic and is associated with type 2 diabetes." JOURNAL OF CLINICAL ENDOCRINOLOGY & METABOLISM, vol. 89, no. 3, March 2004 (2004-03), pages 1359-1364, XP002354158 ISSN: 0021-972X**

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

EP 1 891 237 B9

- **GODING J W ET AL:** "Physiological and pathophysiological functions of the ecto-nucleotide pyrophosphatase/phosphodiesterase family" **BIOCHIMICA ET BIOPHYSICA ACTA. MOLECULAR BASIS OF DISEASE**, AMSTERDAM, NL, vol. 1638, no. 1, 20 May 2003 (2003-05-20), pages 1-19, XP004426713 ISSN: 0925-4439
- **STEFANOVIC VLADISAV ET AL:** "Plasma cell membrane glycoprotein 1 (PC-1): a marker of insulin resistance in obesity, uremia and diabetes mellitus." **CLINICAL LABORATORY**. 2004, vol. 50, no. 5-6, 2004, pages 271-278, XP009057018 ISSN: 1433-6510
- **FRITTITTA LUCIA ET AL:** "A cluster of three single nucleotide polymorphisms in the 3'-untranslated region of human glycoprotein PC-1 gene stabilizes PC-1 mRNA and is associated with increased PC-1 protein content and insulin resistance-related abnormalities" **DIABETES**, vol. 50, no. 8, August 2001 (2001-08), pages 1952-1955, XP002354159 ISSN: 0012-1797 cited in the application
- **MEYRE DAVID ET AL:** "Variants of ENPP1 are associated with childhood and adult obesity and increase the risk of glucose intolerance and type 2 diabetes" **NATURE GENETICS**, vol. 37, no. 8, August 2005 (2005-08), pages 863-867, XP002354160 ISSN: 1061-4036

Description

[0001] The present invention is directed to a method for determining if a subject is at increased risk to develop obesity or pathology related to obesity by determining the presence of a haplotype comprising three specific SNPs in a DNA or RNA sample of this subject and/or an elevated serum ENPP1 protein concentration. The present invention also relates to a kit and to an isolated nucleic sequence, vector or recombinant cell comprises said *ENPP1* gene haplotype. The invention further comprises a method for selecting a compound for the treatment or the prevention of obesity or pathology related to obesity and a method for determining the efficacy of a drug to reduce the risk of obesity or pathology related to obesity in a patient.

[0002] Obesity is a devastating disease. In addition to harming physical health, obesity can wreak havoc on mental health because obesity affects self-esteem, which ultimately can affect a person's ability to interact socially with others. Unfortunately, obesity is not well understood, and societal stereotypes and presumptions regarding obesity only tend to exacerbate the psychological effects of the disease. Because of the impact of obesity on individuals and society, much effort has been expended to find ways to treat obesity, but little success has been achieved in the long-term treatment and/or prevention of obesity.

[0003] Obesity represents the most prevalent of body weight disorders, and it is the most important nutritional disorder in the western world, with estimates of its prevalence ranging from 30% to 50% within the middle-aged population. Obesity, defined as an excess of body fat relative to lean body mass, also contributes to other diseases. For example, this disorder is responsible for increased incidence of diseases such as coronary artery disease, hypertension, stroke, diabetes, hyperlipidemia, and some cancers. Obesity is not merely a behavioral problem, i.e., the result of voluntary hyperphagia. Rather, the differential body composition observed between obese and normal subjects results from differences in both metabolism and neurologic/metabolic interactions. These differences seem to be, to some extent, due to differences in gene expression, and/or level of gene products or activity. The epidemiology of obesity strongly shows that the disorder exhibits inherited characteristics.

[0004] In summary, therefore, obesity, which poses a major, worldwide health problem, represents a complex, highly heritable trait. Given the severity, prevalence, and potential heterogeneity of such disorders, there exists a great need for the identification of those genes that participate in the control of obesity, particularly so as to be able anticipate the risk of developing this pathology or pathology related to obesity and/or to develop better targeted medicaments. This is the object of the present invention.

[0005] Previous work identified a childhood obesity locus on chromosome 6q16.3-q24.2 (lod score = 4.06)¹ that includes a 2.4 Mb type 2 diabetes/obesity-linked region common to all eight published genome scans reporting linkage in this region¹⁻⁸.

[0006] The patent document US 6,465,185 B1 discloses the use of antibodies for the determination of the PC-1 protein and pair of primers for amplifying the exon 4 of the PC-1 protein.

[0007] The present inventors have, by conducting genotype-phenotype association studies in the 580kb *ENPP1* region and ENPP1 mRNA expression studies in 3,147 French Caucasian and in more than 6,000 European subjects individuals.

[0008] The *ENPP1* gene encodes the ecto-nucleotide pyrophosphatase/phosphodiesterase 1 (ENPP1) which is a member of the ecto-nucleotide pyrophosphatase/phosphodiesterase (ENPP) family. The encoded protein is a type II transmembrane glycoprotein comprising two identical disulfide-bonded subunits. This protein has broad specificity and cleaves a variety of substrates, including phosphodiester bonds of nucleotides and nucleotide sugars and pyrophosphate bonds of nucleotides and nucleotide sugars. This protein may function to hydrolyze nucleoside 5' triphosphates to their corresponding monophosphates and may also hydrolyze diadenosine polyphosphates. Mutations in this gene, particularly, have been associated with 'idiopathic' infantile arterial calcification (GLU893TER, ARG774CYS or LEU579PHE), ossification of the posterior longitudinal ligament of the spine (OPLL) (IVS20AS, 1-BP DEL,T,-11) and insulin resistance (LYS121GLN, the "K121Q" mutation).

[0009] The inventors have demonstrated a strongest association to be between variant of this *ENPP1* gene, particularly a three allele risk haplotype (K121Q; IVS20 delT-11 and a SNP localized in the 3'UTR domain sequence (notably A>G +1044 TGA, QdelTG) and childhood obesity and adults with morbid obesity. The inventors have also demonstrated the presence of at least one allele of the risk haplotype was associated with increased serum levels of soluble ENPP1 protein in children suggesting additional elevated ENPP1 expression. In addition, expression of a long ENPP1 mRNA isoform comprising the obesity-associated A>G +1044 TGA SNP was found to be specific for three tissues of major importance for glucose homeostasis: the pancreatic islet D-cells, adipocytes and liver. These findings demonstrate for the first time a primary role for several variants of ENPP in mediating insulin-resistance, both in the development of obesity and type 2 diabetes (T2D), providing common molecular mechanism of both widespread afflictions.

[0010] The inventors have also provided screening assays which can be used to diagnostic the risk of developing obesity and/or pathology related to obesity and to identify molecules for therapeutic treatment.

[0011] The present invention describes a method for determining if a subject is at increased risk to develop obesity or pathology relate to obesity from a biological sample from said subject containing genomic DNA or RNA, said method

comprising the steps of:

- 5 a) determining on at least one gene allele or RNA sequence encoding the protein ENPP1, if said nucleic sequence, or fragments thereof, contains at least one SNP selected from the group of SNPs consisting of the SNPs indicated in table 3 and wherein the variant allele frequency is superior to 5 %; and
- b) observing whether or not the subject is at increased risk to develop obesity or pathology related to obesity by observing if said sequence, or fragments thereof, contains said at least one SNP, the presence of said SNP indicates said subject is at increased risk to develop obesity or pathology related to obesity.

10 **[0012]** It is also described a method for determining if a subject is at increased risk to develop obesity or pathology related to obesity from a biological sample from said subject containing genomic DNA or RNA, said method comprising the steps of:

- 15 a) determining on at least one gene allele or RNA sequence encoding the protein ENPP1, if said nucleic sequence, or fragments thereof, contains at least one SNP selected from the group of SNPs consisting of IVS2 delG +8, K121Q, IVS8 T>G +27, IVS20 delT -11, A>G +1044 TGA and T>G +5954 TGA; and
- b) observing whether or not the subject is at increased risk to develop obesity or pathology related to obesity by observing if said sequence, or fragments thereof, contains said at least one SNP, the presence of said SNP indicates said subject is at increased risk to develop obesity or pathology related to obesity.

20 **[0013]** In a first aspect, the present invention is directed to a method for determining if a subject is at increased risk to develop obesity or pathology related to obesity from a biological sample from said subject containing genomic DNA or RNA, said method comprising the steps of:

- 25 a) determining on at least one gene allele or RNA sequence encoding the protein ENPP 1, if said nucleic sequence, or fragments thereof, contains the haplotype comprising the three following SNPs:

- 30 - K121Q (rs1044498),
- IVS20 delT-11, and
- a SNP localized in the 3'UTR domain sequence and wherein the first 1170 bp 3'UTR domain wild type sequence (including the TGA codon) has the sequence SEQ ID No. 15; and

- 35 b) observing whether or not the subject is at increased risk to develop obesity or pathology related to obesity by observing if said sequence, or fragments thereof, contains said haplotype, the presence of said haplotype indicates said subject is at increased risk to develop obesity or pathology related to obesity.

[0014] In the present description, the term "rs" for example in the wording "rs 1044498" corresponds to the nomenclature of the SNP in the databank dbSNP.

40 **[0015]** The SNP "K121Q (rs1044498)" is intended to designate the SNP rs1044498 corresponding to the ENPP1 variant "K121Q" (121 is the amino acid position in the ENPP1 amino acid sequence having the accession number (A: N.) NP 006199 (SEQ ID No. 14), which corresponds to the position 173 in the ENPP1 amino acid sequence having the accession number AAH59375. The accession number BC059375 corresponds to the mRNA nucleotide sequence encoding the human ENPP protein having the sequence depicted under A.N. AAH59375).

45 **[0016]** The SNP "IVS20 delT-11" correspond to the position -11 before exon 21 of ENPP1 as shown below:

Aaaaattagccaggcgtggtgggtcacacctgtaatcccagcactttgagggggctgcagcggatggatcacctgaggtca
 5 ggagctcaagaccagccttgccaacatggtgaaacccatctctacccaaaacacaaaaattagccaggcgtggtggcagat
 gcctgtagtcccagctactcgggaagctgaggcaggagaatcgcttgaacctggggggcagagttgcgggtgagccgagatt
 gcgccgctgccctctagtctgggtgacagagtgagactccatcttaaaaaataataataataataataataataataataat
 10 atttaaaattgttagataaaatcattctaaacattattcatattagcatagcagaatctgaaaatattgcataaataatgacaattaat
 atctttaaattgtaaagcatttttacactttagttagaaaaaaagatgaatatactagtaggaaaataggggaaggacatgagctga
 cagctagagcttcataattttatgatgtagttcacctttaaataataaaagcaA ttttctctctgtgcctgatatctgagagttcttc
 15 tcatttt(delT)cgttcttcagGACAGTTTCTCTACGGAAGACTTCTCCAAGTGTCTGTAC
 CAGGACTTTAGAATTCCTCTTAGTCCTGTCCATAAATGTTTCATTTTATAAAA
 ATAACACCAAAGTGAGTTACGGGTTCTCTCCCCACCACgtaagtttttctctcctgacc
 20 ttccttttctcctttttgttttcttctgtttataaatcctaccatacattatagggtaat (SEQ ID No: 1).

[0017] The term "haplotype" refers to a combination of alleles present in an individual or a sample. In the context of
 25 the present invention a haplotype preferably refers to a combination of polymorphism alleles found in a given individual
 and which may be associated with a phenotype.

[0018] The term "polymorphism" as used herein refers to the occurrence of two or more alternative genomic sequences
 or alleles between or among different genomes or individuals. "Polymorphic" refers to the condition in which two or more
 30 variants of a specific genomic sequence can be found in a population. A "polymorphic site" is the locus at which the
 variation occurs. A "single nucleotide polymorphism" (SNP) is a single base pair change. Typically a single nucleotide
 polymorphism is the replacement of one nucleotide by another nucleotide at the polymorphic site. Deletion of a single
 nucleotide or insertion of a single nucleotide, also give rise to single nucleotide polymorphisms.

[0019] Variant nucleic will be understood to mean all the alternative nucleic sequence which may naturally exist, in
 particular in human beings, and which correspond in particular to deletions, substitutions and/or additions of nucleotides.
 35 In the present case, the variant nucleic sequence will be in particular partly associated with the increased risk to develop
 obesity or pathology related to obesity.

[0020] As used interchangeably herein, the terms "oligonucleotides", "nucleic acids" and "polynucleotides" include
 RNA, DNA, or RNA/DNA hybrid sequences of more than one nucleotide in either single chain or duplex form.

[0021] As used herein, the term "nucleic acids" and "nucleic acid molecule" is intended to include DNA molecules
 40 (e.g., cDNA or genomic DNA) and RNA molecules (e.g., mRNA) and analogs of the DNA or RNA generated using
 nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-
 stranded DNA. Throughout the present specification, the expression "nucleotide sequence" may be employed to des-
 ignate indifferently a polynucleotide or a nucleic acid. More precisely, the expression "nucleotide sequence" encompasses
 45 the nucleic material itself and is thus not restricted to the sequence information (i.e. the succession of letters chosen
 among the four base letters) that biochemically characterizes a specific DNA or RNA molecule. Also, used interchangeably
 herein are terms "nucleic acids", "oligonucleotides", and "polynucleotides".

[0022] In a preferred embodiment, said pathology related to obesity is selected from the group consisting of type 2
 diabetes, heart disease, coronary artery disease, myocardial infarct, hypertension and lipid related metabolic such as
 hypercholesterolemia and hyperlipidemia.

[0023] In a more preferred embodiment, said pathology related to obesity is type 2 diabetes (T2D).

[0024] Concerning said SNP which is localized in the 3'UTR domain sequence, it is preferred that this SNP is selected
 from the group of SNPs consisting of A>G +1044 TGA (rs7754561), A>C +1092 TGA (rs7754586), C>T +1157 TGA
 (rs7754859), G>T +1101 TGA, T>C +1137 TGA C>T +1157 TGA, G>T +1236 TGA, T>C +1348 TGA, G>A +1350 TGA,
 55 A>G +1539 TGA and T>C +1670 TGA. More preferred is a SNP selected from the group consisting of A>G +1044 TGA,
 A>C +1092 TGA and C>T +1157 TGA, and the SNP A>G +1044 TGA localized in the 3'UTR domain being the most
 preferred.

[0025] Is also preferred a method for determining if a subject is at high increased risk to develop obesity or pathology
 related to obesity, characterized in that said method comprises the step a) and b) of the above method according to the

invention wherein:

- in step a), it is determined whether or not the subject is homozygous for said haplotype comprising said three SNPs; and
- in step b), it is observed whether or not the subject is at high increased risk to develop obesity or pathology related to obesity by observing if the subject is homozygous for said haplotype, the homozygosity indicating said subject is at high increased risk to develop obesity or pathology related to obesity.

[0026] The methods allowing the detection of variation in a gene compared with the natural gene are, of course, highly numerous. They can essentially be divided into two large categories. The first type of method is that in which the presence of a variation is detected by comparing the alternative sequence with the corresponding normal sequence(s), and the second type is that in which the presence of the variation is detected indirectly, for example by evidence of the mismatches due to the presence of the variation.

[0027] Among the methods for the determination of an allelic variability, the methods comprising at least one stage for the so-called PCR (polymerase chain reaction) or PCR-like amplification of the target sequence comprising the three allele risk haplotype according to the invention with the aid of a pair of primers of nucleotide sequences according to the invention are preferred. The amplified products may be treated with the aid of an appropriate restriction enzyme before carrying out the detection or assay of the targeted product.

[0028] PCR-like will be understood to mean all methods using direct or indirect reproductions of nucleic acid sequences, or alternatively in which the labeling systems have been amplified, these techniques are of course known, in general they involve the amplification of DNA by a polymerase; when the original sample is an RNA, it is advisable to carry out a reverse transcription beforehand. There are currently a great number of methods allowing this amplification, for example the so-called NASBA "Nucleic Acid Sequence Based Amplification", TAS "transcription based Amplification System", LCR "Ligase Chain Reaction", "Endo Run Amplification" (ERA), "Cycling Probe Reaction" (CPR), and SDA "Strand Displacement Amplification", methods well known to persons skilled in the art.

[0029] So, also forms part of the present invention a method for determining if a subject is at (high) increased risk to develop obesity or pathology related to obesity according to the invention, wherein the step a) of determining the presence or absence of said three allele risk haplotypes is carried out by a method comprising a step of amplification of the genomic DNA or RNA sequence, preferably DNA sequence, or fragments thereof, susceptible to contain said haplotype SNPs by polymerase chain reaction (PCR).

[0030] In said method wherein the step a) of determining the presence or absence of the three allele risk haplotypes is carried out by a method comprising a step of amplification PCR, the primers used can be all the primers which may be deduced from the nucleotide sequence of the *ENPP1* gene and which may make it possible to detect the presence of said haplotype containing or consisting of the three SNPs K121Q, IVS20 delT-11 and a SNP localized in the 3'UTR domain sequence (the first 1170 bp 3'UTR domain wild type sequence (including the TGA codon) having the sequence SEQ ID No. 15).

[0031] The term "primer" denotes a specific oligonucleotide sequence which is complementary to a target nucleotide sequence and used to hybridize to the target nucleotide sequence. A primer serves as an initiation point for nucleotide polymerization catalyzed by DNA polymerase, RNA polymerase or reverse transcriptase.

[0032] The primers which can be used for the method of the present invention may be designed from the disclosed sequences for any method known in the art. A preferred set of primers is fashioned such that the 3' end of the contiguous span of identity with the target sequence is present at the 3' end of the primer. Such a configuration allows the 3' end of the primer to hybridize to the selected nucleic acid sequence.

[0033] Allele specific primers may be designed such that a polymorphism is at the 3' end of the contiguous span and the contiguous span is present at the 3' end of the primer. Such allele specific primers tend to selectively prime an amplification or sequencing reaction so long as they are used with a nucleic acid sample that contains one of the two alleles present at a polymorphism.

[0034] In a preferred embodiment, said method for determining if a subject is at (high) increased risk to develop obesity or pathology related to obesity according to the invention is characterized in that the step of PCR is carried out from genomic DNA, mRNA or cDNA, eventually after a step of reverse transcription, or fragments thereof, by using the following sets of primers:

- a) 5'-TCATACTCAGGAAGACAGCAA-3' (forward primer SEQ ID No. 2) and 5'-CAATAGCCATGACTCCTAA-3' (reverse primer SEQ ID No. 3) for "K121Q" SNP;
- b) 5'-AGCAT1'TTTACTTTAGTT-3' (forward primer SEQ ID No. 4) and 5'-ATAATGTATGGTAGGATTT-3' (reverse primer SEQ ID No. 5) for IVS20 del T - 11 SNP; and
- c) 5'-ATATTCCTATCCTGCTCACT-3' (forward primer SEQ ID No. 6) and 5'-TGCAGCTGGCCCTTAGGCCG-3' (reverse primer SEQ ID No. 7) for A>G +1044 TGA SNP.

[0035] In a particular embodiment, said method for determining if a subject is at (high) increased risk to develop obesity or pathology related to obesity according to the invention is characterized in that the step of-PCR is carried out from genomic DNA, mRNA or cDNA, eventually after a step of reverse transcription, or fragments thereof, by using the two following sets of primers:

- 5
- a) - 5'-CTTTCCCAATCACTACAGCATTGTCA-3' (forward primer, exon 7, SEQ ID No. 8), and
 - 5'-TTTCAGACCATCCATCAGCATACCAAC-3' (reverse primer exon 12, SEQ ID No. 9); and for the 3'UTR SNP
 - b) - 5'-GTCCTGTGTTTGGACTTTGATTATGA-3' (forward primer exon 23, SEQ ID No. 10), and
 - 5'-CCCTTAGGCCGTTGAAGAATGGTCA-3'(reverse primer 3'UTR, SEQ ID No. 11).
- 10

[0036] In an also preferred embodiment, said method for determining if a subject is at (high) increased risk to develop obesity or pathology related to obesity according to the invention is characterized in that said biological sample is a sample of liver, adipocytes or pancreatic beta-cells when the presence of said haplotype is determined on RNA sequence or fragments thereof.

15 **[0037]** The invention also discloses methods for determining if a subject is at increased risk to develop obesity or pathology related to obesity from a plasma or serum sample from said subject, said method comprising the steps of

- a) determining the concentration of the serum or plasma ENPP1 protein; and
 - b) observing whether or not the subject is at increased risk to develop obesity or pathology related to obesity by observing if said concentration is significantly higher than the level of normal, the presence of said significantly higher concentration indicates said subject is at increased risk to develop obesity or pathology related to obesity.
- 20

[0038] Higher significant concentration of serum or plasma ENPP 1 protein corresponds to at least more than 10 %, preferably more than 15 % of the normal serum or plasma concentration.

25 **[0039]** This invention further describes methods for determining if a subject is at increased risk to develop obesity or pathology related to obesity from a plasma or serum sample from said subject, according to the invention, wherein said concentration of serum or plasma ENPP 1 protein is determined by using antibodies capable of specifically recognizing the ENPP 1 protein or a specific epitopic fragment thereof.

30 **[0040]** In these methods, said antibodies are brought into contact with the serum or plasma to be tested, under conditions allowing the formation of a specific immunological complex between the ENPP1 protein and said antibody, and the immunological complexes formed is quantified.

35 **[0041]** Other aspects encompass a method for determining if a subject is at increased risk to develop obesity or pathology related to obesity from a plasma or serum sample from said subject and from a biological sample containing a genomic DNA or RNA sequence, or fragments thereof, susceptible to contain the *ENPP1* gene three allele risk haplotypes, said method comprising the determination of serum or plasma ENPP1 protein concentration and the observation whether said is serum or plasma ENPP1 protein concentration is significantly higher than the level of normal, and whether said DNA or RNA sequence contains the three allele risk haplotype.

40 **[0042]** In a preferred embodiment for these methods for determining if a subject is at increased(high) risk to develop obesity or pathology related to obesity, the pathology related to obesity is selected from the group consisting of type 2 diabetes, hypertension, hypercholesterolemia and hyperlipidemia particularly type 2 diabetes.

[0043] In another aspect, the present invention also describes a kit for determining if a subject is at increased risk to develop obesity or pathology related to obesity comprising one or more primers or set of primers, or probes, optionally with a portion or all of the necessary reagents and instructions for genotyping a test subject by determining the identity of a nucleotide at the three following SNPs of the *ENPP1* region:

- 45
- K121Q,
 - IVS20 delT-11, and
 - a SNP localized in the 3'UTR domain sequence.

50 **[0044]** The primers or the probes of said kit may optionally be attached to a solid support, or be part of an array or addressable array of polynucleotides. The kit may provide for the determination of the identity of the nucleotide at these three SNPs of the *ENPP1* region position by any method known in the art including, but not limited to, a sequencing assay method, a microsequencing assay method, a hybridization assay method, an allele specific amplification method, or a mismatch detection assay based on polymerases and/or ligases chain reaction.

55 **[0045]** In a preferred embodiment, said kit for determining if a subject is at increased risk to develop obesity or pathology related to obesity, comprising at least one pair or primers or a set of a pair of primers capable of amplifying a fragment of genomic DNA or RNA encoding the protein ENPP1 and susceptible of containing the haplotype comprising the three following SNPs:

- K121Q,
- IVS20 delT-11, and .
- a SNP localized in the 3'UTR domain sequence.

5 **[0046]** In a more preferred embodiment for said kit, the SNP localized in the 3'UTR domain is selected from the group of SNPs consisting of A>G +1044 TGA, A>C +1092 TGA, G>T +1101 TGA, T>C +1137 TGA, C>T +1157 TGA, G>T +1236 TGA, T>C +1348 TGA, G>A +1350 TGA, A>G +1539 TGA and T>C +1670 TGA. Further more preferred for said kit is the SNP localized in the 3'UTR domain is selected from the group of SNPs consisting of A>G +1044 TGA, A>C +1092 TGA and C>T +1157 TGA, the group A>G +1044 TGA and A>C +1092 TGA SNP being also preferred, and A>G +1044 TGA being the most preferred.

10 **[0047]** In a particular embodiment, the invention encompasses a kit for determining if a subject is at increased risk to develop obesity or pathology related to obesity according to the present invention wherein said at least set of a pair of primers is:

- 15 a) 5'-TCATACTCAGGAAGACAGCAA-3' (forward primer SEQ ID No. 2) and 5'-CAATAGCCATGACTCCTAA-3' (reverse primer SEQ ID No. 3) for "K121Q" SNP;
- b) 5'-AGCATTTTTACACTTTAGTT-3' (forward primer SEQ ID No. 4) and 5'-ATAATGTATGGTAGGATTT-3' (reverse primer SEQ ID No. 5) for IVS20 del T - 11 SNP; and
- 20 c) 5'-ATATTCCTATCCTGCTCACT-3' (forward primer SEQ ID No. 6) and 5'-TGCAGCTGGCCCTTAGGCCG-3' (reverse primer SEQ ID No. 7) for A>G +1044 TGA SNP.

[0048] In another particular embodiment, the invention encompasses a kit for determining if a subject is at increased risk to develop obesity or pathology related to obesity according to the present invention wherein said at least set of a pair of primers is:

- 25 a) - 5'-CTTTCCCAATCACTACAGCATTGTCA-3' (forward primer, exon 7, SEQ ID No. 8), and
 - 5'-TTTCAGACCATCCATCAGCATACCAAC-3' (reverse primer exons 12, SEQ ID No. 9); and for the 3'UTR SNP;
- b) - 5'-GTCCTGTGTTTACTTTGATTATGA-3' (forward primer exon 23, SEQ ID No. 10), and
 - 5'-CCCTTAGGCCGTTGAAGAATGGTCA-3'(reverse primer 3'UTR, SEQ ID No. 11).

30 **[0049]** In another aspect, the present invention is directed to an isolated variant nucleic sequence, or to an isolated nucleic acid molecule, or complementary sequence thereof, of a mammal, preferably human, genomic DNA or RNA sequence of the gene coding for ENPP1 protein, characterized in that said isolated variant nucleic sequence comprises the haplotype comprising the three SNPs as defined above K121Q, IVS20 delT-11 and a SNP localized in the 3'UTR domain sequence such as one of the SNPs selected from the group of SNPs consisting of A>G +1044 TGA, A>C +1092 TGA, G>T +1101 TGA, T>C +1137 TGA, C>T +1157 TGA, G>T +1236 TGA, T>C +1348 TGA, G>A +1350 TGA, A>G +1539 TGA and T>C +1670 TGA. The A>G +1044 TGA or A>C +1092 TGA SNP being the more preferred of the SNPs localized in the 3'UTR domain sequence and the SNP A>G +1044 TGA being the most preferred.

35 **[0050]** An "isolated" nucleic acid sequence or molecule is one which is separated from other nucleic acid sequences or molecules which are present in the natural source of the nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the subject from which the nucleic acid is derived. For example, in various embodiments, the isolated variant nucleic sequence of the invention can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. A nucleic acid molecule of the present invention, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the *ENPP1* variant nucleic acid sequence which comprises the haplotype comprising the three SNPs as defined above K121Q, IVS20 delT-11 and a SNP localized in the 3'UTR, the isolated nucleic acid molecules of the present invention can be isolated using standard hybridization and cloning techniques (e.g., as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning. A Laboratory Manual.*, 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989).

45 **[0051]** Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding the isolated variant nucleic sequence of *ENPP1* of the present invention.

50 **[0052]** As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments

can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

[0053] The recombinant expression vectors of the invention can comprise the isolated variant nucleic sequence of the gene coding for ENPP1 protein of the present invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (e.g., in an in vitro transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (e.g., polyadenylation signals). The expression vectors of the invention such as yeast expression vector or baculovirus expression vectors, can be introduced into host cells, in prokaryotic or eukaryotic cell; to thereby produce proteins or peptides. In another embodiment, the expression vector of the present invention is capable of directing expression of the nucleic acid preferentially in mammalian cells or in a particular cell type (e.g., tissue-specific regulatory elements are used to express the nucleic acid).

[0054] Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such term refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

[0055] A host cell can be any prokaryotic or eukaryotic cell. For example, the isolated *ENPP1* variant nucleic sequence of the present invention can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells or human cells). Other suitable host cells are known to those skilled in the art, including *Xenopus laevis* oocytes.

[0056] Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, et al. (Molecular Cloning: A Laboratory Manual, 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989) and other laboratory manuals.

[0057] For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (e.g., resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs.

[0058] The host cells of the invention can also be used to produce non-human transgenic animals. For example, in one embodiment, a host cell, of the invention is a fertilized oocyte or an embryonic stem cell into which the isolated *ENPP1* variant nucleic sequence of the present invention have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous *ENPP1* variant nucleic sequence of the present invention have been introduced into their genome or homologous recombinant animals in which endogenous *ENPP1* nucleic sequence have been altered in order to insert the haplotype variant of the present invention. Such animals are useful for studying the function and/or activity of ENPP1 polypeptide or fragment thereof and for identifying and/or evaluating modulators of ENPP1. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc.. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous *ENPP1* gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into

a cell of the animal, e.g., an embryonic cell of the animal, prior to development of the animal.

[0059] A transgenic animal of the invention can be created by introducing the ENPP1 variant nucleic acid of the present invention into the male pronuclei of a fertilized oocyte, e.g., by microinjection or retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. Alternatively, a nonhuman homologue of the human *ENPP1* gene variant of the present invention, such as a mouse or rat *ENPP1* gene variant, can be used as a transgene. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are well described.

[0060] To create an animal in which a desired nucleic acid has been introduced into the genome via homologous recombination, a vector is prepared which contains at least a portion of a *ENPP1* gene into which the at least three mutation of the haplotype of the present invention has been introduced to thereby alter the *ENPP1* gene.

[0061] So, the present invention relates to transgenic mammal, except human, characterized in that it comprises a transformed cell according to the invention or a cell containing a recombinant functional *ENPP1* gene comprising at least the haplotype comprising the three SNPs as defined above (K121Q; IVS20 delT-11 and a SNP localized in the 3'UTR domain sequence).

[0062] In another aspect, the invention comprises the use of cell or a transgenic mammal according to the present invention, for screening chemical or biochemical compounds for the treatment or the prevention of obesity or pathology related to obesity.

[0063] In a preferred embodiment, said method for selecting a chemical or biochemical compound for the treatment or the prevention of obesity or pathology related to obesity is characterized in that it comprises the step of:

a) bringing into contact a transformed cell or a transgenic mammal according to the invention with said chemical or biochemical compound to be tested; and

b) observing whether the expression product of the ENPP1 gene is decreased or not in said cell or said mammal.

[0064] In a more preferred embodiment, said method for selecting a chemical or biochemical compound for the treatment or the prevention of obesity or pathology related to obesity according to the present invention is characterized in that:

- in step a), the chemical or biochemical compound to be tested is administered to a transgenic mammal to the present invention or to a transgenic mammal overexpressing the ENPP1 protein; and
- in step b), it is observed whether the serum or plasma ENPP1 protein concentration is decreased compared to a reference,

said chemical or biochemical compound being selected if a decrease of said concentration is observed.

[0065] **The present invention also discloses** a method of determining the efficacy of a drug to reduce the risk of obesity or pathology related to obesity in a patient, comprising the steps of

(a) collecting a serum or plasma sample from said patient prior to and subsequent to administering said drug to said patient;

(b) determining the serum or plasma ENPP1 protein concentration in said collected sample, wherein a decrease in the serum or plasma ENPP1 protein concentration subsequent to said treatment is indicative of a treatment that reduces the risk of obesity or pathology related to obesity.

EXAMPLES

[0066] Several of the methods of the present invention are described in the following examples, which are offered by way of illustration and not by way of limitation. Many other modifications and variations of the invention as herein set forth can be made without departing from the spirit and scope thereof.

Legends to figures

[0067]

Figure 1. Pairwise LD between fifty-three SNPs in a 580 kb region including the ENPP1 gene in 421 obese children and 298 control individuals. Regions of high and low LD (Δ) are presented by red and blue shading, respectively. The graph is to the physical map scale.

Figure 2. Allelic association of fifty-three SNPs in the 580 kb region including the ARG1, CRSP3, ENPP3, ENPP1 and CTGF genes. Blue circles correspond to the $-\log_{10}$ (p-value) for the comparison of allelic distribution between

421 obese children and 298 control individuals. Red circles correspond to the $-\log_{10}(\text{p-value})$ for the comparison of allelic distribution between 62 "6q-evidence" families obese children and 298 control individuals. The figure is not to scale according to the physical map location; indeed the SNPs are equidistant to allow easy visualization of the association test results.

Figure 3. Correlation of the ENPP1 protein serum level with Z score of BMI in 279 children.

Figure 4. ENPP1 serum level in 89 lean children according to the presence of the risk haplotype.

Figures 5A and 5B. mRNA expression of PC-1/ENPP1 in human tissues. PCR was performed to detect β -action and ENPP1. Lane 1, Heart; 2 brain; 3, placenta; 4, lung; 5, liver; 6, skeletal muscle; 7, kidney; 8, pancreas (MTC panel) 9, human beta-cells; 10, subcutaneous adipocyte and 11, omental adipocyte. Taking into account that five ENPP1 isoforms were described in databases (<http://www.ncbi.nih.gov/IEB/Research/Acembly/av.cgi?db=human&1=ENPP1>), we first designed primers amplifying a coding region between exon 7 and 12 which is common to at least 3 transcripts, and found ubiquitous expression (figure 5A). We then designed primers specific to the long mRNA isoform, characterized by a larger 3'UTR (1,170 bases downstream the TGA, including the obesity-associated SNP A/G +1044 TGA) and found an expression specific to pancreas and beta-cells, subcutaneous and omental adipocyte and liver (figure 5B).

Example 1: Methods

a) Subjects

[0068] Phenotypic characteristics and type of study performed on each sample are summarized in table 1.

Table 1. Description of the French and Austrian samples used in the study. Sex ratio: female/male

Population	Effective	Age (y.)	BMI (kg/m ²)	Sex ratio	Type of study
Obese children	N=529	10.4±3.9	28.2±6.4	281/248	Association, 6q linkage, TDT, phenotypes, serum
Parents of ob. ch.	N=464	40.7±5.8	31.4±7.7	232/232	Phenotypes
Morbidly obese	N=696	45.8±12.0	47.4±7.4	520/176	Association
Moderately obese	N=717	49.1±14.7	34.6±4.0	401/316	Association
Obese adults	N=325	35.6±12.1	32.3±9.1	134/191	TDT
FrenchT2D	N=752	59.3±11.3	28.1±4.3	379/373	Association
Austrian T2D	N=503	56.7±10.2	30.4±6.3	208/295	Association
Lean children	N=198	14.6±2.2	18.3±2.2	99/99	Association, serum
Control 1	N=556	55±6.0	22±1.8	254/302	Association
Control 2	N=649	50.9±12.7	22.9±2.3	394/255	Association
Austrian control	N=758	52.2±5.8	26.8±4.0	293/465	Association

[0069] Five hundred and twenty nine unrelated obese children were studied. We collected 336 pedigrees with at least one obese child at the CNRS-Institut Pasteur Unit and at the Jeanne de Flandres Hospital in Lille. One hundred and six additional obese children were recruited at the Children's Hospital, Toulouse, and 87 at the Trousseau Hospital, Paris. Children with a BMI greater than the 97th percentile of BMI for age and sex reported on the tables of Rolland-Cachera et al.¹² (French general population) were diagnosed as obese according to the recommendation of the European Childhood Obesity Group (ECOG)⁴². We used a set of 696 unrelated morbidly obese (BMI ≥ 40 kg/m²) and 717 unrelated moderately obese (BMI between 30 and 40 kg/m²) adult patients. We also used for TDT 87 pedigrees with adult obesity (one proband with a BMI of greater than 40 kg/m², at least one additional sibling with a BMI of greater than 27 kg/m² and at least one parent). These individuals were collected at the Department of Nutrition of the Hôtel Dieu Hospital in Paris or at the CNRS-Institut Pasteur Unit in Lille. The 752 T2D subjects were recruited at the Sud Francilien Hospital in Corbeil-Essonnes or at the CNRS-Institut Pasteur Unit in Lille. Glucose Intolerance and type 2 diabetes status were defined according to the WHO 1999 criteria. The first set of 556 control subjects was obtained from the SUVIMAX population⁴³. The second set of 649 control individuals was recruited at the CNRS-Institut Pasteur Unit in Lille and through the "Fleurbaix-Laventie Ville Santé" study⁴⁴. The set of 1,261 unrelated Austrian subjects is a random subset of a previously described population⁴⁵. In order to justify the use of non-age matched controls and to confirm there was no age-cohort effect on SNP and haplotype frequencies, we genotyped the K121Q, IVS20 delT -11 and A>G +1044 TGA polymorphisms in 198 control children from the "Fleurbaix-Laventie Ville Santé" study⁴⁴ and found very concordant frequencies for the Q allele (13.8%) and the risk haplotype (7.8 %) in comparison with control adults. The 458 control

trios (2 parents and one lean child) were issued from the "Fleurbaix-Laventie Ville Santé" cohort. Informed written consent was obtained from all the subjects before participation. The genetic study was approved by the Ethical Committee of Hotel Dieu in Paris and CHRU in Lille. The Z score of BMI was obtained according to Cole's method⁴⁶. During oral glucose tolerance test (OGTT), subjects received after a 12h overnight fast: 1) 1g glucose/kg if subject's weight lower than 50 kg; 2) 75g glucose if subject's weight higher than 50 kg. Blood samples were taken at 0, 30, 60, 90 and 120 min for the measurement of plasma glucose and insulin concentrations. Quantitative measurements of plasma insulin were carried out using double-antibody radioimmunoassays. Serum glucose concentrations were measured using a glucose oxidase procedure. Insulinogenic index was calculated according to Seltzer et al.⁴⁷.

b) Linkage analysis

[0070] In sixty-eight nuclear families, comprising 306 individuals, five polymorphic markers (D6S1720, D6S434, D6S287, D6S1656, D6S292) covering the linkage interval of the initial genome-wide study¹ were used for genotyping. Two-point and multipoint analyses were performed using the MLS test⁴⁸ implemented in the GeneHunter software⁴⁹.

c) Mutation screening

[0071] The twenty-five exons and UTRs (upstream 5' UTR and downstream 3' UTR) were screened using DHPLC (Transgenomic, San Jose, Calif., USA) in 48 unrelated obese children randomly selected among families contributing to the linkage at the chromosome 6q16-q24 locus and in 24 unrelated non-obese normoglycemic control subjects selected from French pedigrees. Variant profiles were sequenced using an automated ABI Prism 3700 DNA sequencer in combination with the Big Dye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems, Foster City, Calif., USA).

d) genotyping

[0072] Genotyping of microsatellite markers was carried out using a fluorescence-based semi-automated technique on automated DNA sequencing machines (ABI 377, PE ABI). Primers were synthesized by MWG Biotech. SNPs were genotyped with the LightCycler™, LightTyper™ (Roche Diagnostics, Basel, Switzerland), with the TaqMan™ (Applied Biosystems, Foster City, Calif. USA) or by direct sequencing. Probes for LightCycler™ and LightTyper™ were synthesized by TIB Molbiol™ Syntheselabor Germany. For coverage of the 580 kb region, SNPs with an allele frequency > 10% in Caucasians were selected using the Applied Biosystem SNP Viewer II software. In order to avoid SNP genotyping errors 10% of DNA samples were systematically re-genotyped for further verification. We found concordance rates of 100% for all SNPs.

e) Statistical analyses

[0073] Comparison of allele frequencies between cases and controls was achieved using the χ^2 test and the p value was empirically computed with the program CLUMP⁵⁰. Haplotype frequencies were determined and were compared between groups with the UNPHASED software <http://www.mrc-bsu.cam.ac.uk/opersonalfrank/>⁵¹. The effect of haplotypes on qualitative or quantitative trait variation was evaluated using the subprograms cocophase and qt-phase of UNPHASED software⁵¹. Independence of association was tested with the software THESIAS³². THESIAS also implements an EM algorithm and allows for likelihood testing of models of haplotype effect in a linear framework. We used this program to test whether the effect of each SNP, on the obesity status, was independent from the effect of K121Q alone. Haplotype analysis was preferred over multiple logistic regression as these SNPs were in partial LD. TDT analyses on SNPs and haplotypes were performed by the TDT method implemented in the UNPHASED software⁵¹. Because of low LD between the three SNPs of the risk haplotype, we only used the unambiguous haplotypes to obtain a true TDT robust to population stratification. In order to evaluate the effect of the risk haplotype on linkage, we used the Genotype IBD Sharing Test (GIST) procedure²³, which tests the correlation between the familial NPL score and an index of the risk haplotype frequency in the family.

f) Linkage disequilibrium analysis

[0074] LD (Linkage disequilibrium) among twenty-four identified ENPP1 SNPs and among fifty-three SNPs in the 580 kb ENPP1 region was investigated. Pairwise delta (correlation coefficient between SNPs) was estimated from genotypes and the results were visualized by the GOLD program (<http://www.sph.umich.edu/csg/abecasis/GOLD/>).

g) ENPP1 mRNA expression

[0075] Human cDNAs from MTC Panel (BD Biosciences Contech), subcutaneous and omental adipocytes (provided by G Fruhbeck, university of Navarra Pamplona Spain) and FACS purified pancreatic beta cells (provided by the Human Pancreatic Cell Core Facility, University Hospital, Lille, France⁵²) were used for mRNA expression analysis. The beta cell purity was confirmed by immunocytochemistry (98% insulin positive cells) and PCR (absence of amplification with chymotrypsin primers, specific for exocrine cells). PCR was performed in a 25 μ l mixture containing 10 mM Tris-HCl, 1.5 mM MgCl₂, 50 mM KCl, 10 mM of each dNTPs, 2.5 U of Taq polymerase (Promega), 30 mM of both forward primer and reverse primer, and 3 μ l of single-strand cDNA. The forward primer for ENPP1 was (SEQ ID No. 8) 5'-CTTTC-CCCAATCACTACAGCATTGTCA-3' (exon 7); and the reverse primer was (SEQ ID No. 9) 5'-TTTCAGACCATCCAT-CAGCATACCAAC-3' (exon 12); and for beta-actin, a forward primer (SEQ ID No. 12: 5'-CGTCATACTCCTGCTTGCT-GATCCACATCTGC-3') and a reverse primer (SEQ ID No. 13: 5'-ATCTGGCACCACACCTTCTACAATGAGCTGCG-3') were used. The ENPP1 isoform, characterized by a larger 3'UTR (1170 bases downstream of the TGA stop codon) was amplified using SEQ ID No. 10 5'-GTCCTGTGTTTGACTTTGATTATGA-3' (exon 23) and (SEQ ID No. 11 5'-CCCTTAG-GCCGTTGAAGAATGGTCA-3' (3'UTR)). The mixtures were heated at 95°C for 2 min and subjected to 45-cycle amplification at 95°C for 30 s, 68°C for 2 min., and 72°C for 2 min, and then 10 min at 72°C. PCR products were separated on 2 % (w/v) agarose gel and visualized using ethidium bromide and UV transillumination.

h) ENPP1 protein serum assays

[0076] ENPP1 serum level was measured using the ELISA described by Rutsch et al.²⁴.

Example 2: Search for positional candidate genes on chromosome 6q163-q24.2

[0077] The phenotypic characteristics of 62 "6q-evidence" families (defined by an individual pedigree Gene Hunter Zscore > 1.0 in the 2-LOD drop interval flanked by the micro-satellite markers D6S434 and D6S1704) were first compared with the remaining 35 families, all issued from our previously published genome scan for childhood obesity' .

Table 2. Phenotypic characteristics of "6q-evidence" families

Trait	Obese children from "6q-evidence" families	Obese children from other families	p-value of the T-test or Chi2 test
Age	11.52 \pm 3.15 (N=138)	11.54 \pm 3.41 (N=70)	0.95
Gender (M/F)	63/75 (N=138)	40/30 (N=70)	0.11
ZBMI	4.20 \pm 0.11 (N=138)	3.74 \pm 0.1 (N=70)	0.004
AUGC	773.04 \pm 17.32 (N=95)	709.71 \pm 17.30 (N=50)	0.07
Ins Ind	25.44 \pm 3.01 (N=71)	41.33 \pm 7.67 (N=30)	0.02
% IG/T2D	3.1% (N=131)	0% (N=69)	0.18
Trait	Parents from "6q-evidence" families	Parents from other families	p-value of the Chi2 test
% T2D	13.8 % (N=123)	3.2 % (N=62)	0.018
ZBMI: Z score of BMI. AUGC: area under the glycemia curve during 2h oral glucose tolerance test. Ins Ind: Insulinogenic Index (mU.mmol ⁻¹). %IG/T2D: percentage of obese children displaying glucose intolerance or type 2 diabetes mellitus. %T2D: percentage of parents displaying type 2 diabetes mellitus.			

[0078] As shown in Table 2, the "6q-evidence" obese children have a significantly higher Z score of BMI (deviation of the body mass index compared to a French reference population¹²). After adjustment of parameters for sex, age and BMI, their blood glucose levels after glucose administration is higher, and their insulinogenic index, which estimates the first-phase insulin secretion, is significantly lower than in children of other families. Moreover, 3.1 % of the "6q-evidence" obese children are either glucose intolerant or diabetic: in contrast, none of the obese children in other families is glucose intolerant. We also investigated the previous generation, and found that 13.8% of 6q linked parents have type 2 diabetes mellitus (T2D) as compared to 3.2% of parents in other families (p=0.02). These data led us to conclude that the obesity susceptibility gene(s) on chromosome 6q may be also involved in glucose homeostasis.

[0079] Based on our genome scan for childhood obesity¹ the 2-LOD drop interval for the 6q16.3-q24.2 linkage peak

covers 41.4 Mb and includes 166 referenced genes. In order to focus the search for obesity susceptibility genes to a more restricted chromosomal area, the peaks of 7 published reports of linkage on chromosome 6q16.1-q27 with either obesity², insulin secretion^{3,4} or T2D⁵⁻⁸, all performed in adults, were analyzed. A 2.4 Mb interval was identified as common to all of the genome scans, spanning the region between markers D6S1656 and D6S270. Among the twenty

two genes mapping within this interval, the best physiological candidate was ectonucleotide pyrophosphatase/phosphodiesterase ENPP1 (also known as the Plasma Cell glycoprotein-1 PC-1) (OMIM reference 173335). ENPP1 is believed to directly inhibit insulin-induced conformational changes of the insulin receptor, thereby affecting its activation and downstream signaling^{9,13}. In this regard, a functional missense mutation (K121Q), which was proposed to more effectively inhibit insulin signalling¹¹, was associated with insulin resistance or T2D in various ethnic groups¹⁴⁻¹⁸.

[0080] The microsatellite marker D6S1656 in intron 1 of the ENPP1 gene, linked to childhood obesity in our initial genome scan, was analyzed in a second replication set of 68 families having at least two sibs with BMI > 95th percentile (corrected for gender and age) and modest evidence of linkage (MLS=0.83, p=0.04) was observed. In contrast, no linkage with obesity was found with 4 other microsatellite markers of this region in the replication set (data not shown). Allele 10 of D6S1656 was also significantly under-transmitted to affected children in both the initial and replication sets (44 transmitted vs 69 untransmitted, p=0.02, data not shown). These results suggested that functional polymorphisms in partial linkage disequilibrium (LD) may be located nearby.

Example 3: Sequence variation within the *ENPP1* gene and association with obesity due to ENPP1 SNPs

[0081] ENPP1 has 25 exons and spans over 83.2 Kb. In order to characterise sequence variation within the gene, sequencing of all coding regions, and 1.3 Kb upstream of the ATG start codon and downstream of the TGA stop codon was carried out in 48 obese children from "6q-evidence" families and in 24 non-obese adults. Eight single-nucleotide polymorphisms (SNPs) were identified in the upstream sequence, six in intron/exon junctions, four were missense, three were synonymous mutations and twenty were identified in the downstream sequence (see table 3).

Table 3. Polymorphisms in 5' and 3' proximal regions, exons and boundaries of ENPP1 gene

SNP	Description	Variant allele frequency (%)
1	C>T -1125 ATG	0.7
2	T>C -931 ATG	0.7
3 (rs 1800949)	C>T -768 ATG	20.7
4	G>C -406 ATG	0.7
5	C>T -359 ATG	1.4
6	G>T -190 ATG	7.7
7 (rs12212106)	C>G -187 ATG	20.4
8	C>T -165 ATG	4.5
9	IVS2 delG +8	6.8
10 (rs 1044498)	K121Q	16.9
11	IVSS T>G -20	2.8
12	IVS7 T>G -28	0.7
13 (rs9493113)	IVS8 T>G +27	8.3
14	IVS9 G>C +24	2.1
15 (rs9483347)	K387K	0.7
16	IVS20 delT-11	25.3
17 (rs7750837)	S650S	0.7
18	R722C	1.5
19 (rs1805138)	T727N	0.7
20 (rs8192683)	R834T	0.7
21 (rs1804025)	A835A	3.6
22 (rs1044548)	G>A +112 TGA	7.6
23 (rs11964389)	G>C +121 TGA	3.5
24 (rs1044558)	C>T +164 TGA	7.6
25	G>A +395 TGA	7.8
26 (rs1044582)	T>A +457 TGA	7.8
27	InsAA DelT +700 TGA	7.7

(continued)

SNP	Description	Variant allele frequency (%)
28(rs12212882)	A>G +735 TGA	4.2
29	A>T +772 TGA	8.4
30	C>T +777 TGA	7.7
31	Ins 13bp +828 TGA	4.2
32 (rs7754561)	A>G+1044 TGA	26.1
33 (rs7754586)	A>C +1092 TGA	21.6
34 (rs 11154647)	G>T +1101 TGA	5.9
35 (rs11154648)	T>C +1137 TGA	6
36 (rs7754859)	C>T +1157 TGA	27.6
37 (rs11154649)	G>T +1236 TGA	6
38 (rs9483349)	T>C +1348 TGA	3.6
39 (rs9493120)	G>A +1350 TGA	7.1
40 (rs9493121)	A>G +1539 TGA	7.1
41 (rs9493122)	T>C +1670 TGA	4.3

[0082] Among these forty-one polymorphisms, twenty-four were present in the public databases and twenty-two of the variants had a minor allele frequency (MAF) higher than 5 %. For the nineteen infrequent polymorphisms, their transmission was inspected within pedigrees. When the rare allele was present in unaffected sibs the variant was not investigated further (n=14). The remaining five rare SNPs did not show any significant association with obesity in case-control analyses (data not shown). Pairwise LD among the twenty-two most common SNPs, plus two others: SNP23 (3.5 %), previously described as influencing mRNA stability¹⁹ and a 13 bp insertion (SNP31, Ins 13bp +828 TGA, prevalence 4.2 %) was used to select a set of ten "haplotype tagging" SNPs to be typed in the whole set of samples.

[0083] To further enhance the chances of identifying functional SNPs, twenty-five fragments were identified that were more than 100bp long, showing a high degree of homology (>70 %) across Fugu rubripes, rat and human genomes and within the 198 Kb of intronic and 5' and 3' intergenic regions of the ENPP1 gene. These fragments were sequenced and a further eleven SNPs of MAF >5% were identified. These SNPs were initially used to genotype a test set of 421 obese children and 298 control individuals. One SNP in the 3' region of ENPP1 (T>G +5954 TGA) that showed a trend for association with childhood obesity (p<0.1) was added to the set of SNPs analyzed in the whole sample set.

[0084] Altogether, eleven SNPs were genotyped in 2,430 individuals, made up of 529 unrelated obese children, 696 unrelated morbidly obese adults, and 1,205 lean normoglycaemic adults. Lean adults were used as controls for both sets of cases because they demonstrated a long term resistance against obesity. As shown in Table 4, we observed associations between severe forms of obesity and six of the SNPs: IVS2 delG +8, K121Q, IVS8 T>G +27, Ins20 delT -11, A>G +1044 TGA and T>G +5954 TGA (see Table 4, 0.00008<p<0.03; 1.21>OR>1.37). The global p-value, assessed by 10⁵ permutations of the obesity status among individuals, was 0.001. Analysis of the pooled data identified the strongest association with severe forms of obesity of the K121Q SNP (OR+1.37, 95% confidence interval [1.17-1.61], p=0.00008). The odds ratio under a recessive model was increased at 3.29 [1.83 - 5.93] (p=0.00003) with a significant departure from the additive model (p=0.02).

[0085] To test this association further, 184 multiplex families were genotyped for this SNP. This set consisted of the 97 childhood obesity genome scan families plus 87 nuclear families with adult severe obesity, which showed linkage of serum leptin levels to chromosome 6q24²⁰. Using the Transmission Disequilibrium Test (TDT) the 121 Q-allele was significantly over-transmitted to obese offspring (transmitted: 76, non-transmitted: 48, p=0.01), supporting the case-control result.

[0086] To confirm ENPP1 specificity for these associations, a total of fifty-three SNPs in the chromosome 6q region, spanning 580 kb and including the ARG1, CRSP3, ENPP3, ENPP1, and CTGF genes, were typed in the initial set of 421 obese children and 298 control individuals used above (average density: 1 SNP/10.9 kb). Using the GOLD software, three distinct regions of linkage disequilibrium (LD) were noted; the first one contained ARG1, CRSP3 and ENPP3, the second one ENPP1 alone, and the third the CTGF gene and 176Kb of non coding region (figure 1). Comparison of the allelic distribution of the fifty-three SNPs between obese children and controls, and also between sixty-two "6q-evidence" obese children and controls revealed that eight SNPs were associated with both childhood obesity and "6q-evidence" childhood obesity (p<0.05). Seven of the eight SNPs (K121Q, Celera dbSNP hcV1207989, C>T +164 TGA, Celera dbSNP hcV1207974, A>G +1044 TGA, G>T +1101 TGA, C>T +1157 TGA) mapped within ENPP1 (figure 2). The eighth associated SNP was located in an intergenic region, 69 kb 3' of CTGF, and not located in a fragment showing a high degree of homology across Fugu rubripes, rodent and human genomes and with no known transcription factor binding

EP 1 891 237 B9

site. These data support the hypothesis that the observed association with childhood obesity is due to ENPP1 SNPs.

Table 4: Positive association ($p < 0.05$) of genotypes and alleles of ENPP1 gene SNPs with obesity. Cases Set 1 = 529 French Caucasian children with BMI higher than the obesity threshold of the 97th percentile. Control Set 1 = 556 non obese and normoglycemic French Caucasian adults. Cases Set 2 = 696 French Caucasian adults with BMI \geq 40. Control Set 2 = 649 non obese and normoglycemic French Caucasian adults.

Cohorts	Genotypes n (frequency)			Allele	Co-dom.	Dominant	Recessive	HW	
	CC	CT	TT	freq. OR (p-val.)	OR (p-val.)	OR (p-val.)	OR (p-val.)	(p-val.)	
C>T-768ATG									
Set 1	Obese children	310 (0.65)	145 (0.30)	23 (0.05)	0.91 (0.42)	0.91 (0.44)	0.93 (0.58)	0.77 (0.37)	0.25
	Control 1	312 (0.63)	152 (0.31)	30 (0.06)					0.06
Set 2	Morbidly obese	385 (0.60)	225 (0.35)	30 (0.05)	1.05 (0.59)	1.05 (0.59)	1.07 (0.58)	1.08 (0.77)	0.82
	Control 2	375 (0.62)	206 (0.34)	27 (0.04)					0.90
Set 1+2	Obese	695 (0.62)	370 (0.33)	53 (0.05)	0.99 (0.92)	0.98 (0.92)	1.01 (0.93)	0.92 (0.67)	0.66
	Control 1+2	687 (0.62)	358 (0.33)	57 (0.05)					0.24
G>T-190ATG									
Set 1	Obese children	436 (0.88)	59 (0.12)	1 (0.00)	1.26 (0.24)	1.25 (0.24)	1.28 (0.22)	1.05 (0.97)	1
	Control 1	456 (0.90)	48 (0.10)	1 (0.00)					1
Set 2	Morbidly obese	589 (0.90)	63 (0.10)	3 (0.00)	0.77 (0.12)	0.79 (0.13)	0.75 (0.11)	0.88 (0.87)	0.41
	Control 2	518 (0.87)	74 (0.12)	3 (0.01)					0.74
Set 1+2	Obese	1025 (0.89)	122 (0.11)	4 (0.00)	0.95 (0.71)	0.95 (0.71)	0.95 (0.70)	0.95 (0.94)	0.78
	Control 1+2	974 (0.89)	122 (0.11)	4 (0.00)					0.79
IVS2delG+8									
Set 1	Obese children	378 (0.88)	48 (0.11)	2 (0.01)	1.34 (0.14)	1.54 (0.14)	1.31 (0.20)	6.55 (0.11)	0.66
	Control 1	495 (0.91)	50 (0.09)	0 (0.00)					0.62
Set 2	Morbidly obese	610 (0.89)	74 (0.11)	3 (0.00)	1.32 (0.12)	1.34 (0.12)	1.31 (0.15)	2.71 (0.37)	0.49
	Control 2	551 (0.91)	52 (0.09)	1 (0.00)					1
Set 1+2	Obese	988 (0.89)	122 (0.11)	5 (0.00)	1.33 (0.03)	1.39 (0.03)	1.30 (0.05)	5.29 (0.09)	0.58
	Control 1+2	1046 (0.91)	102 (0.09)	1 (0.00)					0.72

EP 1 891 237 B9

(continued)

	K121Q	KK	KQ	QQ						
5	Set 1	Obese children	351 (0.67)	153 (0.29)	22 (0.04)	1.45 (0.001)	1.56 (0.001)	1.41 (0.01)	3.63 (0.001)	0.31
		Control 1	405 (0.74)	136 (0.25)	7 (0.01)					0.28
10	Set 2	Morbidly obese	479 (0.70)	176 (0.26)	25 (0.04)	1.31 (0.01)	1.41 (0.01)	1.26 (0.07)	3.05 (0.004)	0.09
		Control 2	467 (0.75)	148 (0.24)	8 (0.01)					0.38
15	Set 1+2	Obese	830 (0.69)	329 (0.27)	47 (0.04)	1.37 (0.00008)	1.48 (0.00009)	1.32 (0.002)	3.29 (0.00003)	0.06
		Control 1+2	872 (0.75)	284 (0.24)	15 (0.01)					0.16
		NSBT>G+27		TT	TG	GG				
20	Set 1	Obese children	377 (0.78)	105 (0.22)	2 (0.00)	1.35 (0.04)	1.40 (0.03)	1.39 (0.04)	2.26 (0.49)	0.07
		Control 1	427 (0.83)	86 (0.17)	1 (0.00)					0.16
25	Set 2	Morbidly obese	539 (0.81)	118 (0.18)	9 (0.01)	1.26 (0.09)	1.32 (0.09)	1.23 (0.15)	2.83 (0.10)	0.39
		Control 2	508 (0.84)	94 (0.16)	3 (0.00)					0.79
30	Set 1+2	Obese	916 (0.80)	223 (0.19)	11 (0.01)	1.30 (0.009)	1.35 (0.009)	1.30 (0.01)	2.81 (0.06)	0.64
		Control 1+2	935 (0.84)	180 (0.16)	4 (0.00)					0.17
		IVS20de1T-11		T/T	T/delT	delT/ delT				
35	Set 1	Obese children	293 (0.56)	190 (0.36)	43 (0.08)	1.10 (0.32)	1.20 (0.31)	0.99 (0.92)	1.96 (0.01)	0.14
		Control 1	307 (0.55)	224 (0.40)	23 (0.05)					0.03
40	Set 2	Morbidly obese	382 (0.55)	255 (0.37)	57 (0.08)	1.31 (0.002)	1.32 (0.003)	1.31 (0.01)	1.85 (0.007)	0.12
		Control 2	396 (0.62)	214 (0.33)	32 (0.05)					0.64
45	Set 1+2	Obese	675 (0.55)	445 (0.37)	100 (0.08)	1.21 (0.004)	1.26 (0.004)	1.15 (0.09)	1.89 (0.0002)	0.03
		Control 1+2	703 (0.59)	438 (0.36)	55 (0.05)					0.22
		G>C+121TGA		GG	GC	CC				
50	Set 1	Obese children	443 (0.96)	17 (0.04)	1 (0.00)	0.86 (0.63)	0.87 (0.64)	0.85 (0.63)	0.89 (0.94)	0.17
		Control 1	398 (0.95)	18 (0.04)	1 (0.01)					0.21

55

EP 1 891 237 B9

(continued)

		G>C+121TGA	GG	GC	CC					
5	Set 2	Morbidly obese	601 (0.94)	35 (0.06)	2 (0.00)	1.62 (0.09)	1.88 (0.09)	1.55 (0.13)	3.98 (0.21)	0.11
		Control 2	478 (0.96)	19 (0.04)	0 (0.00)					1
10	Set 1+2	Obese	1044 (0.95)	52 (0.05)	3 (0.00)	1.24 (0.30)	1.26 (0.31)	1.21 (0.37)	2.52 (0.41)	0.05
		Control 1+2	876 (0.96)	37 (0.04)	1 (0.00)					0.34
		C>T+164TGA	CC	CT	TT					
15	Set 1	Obese children	436 (0.89)	55 (0.11)	1 (0.00)	0.99 (0.97)	0.99 (0.97)	0.99 (0.96)	1.10 (0.95)	1
		Control 1	478 (0.89)	61 (0.11)	1 (0.00)					1
20	Set 2	Morbidly obese	607 (0.87)	83 (0.12)	6 (0.01)	1.40 (0.04)	1.49 (0.05)	1.35 (0.09)	5.54 (0.07)	0.12
		Control 2	561 (0.90)	60 (0.10)	1 (0.00)					
25	Set 1+2	Obese	1043 (0.88)	138 (0.12)	7 (0.00)	1.20 (0.14)	1.25 (0.14)	1.17 (0.22)	3.49(0.1)	0.32
		Control 1+2	1039 (0.90)	121 (0.10)	2 (0.00)					0.77
		Ins+828TGA	-/-	-/Ins	Ins/Ins					
30	Set 1	Obese children	454 (0.90)	50 (0.10)	2 (0.00)	1.17 (0.45)	1.15 (0.45)	1.19 (0.42)	1.00 (1)	0.64
		Control 1	456 (0.91)	42 (0.09)	2 (0.00)					0.28
35	Set 2	Morbidly obese	586 (0.90)	65 (0.10)	2 (0.00)	0.87 (0.40)	0.87 (0.40)	0.86 (0.39)	0.93 (0.94)	0.7
		Control 2	547 (0.88)	71 (0.12)	2 (0.00)					
40	Set 1+2	Obese	1040 (0.90)	115 (0.10)	4 (0.00)	0.98 (0.89)	0.98 (0.89)	0.98 (0.88)	0.96 (0.96)	0.56
		Control 1+2	1003 (0.90)	113 (0.10)	4 (0.00)					0.56
		A>G+1044TGA	AA	AG	GG					
45	Set 1	Obese children	260 (0.49)	213 (0.40)	56 (0.11)	1.09 (0.38)	1.14 (0.38)	0.99 (0.94)	1.52 (0.06)	0.22
		Control 1	269 (0.49)	243 (0.44)	38 (0.07)					0.1
50	Set 2	Morbidly obese	328 (0.50)	257 (0.39)	77 (0.11)	1.36 (0.0004)	1.36 (0.0006)	1.36 (0.006)	1.97 (0.0008)	0.02
		Control 2	369 (0.57)	233 (0.36)	44 (0.07)					0.40
55	Set 1+2	Obese	588 (0.49)	470 (0.40)	133 (0.11)	1.22 (0.002)	1.26 (0.002)	1.17 (0.05)	1.76 (0.0002)	0.01 1

(continued)

		A>G+1044TGA	AA	AG	GG				
5	Control 1+2	638 (0.53)	476 (0.40)	82 (0.07)					0.66
		T>G +5954 TGA	TT	TG	GG				
10	Set 1 Obese children	366 (0.77)	100 (0.21)	10 (0.02)	1.43 (0.01)	1.47 (0.02)	1.41 (0.03)	2.9 (0.06)	0.30
	Control 1	426 (0.82)	87 (0.17)	4 (0.01)					1
15	Set 2 Morbidly obese	532 (0.79)	126 (0.19)	11 (0.02)	1.24 (0.09)	1.26 (0.10)	1.23 (0.14)	1.85 (0.22)	0.24
	Control 2	536 (0.83)	106 (0.16)	6 (0.01)					0.81
	Set 1+2 Obese	898 (0.78)	226 (0.20)	21 (0.02)	1.32 (0.004)	1.34 (0.005)	1.30 (0.01)	2.25 (0.03)	0.15
20	Control 1+2	962 (0.83)	193 (0.16)	10 (0.01)					0.86

Example 4: Characterization of a three allele risk haplotype strongly associated with severe forms of obesity

[0087] A two SNP analysis between the K 121 Q polymorphism and the 6 other obesity-associated SNPs was used to assess whether these SNPs had an independent effect on the risk of obesity. A likelihood ratio test showed that only IVS20de1T-1 and G+1044TGA SNPs significantly modulate the effect of K121Q (pb=0.03 and p=0.04 respectively). It is worth pointing out that for these two SNPs, the model with the best fit was also a recessive one. This may account for the small observed deviation from Hardy-Weinberg Equilibrium (HWE) for these SNPs as genotyping errors have been ruled out by resequencing (data not shown). Such a deviation can be expected in affected "extreme" individuals when the genetic model is not strictly additive ²¹.

[0088] In order to estimate the potential cumulative effect of combinations of the three SNPs on the risk of obesity, haplotype analysis was performed using the K121Q, IVS20 delT-11 and A>G +1044 TGA SNP data from the whole set of 2,430 French Caucasian subjects. Eight haplotypes were predicted, five of them having a MAF >5 %. As shown in Table 5, the three allele wild-type haplotype (K121Q/IVS20 delT-11/A>G +1044 TGA, KTA), was less frequent in obese subjects than in controls (60.3 % vs 64.0 %, p=0.002). In contrast, the three allele risk haplotype (K121Q/IVS20 delT-11/A>G +1044 TGA, QdelTG), was strongly associated with severe forms of obesity (11.2 % vs 7.5 %, OR=1.58, p=0.00001, empirical p-value<0.0001 for 10⁵ simulations). Interestingly, the risk haplotype effect was of similar magnitude in both morbidly obese adults and childhood obesity (10.8 % vs 7.9 %, OR=1.50, p=0.006, and 11.7 % vs 7.1 %, OR=1.69, p=0.0006, respectively). The association was also supported by TDT analysis in the 184 families (Chi2=5.68, p=0.01, Table 5).

Table 5. Haplotype analysis of 1225 morbidly obese/obese children and 1205 control subjects. Haplotypes comprising three polymorphisms (K121Q, IVS20 delT-11, A>G+1044 TGA) are shown. TDT was performed for each haplotype in 97 pedigrees with childhood obesity and in 87 pedigrees with adult obesity .

Haplotypes			Case/control Test				TDT in 184 obesity pedigrees	
K121Q	IVS20 delT-11	A>G+1044TG A	Non obese	Obese	Chi-square	p-value	Chi-square	p-value
K	T	A	64.0	60.3	8.99	0.003	2.7	0.1
		G	7.0	7.6	0.51	0.47	0.09	0.75
	DelT	A	3.5	2.7	1.26	0.26	0	1
		G	11.7	11.9	0.18	0.67	0.73	0.39

(continued)

Haplotypes			Case/control Test				TDT in 184 obesity pedigrees	
K121Q	IVS20 delT-11	A>G+1044TGA	Non obese	Obese	Chi-square	p-value	Chi-square	p-value
Q	T	A	5.5	4.8	0.15	0.70	2.30	0.13
		G	0.6	0.6	0.19	0.66	0	1
	DelT	A	0.2	0.9	5.23	0.02	0.20	0.65
		G	7.5	11.2	18.78	0.00001	5.68	0.01

[0089] TDT analysis in 458 french Caucasian control trios (2 parents and one lean child) didn't show an excess of transmission of the risk haplotype ($\text{Chi}^2=0.53$, $p=0.46$, data not shown) and excluded a transmission ratio distortion of the risk haplotype in unaffected children²². Finally, we evaluated the effect of the haplotypes in an additional set of 717 adult subjects with a less severe form of obesity (BMI between 30 and 40 kg/m²), compared against the 649 non-obese normoglycemic control subjects used previously. Again, the risk haplotype was associated with obesity (10.3 % vs 7.9 %, $\text{OR}=1.37$, $p=0.02$).

[0090] Several approaches were used to test the impact of the ENPP1 risk haplotype on the linkage with childhood obesity observed in the genome scan. We first detected a higher frequency of the risk haplotype in individuals from "6q-evidence" families (16.2 % vs 7.1 %, $\text{OR}=2.37$, $p=0.004$) compared to other families (12.2 % vs 7.1 %, $\text{OR}=1.65$, $p=0.16$). The affected sib-pairs sharing the obesity-associated risk haplotype (15 from a total of 135), when analyzed separately, gave a maximal multipoint MLS LOD score of 3.49 at the position of the initial maximum linkage (marker D6S287). Conversely, after removing these 15 affected sib-pairs, the multipoint MLS LOD score dropped from 4.06 to 1.6 at marker D6S287, and a new maximal MLS LOD score of 2.63 appeared 16-Mb centromeric to the original peak of linkage, at the marker D6S301. To evaluate the significance of these variations, the Genotype IBD Sharing Test (GIST) was used²³. This test suggested a trend to for a possible effect of the haplotype under an additive model ($p=0.07$) which became significant for a recessive model ($p=0.03$), but this was based upon only three affected sib-pairs sharing two copies of the haplotype. Altogether, these data suggest some contribution of the ENPP1 obesity risk haplotype to the observed linkage with childhood obesity on chromosome 6q, though the GIST result demonstrates that it is relatively modest.

[0091] According to these data, at least three ENPP1 SNPs are involved in the association with obesity. The Q121 variant is believed to inhibit insulin signalling¹¹ more effectively than the wild-type version, but the functional effects of the risk haplotype are unknown.

Example 5: Highly significant increase of ENPP1 levels associated with K121Q, IVS20delT-11 and G+1044TGA alleles

[0092] The ENPP1 protein has a proteolytic cleavage site, is cleaved at the surface of cells, and is known to be present in the circulation²⁴. Thus, the protein serum level represents a good estimation of its tissue expression²⁴. Serum ENPP1 protein levels measured in 279 children encompassing a wide weight range showed a positive correlation with the Z score of BMI (Pearson correlation coefficient=0.1, $p=0.05$, Figure 3). Eighty-nine lean children were then selected (mean BMI, 18.4 ± 2.5 kg/m²; mean age, 13.3 ± 2.6 y; $n=50/39$ girl/boy), to fix the confounder BMI, and analyzed for the effect of the three SNPs on ENPP1 levels. The presence of at least one copy of Q121, IVS20delT-11 and G+1044TGA alleles was associated with a highly significant increase of ENPP1 levels (28.6 ng/ml vs 24.1 ng/ml, $p=0.008$) (Figure 4), suggesting that the obesity-associated haplotype not only impairs insulin binding but also enhances ENPP1 levels of expression.

[0093] The inventors then assessed the contribution of the risk haplotype on the variation of obesity-related phenotypes in 474 obese children where data were available. Obese children with the obesity-associated risk haplotype showed a 0.17 mmol/l increase in fasting glycemia ($p=0.002$) with a higher prevalence of glucose intolerance/T2D ($\text{OR}=3.43$, $p=0.02$). Parents carrying the risk haplotype had a 2.35 fold increased risk to develop T2D ($p=0.005$). Moreover, the risk was higher in the subset of obese parents ($\text{OR}=3.26$, $p=0.0005$) whereas no increase in the risk of T2D was observed in the non obese parents carrying the risk haplotype ($\text{OR}=0.86$, $p=0.9$). To confirm the association of the ENPP1 SNPs with T2D, an additional non overlapping cohort of 752 unrelated T2D French Caucasian subjects with familial history of the disease was compared to the previously used 556 middle aged non obese normoglycemic subjects (average age: 55 ± 6 years). A significant excess of the risk haplotype was identified in the T2D group (10.7% vs 7.1%, $\text{OR}=1.44$,

p=0.005), further supporting a potential effect of ENPP1 SNPs on glucose homeostasis in French Caucasians. This finding was then replicated using 1261 unrelated Austrian subjects consisting of 503 T2D subjects and 758 non obese normoglycemic subjects (9.8 % vs 6.3 %, OR=1.68, p=0.001). Because of potential differences in allele frequency and effect size between the French and Austrian cohorts, the Mantel Haenszel adjusted odds ratio, under a fixed effects model, was used to evaluate the risk of T2D in the pooled cohorts. The pooled data, with 2569 individuals of European origin, strengthened the association of the risk haplotype with T2D (10.4 % vs 6.6%, combined OR=1.56, p=0.00002). In summary, these findings indicate that both obesity and T2D, especially in obese subjects, are associated with ENPP1 genetic variability and specifically, a single three allele risk haplotype. This provides genetic evidence for the recently described link between obesity in childhood and the high risk for T2D in the teens or early adulthood²⁵, providing the first common molecular mechanism for this deleterious association.

[0094] The tissue distribution of ENPP1 transcripts was examined next. RT-PCR was performed on cDNAs from a wide range of human tissues including brain, muscle, liver, adipocyte (subcutaneous and omental), and purified pancreatic islet beta-cells. Taking into account that five ENPP1 isoforms are known, (<http://www.ncbi.nih.gov/IEB/Research/Acm-bly/av.cgi?db=human&1=ENPP1>), we designed primers to amplify the region between exons 7 and 12, which is common to at least 3 transcripts, and found ubiquitous expression (see figure 5). Primers were then designed that were specific to the long mRNA isoform, characterized by a larger 3'UTR with 1170 bases downstream of the TGA stop codon and including the obesity-associated SNP A>G +1044 TGA. The long form was found to be only expressed in pancreatic β -cells, adipocytes and liver, three key tissues for glucose homeostasis (see supplementary figure 1).

[0095] This study provides the first evidence for a primary role of variants of ENPP1 gene variants in the development of both childhood and adult morbid and less severe forms of obesity and strong evidence for an association with T2D that could be increased by the concurrent presence of obesity. This extensive analysis of the whole ENPP1 locus in the context of metabolic diseases revealed that this effect is mainly the result of a newly described three allele risk haplotype that was a more potent predictor than individual SNPs in all the studied populations.

[0096] The likely contribution of the obesity-associated ENPP1 risk haplotype to our observed linkage with childhood obesity on 6q was assessed by several methods including the recently described Genotype IBD Sharing Test (GIST) procedure²³. A significantly higher frequency of the risk haplotype was observed in affected individuals from "6q-evidence" families together with a significant correlation between the haplotype specific index and the NPL score in the GIS test. In this regard, Jenkinson et al. recently presented preliminary data which support a contribution of ENPP1 SNPs²⁶ to the linkage for insulin fasting levels observed in Mexican-Americans³. As we also found evidence for association between ENPP1 SNPs and T2D, this gene may also contribute to the linkage peaks with T2D and related traits found in other populations³⁻⁸. However, in spite of the strong association of the three allele haplotype with severe forms of obesity in the French Caucasian population (OR=1.58, p=0.00001) and the evidence for its contribution to the 6q linkage signal, we only found a moderate excess of transmission of this haplotype to affected offspring as measured by the TDT (p=0.01). This suggests that the risk haplotype is necessary for the linkage, thus showing a higher frequency in "linked" individuals, but not sufficient to fully explain it. That implies other contributing sequences in the 6q linked region.

[0097] It is plausible that additional SNPs in the non-coding regions of the ENPP1 locus account for a part of the observed linkage. For instance, the SNP T>G +5954 TGA, included in a highly conserved region across Fugu rubripes, rat and human genomes, showed association with severe forms of obesity. Using Genomatix (<http://www.genomatix.de/>) we found that the polymorphism was part of a sequence (aaaaaattcTT(>G)ATgacaccta) recognized by the Insulin Promoter Factor IPF1²⁷. The substitution from T to G predicts loss of binding IPF1, suggesting that this SNP could be functional. Alternatively, more than one gene may explain the linkage with obesity and T2D on 6q. Such a situation has been already described on chromosome 16 for Crohn's disease^{28,29}. In this regard, the 16 Mb centromeric linkage peak shift when ENPP1 risk obesity-associated haplotype carriers are discarded coincides with the location of Single-minded 1 (SIM1), another obesity candidate gene^{30,31}.

[0098] The putative functionality of the three obesity-associated SNPs was primarily assessed by statistical means, including a likelihood ratio test of haplotype effects, using THESIAS software³². Among these SNPs, the ENPP1 K121Q exon 4 missense mutation has been previously associated with insulin resistance in rather small size samples from Sicily, Scandinavia and India¹⁴⁻¹⁷, and with T2D in Dominicans¹⁸, but has never been confirmed in large scale studies. The present study involves a total of 6,147 subjects, all of European origin, and confirms the T2D findings and expands them to include obesity. The other two strongly obesity-associated SNPs are located in the exon 21 boundary (IVS20de1T-11) and in the 3'UTR (A>G +1044 TGA). Interestingly, the latter belongs to an isoform that is specifically expressed in three highly insulin-responsive human tissues (pancreatic islet β -cell, adipocyte and liver) which are thought to play a major role in T2D pathophysiology. Highlighting the importance of ENPP1 action in the liver, mice given an adenovirus expression construct overexpressing this gene in hepatocytes show insulin resistance and glucose intolerance¹⁰. Although the exonic 121 Q amino acid substitution was proposed to directly inhibit insulin receptor by a non enzymatic mechanism¹¹, the other non-coding SNPs may have their effect by modifying gene expression, protein production or splicing, both increasing the protein's putative deleterious effect on insulin signaling. This hypothesis is favoured by the increased serum protein levels in children carrying the ENPP1 obesity risk haplotype and also by the rise of ENPP1

levels with adiposity. Whether this effect is haplotype specific or is unrelated to genetic phase needs to be ascertained in larger family cohorts.

[0099] Our study has demonstrated that ENPP1 mRNA is present in cellular types responsive to insulin, and higher protein expression may mimic to a certain extent the effects of insulin receptor inactivation in the brain where insulin has potent anorectic actions³³, or in the skeletal muscle, both leading to an increased fat mass^{34,35}. In addition, insulin receptor knockout mice in liver³⁶ or in β -cell³⁷ are glucose intolerant. Conversely, Um et al. recently demonstrated in S6K1-deficient mice that protection against age and diet-induced obesity was associated with an enhanced insulin sensitivity³⁸. In the current study obese children carrying the ENPP1 risk haplotype have a strong family history of T2D and are more often intolerant to glucose load. This suggests that the exaggerated insulin resistance conferred by inherited increased ENPP1 expression in the liver, muscle and adipose tissue (and possibly in the brain and the pancreatic β -cell), in the context of a very strong Westernized obesogenic environment, may modify nutrition partitioning and contribute to excessive fat accumulation. In this regard Barroso et al. recently showed that the ENPP1 Q121 allele was associated with increased BMI in the UK general population³⁹. The data presented here support the view of a causative effect of primary insulin resistance on childhood obesity which can be viewed as a "pre-diabetic" state²⁵. According to this hypothesis, insulin resistance-induced fasting hyperinsulinemia was shown to be a strong predictor for the subsequent development of obesity in children of various ethnic groups^{40,41}.

[0100] In conclusion, this study strongly supports a genetic link between ENPP1 gene variants and chromosome 6q-linked childhood polygenic obesity and also with adult obesity and T2D. Our data provide an insight into the molecular basis for the physiologic association between insulin resistance and obesity, and therefore present a new perspective for prevention, and treatment of these conditions.

REFERENCES

[0101]

1. Meyre, D. et al. A genome-wide scan for childhood obesity-associated traits in French families shows significant linkage on chromosome 6q22.31-q23.2. *Diabetes* 53, 803-11(2004).
2. Atwood, L.D. et al. Genomewide Linkage Analysis of Body Mass Index across 28 Years of the Framingham Heart Study. *Am. J. Hum. Genet.* 71, 1044-50 (2002).
3. Duggirala, R. et al. A major locus for fasting insulin concentrations and insulin resistance on chromosome 6q with strong pleiotropic effects on obesity-related phenotypes in nondiabetic Mexican Americans. *Am. J. Hum. Genet.* 68, 1149-64 (2001).
4. Abney, M., Ober, C. & McPeck, M.S. Quantitative-trait homozygosity and association mapping and empirical genomewide significance in large, complex pedigrees: fasting serum- insulin level in the Hutterites. *Am. J. Hum. Genet.* 70, 920-34 (2002).
5. Gelder Ehm, M.G. et al. Genomewide search for type 2 diabetes susceptibility genes in four American populations. *Am. J. Hum. Genet.* 66, 1871-81 (2000).
6. Ghosh, S. et al. The Finland-United States investigation of non-insulin-dependent diabetes mellitus genetics (FUSION) study. I. An autosomal genome scan for genes that predispose to type 2 diabetes. *Am. J. Hum. Genet.* 67, 1174-85 (2000).
7. Demenais, F. et al. A meta-analysis of four European genome screens (GIFT Consortium) shows evidence for a novel region on chromosome 17p11.2-q22 linked to type 2 diabetes. *Hum Mol Genet* 12, 1865-73 (2003).
8. Xiang, K. et al. Genome-Wide Search for Type 2 Diabetes/Impaired Glucose Homeostasis Susceptibility Genes in the Chinese: Significant Linkage to Chromosome 6q21-q23 and Chromosome 1q21-q24. *Diabetes* 53, 228-34 (2004).
9. Maddux, B.A. et al. Membrane, glycoprotein PC-1 and insulin resistance in non-insulin- dependent diabetes mellitus. *Nature* 373,448-51 (1995).
10. Dong, H. et al. Increased Hepatic Levels of the Insulin Receptor Inhibitor, PC-1/NPP1, Induce Insulin Resistance and Glucose Intolerance. *Diabetes* 54, 367-72 (2005).
11. Costanzo, B.V. et al. The Q allele variant (GLN121) of membrane glycoprotein PC-1 interacts with the insulin receptor and inhibits insulin signaling more effectively than the common K allele variant (LYS121). *Diabetes* 50, 831-6 (2001).
12. Rolland-Cachera, M.F. et al. Body Mass Index variations: centiles from birth to 87 years. *Eur J Clin Nutr* 45, 13-21 (1991).
13. Maddux, B.A. & Goldfine, I.D. Membrane glycoprotein PC-1 inhibition of insulin receptor function occurs via direct interaction with the receptor alpha-subunit. *Diabetes* 49, 13-19 (2000).
14. Pizzuti, A. et al. A polymorphism (K121Q) of the human glycoprotein PC-1 gene coding region is strongly associated with insulin resistance. *Diabetes* 48, 1881-4 (1999).

15. Gu, H.F. et al. Association between the human glycoprotein PC-1 gene and elevated glucose and insulin levels in a paired-sibling analysis. *Diabetes* 49, 1601-3 (2000).
16. Kubaszek, A., Pihlajamaki, J., Karhapaa, P., Vauhkonen, I. & Laakso, M. The K121Q polymorphism of the PC-1 gene is associated with insulin resistance but not with dyslipidemia. *Diabetes Care* 26, 464-7 (2003).
- 5 17. Abate, N. et al. Genetic polymorphism PC-1 K121Q and ethnic susceptibility to insulin resistance. *J Clin Endocrinol Metab* 88, 5927-34 (2003).
18. Hamaguchi, K. et al. The PC-1 Q121 allele is exceptionally prevalent in the Dominican Republic and is associated with type 2 diabetes. *J Clin Endocrinol Metab* 89, 1359-64 (2004).
19. Frittitta, L. et al. A cluster of three single nucleotide polymorphisms in the 3'-untranslated region of human glycoprotein PC-1 gene stabilizes PC-1 mRNA and is associated with increased PC-1 protein content and insulin resistance-related abnormalities. *Diabetes* 50, 1952-5 (2001).
- 10 20. Hager, J. et al. A genome-wide scan for human obesity genes reveals a major susceptibility locus on chromosome 10. *Nat Genet* 20, 304-8 (1998).
21. Deng, H.W., Chen, W.M. & Recker, R.R. QTL fine mapping by measuring and testing for Hardy-Weinberg and linkage disequilibrium at a series of linked marker loci in extreme samples of populations. *Am J Hum Genet* 66, 1027-45 (2000).
- 15 22. Eaves, I.A. et al. Transmission ratio distortion at the INS-IGF2 VNTR. *Nat Genet* 22, 324-5 (1999).
23. Li, C., Scott, L.J. & Boehnke, M. Assessing whether an allele can account in part for a linkage signal: the Genotype-IBD Sharing Test (GIST). *Am J Hum Genet* 74, 418-31 (2004).
- 20 24. Rutsch, F. et al. PC-1 nucleoside triphosphate pyrophosphohydrolase deficiency in idiopathic infantile arterial calcification. *Am J Pathol* 158, 543-54 (2001).
25. Weill, J., Vanderbecken, S. & Froguel, P. Understanding the rising incidence of type 2 diabetes in adolescence. *Arch Dis Child* 89, 502-4 (2004).
26. Jenkinson, C.P. et al. Comprehensive analysis of snps in ENPP1: association with diabetes. ASHG abstract (2004).
- 25 27. Ohlsson, H., Karlsson, K. & Edlund, T. IPF 1-, a homeodomain-containing transactivator of the insulin gene. *Embo J* 12, 4251-9 (1993).
28. Hugot, J.P. et al. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 411, 599-603 (2001).
- 30 29. Hampe, J. et al. Evidence for a NOD2-independent susceptibility locus for inflammatory bowel disease on chromosome 16p. *Proc Natl Acad Sci U S A* 99, 321-6 (2002).
30. Holder, J.L., Jr., Butte, N.F. & Zinn, A.R. Profound obesity associated with a balanced translocation that disrupts the SIM1 gene. *Hum Mol Genet* 9, 101-8. (2000).
31. Michaud, J.L. et al. Sim1 haploinsufficiency causes hyperphagia, obesity and reduction of the paraventricular nucleus of the hypothalamus. *Hum Mol Genet* 10, 1465-73 (2001).
- 35 32. Tregouet, D.A., Escolano, S., Tiret, L., Mallet, A. & Golmard, J.L. A new algorithm for haplotype-based association analysis: the Stochastic-EM algorithm. *Ann Hum Genet* 68, 165-77 (2004).
33. Schwartz, M.W. Progress in the search for neuronal mechanisms coupling type 2 diabetes to obesity. *J Clin Invest* 108, 963-4 (2001).
- 40 34. Bruning, J.C. et al. Role of brain insulin receptor in control of body weight and reproduction. *Science* 289, 2122-5 (2000).
35. Kim, J.K. et al. Redistribution of substrates to adipose tissue promotes obesity in mice with selective insulin resistance in muscle. *J Clin Invest* 105, 1791-7 (2000).
36. Michael, M.D. et al. Loss of insulin signaling in hepatocytes leads to severe insulin resistance and progressive hepatic dysfunction. *Mol Cell* 6, 87-97 (2000).
- 45 37. Kulkarni, R.N. et al. Tissue-specific knockout of the insulin receptor in pancreatic beta cells creates an insulin secretory defect similar to that in type 2 diabetes. *Cell* 96, 329-39 (1999).
38. Um, S.H. et al. Absence of S6K1 protects against age- and diet-induced obesity while enhancing insulin sensitivity. *Nature* 431, 200-205 (2004).
- 50 39. Barroso, I. et al. Candidate Gene Association Study in Type 2 Diabetes Indicates a Role for Genes Involved in beta-Cell Function as Well as Insulin Action. *PLoS Biol* 1, E20 (2003).
40. Odeleye, O.E., de Courten, M., Pettitt, D.J. & Ravussin, E. Fasting hyperinsulinemia is a predictor of increased body weight gain and obesity in Pima Indian children. *Diabetes* 46, 1341-5 (1997).
41. Johnson, M.S., Figueroa-Colon, R., Huang, T.T., Dwyer, J.H. & Goran, M.I. Longitudinal changes in body fat in African American and Caucasian children: influence of fasting insulin and insulin sensitivity. *J Clin Endocrinol Metab* 86, 3182-7 (2001).
- 55 42. Poskitt, E.M. Defining childhood obesity: the relative body mass index (BMI). European Childhood Obesity group. *Acta Paediatr* 84, 961-3 (1995).

43. Hercberg, S. et al. A primary prevention trial using nutritional doses of antioxidant vitamins and minerals in cardiovascular diseases and cancers in a general population: the SU.VI.MAX study-design, methods, and participant characteristics. *SUpplémentation en Vitamines et Minéraux AntioXydants. Control Clin Trials* 19, 336-51 (1998).
- 5 44. Lafay, L. et al. Determinants and nature of dietary underreporting in a free-living population: the Fleurbaix Laventie Ville Sante (FLVS) Study. *Int J Obes Relat Metab Disord* 21, 567-73 (1997).
45. Oberkofler, H. et al. Complex haplotypes of the PGC-lalpha gene are associated with carbohydrate metabolism and type 2 diabetes. *Diabetes* 53, 1385-93 (2004).
46. Cole, T.J. The LMS method for constructing normalized growth standards. *Eur J Clin Nutr* 44, 45-60 (1990).
- 10 47. Seltzer, H.S., Allen, E.W., Herron, A.L., Jr. & Brennan, M.T. Insulin secretion in response to glycemic stimulus: relation of delayed initial release to carbohydrate intolerance in mild diabetes mellitus. *J Clin Invest* 46, 323-35 (1967).
48. Risch, N. Linkage strategies for genetically complex traits. I. Multilocus models. *Am J Hum Genet* 46, 222-228 (1990).
49. Kruglyak, L. & Lander, E.S. Complete multipoint sib-pair analysis of qualitative and quantitative traits. *Am J Hum Genet* 57, 439-54 (1995).
- 15 50. Sham, P.C. & Curtis, D. Monte Carlo tests for associations between disease and alleles at highly polymorphic loci. *Ann Hum Genet* 59 (Pt 1), 97-105 (1995).
51. Dudbridge, F., Koeleman, B.P., Todd, J.A. & Clayton, D.G. Unbiased application of the transmission/disequilibrium test to multilocus haplotypes. *Am J Hum Genet* 66, 2009-12 (2000).
- 20 52. Lukowiak, B. et al. Identification and purification of functional human beta-cells by a new specific zinc-fluorescent probe. *J Histochem Cytochem* 49, 519-28 (2001).

SEQUENCE LISTING

[0102]

- 25 <110> Centre National de la Recherche Scientifique (CNRS)
 Université de Droit et de la Santé de Lille 2
 FROGUEL Philippe
 MEYRE David
 30 DINA Christian
 BOUTIN Philippe
- <120> ENPP1 (PC-1) gene halotype associated with the risk of obesity and type 2 diabetes and their applications
- 35 <130> D23085
- <150> EP 05291298.7
 <151> 2005-06-17
- 40 <160> 15
- <170> PatentIn version 3.1
- 45 <210> 1
 <211> 832
 <212> DNA
 <213> homo sapiens
- 50 <220>
 <223> Fragment of the genomic DNA of the ENPP1 gene including the SNP "IVS20 delT-11" in the position -11 before exon 21
- <400> 1
- 55

EP 1 891 237 B9

```

aaaaattagc caggcgtggt gggtcacacc tgtaatccca gcactttgag ggggctgcag      60
cggatggatc acctgaggtc aggagctcaa gaccagcctt gccaacatgg tgaaacccca      120
tctctaccaaa aaacacaaaa attagccagg cgtgggtggca gatgcctgta gtcccagcta      180
5  ctcgggaagc tgaggcagga gaatcgcttg aacctggggg gcagagtgtc ggtgagccga      240
gattgcgccg ctgccctcta gtctgggtga cagagtgaga ctccatctta aaaaaataat      300
aataataaat aaataaataa ataaataaat atttaaaatt gtgtagataa aatcattcta      360
aacattatth catattagca tagcagaatc tgaaaatatt tgcataaata tgacaattaa      420
tatctttaat attgtaaagc atttttacac tttagttaga aaaaaagatg aatatactag      480
10 taggaaaata gggaaggaca tgagctgaca gctagagctt cataatttta tgatgtagtt      540
cacctttaaata tattaataaa gcaatthttct tctctgtgcc tgatatctga gagttcttct      600
cattttcgtt cttcaggaca gtttctctac ggaagacttc tccaactgtc tgtaccagga      660
ctttagaatt cctcttagtc ctgtccataa atgttcattt tataaaaaata acaccaaagt      720
gagttacggg ttctctctcc caccacgtaa gthttttctct ctctgacct tccctthttct      780
15 cctthtttgth ttctthttctg thttataaatc ctaccataca ttatagggha at      832

```

```

<210> 2
<211> 21
20 <212> DNA
    <213> artificial sequence

<220>
<223> Primer derived from the genomic DNA encoding the human protein ENPP1
25

<400> 2
tcatactcag gaagacagca a 21

<210> 3
<211> 19
30 <212> DNA
    <213> artificial sequence

<220>
<223> Primer derived from the genomic DNA encoding the human protein ENPP1
35

<400> 3
caatagccat gactcctaa      19

40 <210> 4
<211> 20
<212> DNA
<213> artificial sequence

45 <220>
<223> Primer derived from the genomic DNA encoding the human protein ENPP1

<400> 4
50 agcattthta cactthtagth      20

<210> 5
<211> 19
<212> DNA
<213> artificial sequence
55

<220>
<223> Primer derived from the genomic DNA encoding the human protein ENPP1

```

EP 1 891 237 B9

<400> 5
ataatgtatg gtaggattt 19

5
<210> 6
<211> 20
<212> DNA
<213> artificial sequence

10
<220>
<223> Primer derived from the genomic DNA encoding the human protein ENPP1

<400> 6
atattcctat cctgctcact 20

15
<210> 7
<211> 20
<212> DNA
<213> artificial sequence

20
<220>
<223> Primer derived from the genomic DNA encoding the human protein ENPP1

<400> 7
tgcagctggc ccttaggccg 20

25
<210> 8
<211> 27
<212> DNA
<213> artificial sequence

30
<220>
<223> Primer derived from the genomic DNA encoding the human protein ENPP1

<400> 8
ctttcccaaa tcactacagc attgtca 27

35
<210> 9
<211> 27
<212> DNA
<213> artificial sequence

40
<220>
<223> Primer derived from the genomic DNA encoding the human protein ENPP1

45
<400> 9
tttcagacca tccatcagca taccaac 27

50
<210> 10
<211> 25
<212> DNA
<213> artificial sequence

55
<220>
<223> Primer derived from the genomic DNA encoding the human protein ENPP1

<400> 10
gtcctgtgtt tgactttgat tatga 25

EP 1 891 237 B9

<210> 11
<211> 25
<212> DNA
<213> artificial sequence
5
<220>
<223> Primer derived from the genomic DNA encoding the human protein ENPP1

<400> 11
10 cccttaggcc gttgaagaat ggtca 25

<21.0> 12
<211> 32
<212> DNA
15 <213> artificial sequence

<220>
<223> Primer derived from the genomic DNA encoding the human protein beta-actin
20 <400> 12
cgtcatactc ctgcttgctg atccacatct gc 32

<210> 13
<211> 32
25 <212> DNA
<213> artificial sequence

<220>
<223> Primer derived from the genomic DNA encoding the human protein beta-actin
30 <400> 13
atctggcacc acaccttcta caatgagctg cg 32

<210> 14
35 <211> 873
<212> PRT
<213> Homo sapiens

<220>
40 <223> Human ENPP1 protein (Genbank Accession Number NP_006199)

<400> 14

45

50

55

EP 1 891 237 B9

	Met	Asp	Val	Gly	Glu	Glu	Pro	Leu	Glu	Lys	Ala	Ala	Arg	Ala	Arg	Thr
	1				5					10					15	
	Ala	Lys	Asp	Pro	Asn	Thr	Tyr	Lys	Val	Leu	Ser	Leu	Val	Leu	Ser	Val
5				20					25					30		
	Cys	Val	Leu	Thr	Thr	Ile	Leu	Gly	Cys	Ile	Phe	Gly	Leu	Lys	Pro	Ser
			35					40					45			
	Cys	Ala	Lys	Glu	Val	Lys	Ser	Cys	Lys	Gly	Arg	Cys	Phe	Glu	Arg	Thr
		50					55					60				
10	Phe	Gly	Asn	Cys	Arg	Cys	Asp	Ala	Ala	Cys	Val	Glu	Leu	Gly	Asn	Cys
	65					70					75				80	
	Cys	Leu	Asp	Tyr	Gln	Glu	Thr	Cys	Ile	Glu	Pro	Glu	His	Ile	Trp	Thr
					85					90					95	
	Cys	Asn	Lys	Phe	Arg	Cys	Gly	Glu	Lys	Arg	Leu	Thr	Arg	Ser	Leu	Cys
				100						105				110		
15	Ala	Cys	Ser	Asp	Asp	Cys	Lys	Asp	Lys	Gly	Asp	Cys	Cys	Ile	Asn	Tyr
			115					120					125			
	Ser	Ser	Val	Cys	Gln	Gly	Glu	Lys	Ser	Trp	Val	Glu	Glu	Pro	Cys	Glu
			130				135					140				
20	Ser	Ile	Asn	Glu	Pro	Gln	Cys	Pro	Ala	Gly	Phe	Glu	Thr	Pro	Pro	Thr
	145					150					155					160
	Leu	Leu	Phe	Ser	Leu	Asp	Gly	Phe	Arg	Ala	Glu	Tyr	Leu	His	Thr	Trp
					165					170					175	
	Gly	Gly	Leu	Leu	Pro	Val	Ile	Ser	Lys	Leu	Lys	Lys	Cys	Gly	Thr	Tyr
			180						185					190		
25	Thr	Lys	Asn	Met	Arg	Pro	Val	Tyr	Pro	Thr	Lys	Thr	Phe	Pro	Asn	His
			195					200					205			
	Tyr	Ser	Ile	Val	Thr	Gly	Leu	Tyr	Pro	Glu	Ser	His	Gly	Ile	Ile	Asp
		210				215						220				
30	Asn	Lys	Met	Tyr	Asp	Pro	Lys	Met	Asn	Ala	Ser	Phe	Ser	Leu	Lys	Ser
	225					230					235					240
	Lys	Glu	Lys	Phe	Asn	Pro	Glu	Trp	Tyr	Lys	Gly	Glu	Pro	Ile	Trp	Val
				245						250					255	
	Thr	Ala	Lys	Tyr	Gln	Gly	Leu	Lys	Ser	Gly	Thr	Phe	Phe	Trp	Pro	Gly
				260					265					270		
35	Ser	Asp	Val	Glu	Ile	Asn	Gly	Ile	Phe	Pro	Asp	Ile	Tyr	Lys	Met	Tyr

40

45

50

55

EP 1 891 237 B9

5 Arg Asn Gln Glu Ile Leu Ile Pro Thr His Phe Phe Ile Val Leu Thr
 770 775 780
 Ser Cys Lys Asp Thr Ser Gln Thr Pro Leu His Cys Glu Asn Leu Asp
 785 790 795 800
 Thr Leu Ala Phe Ile Leu Pro His Arg Thr Asp Asn Ser Glu Ser Cys
 805 810 815
 Val His Gly Lys His Asp Ser Ser Trp Val Glu Glu Leu Leu Met Leu
 820 825 830
 10 His Arg Ala Arg Ile Thr Asp Val Glu His Ile Thr Gly Leu Ser Phe
 835 840 845
 Tyr Gln Gln Arg Lys Glu Pro Val Ser Asp Ile Leu Lys Leu Lys Thr
 850 855 860
 15 His Leu Pro Thr Phe Ser Gln Glu Asp
 865 870

<210> 15

<211> 1173

20 <212> DNA

<213> Homo sapiens

<220>

<223> 3'UTR domain of the human ENPP1 gene

25

<400> 15

30 tgatatgttt tttatcccca aacacccatga atctttttga gagaacctta tattttatat 60
 agtcctctag ctacactatt gcattgttca gaaactgtcg accagagtta gaacggagcc 120
 ctcggtgatg cggacatctc agggaaactt gcgtactcag cacagcagtg gagagtgttc 180
 ctggttgaatc ttgcacatat ttgaatgtgt aagcattgta tacattgatc aagttcgggg 240
 gaataaagac agaccacacc taaaactgcc tttctgcttc tcttaaagga gaagtagctg 300
 tgaacattgt ctggatacca gatatttgaa tctttcttac tattggtaat aaaccttgat 360
 35 ggcattgggc aaacagtaga cttatagtag ggttggggta gcccatgtta tgtgactatc 420
 tttatgagaa ttttaaagtg gttctggata tcttttaact tggagtttca tttcttttca 480
 ttgtaatcaa aaaaaaatt aacagaagcc aaaatacttc tgagaccttg tttcaatctt 540
 tgctgtatat cccctcaaaa tccaagttat taatcttatg tgttttcttt ttaatttttt 600
 gattggattt ctttagattt aatggttcaa atgagttcaa ctttgagggga cgatctttga 660
 atatacttac ctattataaa atcttacttt gtatttgtat ttaaaaaaga aaaatattcc 720
 40 tatcctgctc actggtaatt aacataggtt taaaatggct tcaaatgtgg ccctatagac 780
 ggttaaaatt gtaccttacc ttggcaaac ttcagagcac cagtcagtg atgcaagggtg 840
 ccatttttta ttgagatgct tagaatgttt ctttctgtgc acaagactta ccctaccagc 900
 agcagagcca ttctctgttg agtggttcat tttgaagttc cacagattga agagaacatg 960
 ccaccaatca cctcacatct tcttgggtga catgataaat gacacaatga acttgatttc 1020
 45 tttactacct tgactgtagc tttttatccc tacctgtgaa ccttcaaaga ctgcattaac 1080
 ttttaggcta cataggtcaa ttgaggtata atatcagtac accaaagatt tttatatgtc 1140
 cttcgtgtga ccattcttca acggcctaag ggc 1173

50

Claims

1. Method for determining if a subject is at increased risk to develop obesity or pathology related to obesity from a biological sample from said subject containing genomic DNA or RNA, said method comprising the steps of:

55

a) determining on at least one gene allele or RNA sequence encoding the protein ENPP1, if said sequence, or fragments thereof, contains the haplotype comprising the three following SNPs:

- K121Q,
- IVS20 delT-11, and
- a SNP localized in the 3'UTR domain sequence and wherein the first 1170 bp.3'UTR domain wild type sequence (including the TGA _codon) has the sequence SEQ ID No. 15; and

- 5
- b) observing whether or not the subject is at increased risk to develop obesity or pathology related to obesity by observing if said sequence, or fragments thereof, contains said haplotype, the presence of said haplotype indicates said subject is at increased risk to develop obesity or pathology related to obesity.
- 10
2. Method according to claim 1, wherein the pathology related to obesity is selected from the group consisting of type 2 diabetes, hypertension, hypercholesterolemia and hyperlipidemia.
3. Method according to claim 1 or 2, wherein the pathology related to obesity is type 2 diabetes.
- 15
4. Method according to one of claims 1 to 3, wherein said SNP localized in the 3'UTR domain is selected from the group of SNPs consisting of A>G +1044 TGA, A>C +1092 TGA and C>T +1157 TGA.
5. Method according to claim 4, wherein said SNP localized in the 3'UTR domain is selected from the group of SNPs consisting of A>G +1044 TGA and A>C +1092 TGA.
- 20
6. Method according to claim 5, wherein said SNP localized in the 3'UTR domain is of A>G +1044 TGA.
7. Method for determining if a subject is at high increased risk to develop obesity or pathology related to obesity, **characterized in that** said method comprises the step a) and b) of the method according to one of claims 1 to 6, wherein:
- 25
- in step a), it is determined whether or not the subject is homozygous for said haplotype comprising said three SNPs; and
 - in step b), it is observed whether or not the subject is at high increased risk to develop obesity or pathology related to obesity by observing if the subject is homozygous for said haplotype, the homozygosity indicating said subject is at high increased risk to develop obesity or pathology related to obesity.
- 30
8. Method according to one of claims 1 to 7, wherein the step a) of determining the presence or absence of said three allele risk haplotype is carried out by a method comprising a step of amplification of the genomic DNA or RNA sequence, or fragments thereof, susceptible to contain said haplotype SNPs by polymerase chain reaction (PCR).
- 35
9. Method according to claim 7, wherein the step of PCR is carried out from genomic DNA, mRNA or cDNA, eventually after a step of reverse transcription, or fragments thereof, by using the following sets of primers:
- 40
- a) 5'-TCATACTCAGGAAGACAGCAA-3' (forward primer SEQ ID No. 2) and 5'-CAATAGCCATGACTCCTAA-3' (reverse primer SEQ ID No. 3) for "K121Q" SNP;
 - b) 5'-AGCATTTTTACTTTAGTT-3' (forward primer SEQ ID No. 4) and 5'-ATAATGTATGGTAGGATTT-3' (reverse primer SEQ ID No. 5) for IVS20 del T - 11 SNP; and
 - c) 5'-ATATTCCTATCCTGCTCACT-3' (forward primer SEQ ID No. 6) and 5'-TGCAGCTGGCCCTTAGGCCG-3' (reverse primer SEQ ID No. 7) for A>G +1044 TGA SNP.
- 45
10. Method according to one of claims 1 to 9, wherein said biological sample is a sample of liver, adipocytes or pancreatic beta-cells when the presence of said haplotype is determined on RNA sequence or fragments thereof.
- 50
11. Method for determining if a subject is at increased risk to develop obesity or pathology related to obesity from a plasma or serum sample from said subject, said method comprising the steps of:
- a) determining the concentration of the serum or plasma ENPP1 protein and further the presence of the haplotype comprising the three SNPs as defined in claims 1 and 4 to 6; and
 - b) observing whether or not the subject is at increased risk to develop obesity or pathology related to obesity by observing if said concentration is significantly higher than the level of normal, the presence of said significantly higher concentration indicates said subject is at increased risk to develop obesity or pathology related to obesity.
- 55

12. Method according to claim 11, wherein said higher significant concentration of serum or plasma ENPP1 protein corresponds to at least more than 10 %, preferably more than 15 % of the normal serum or plasma concentration.

13. Method according to claim 11 or 12, wherein said concentration of serum or plasma ENPP1 protein is determined by using antibodies capable of specifically recognizing the ENPP1 protein.

14. Method according to claim 13, **characterized in that** said antibodies are brought into contact with the serum or plasma to be tested, under conditions allowing the formation of a specific immunological complex between the ENPP1 protein and said antibody, and **in that** the immunological complexes formed is quantified.

15. Method according to one of claims 11 to 14, wherein the pathology related to obesity is selected from the group consisting of type 2 diabetes, hypertension, hypercholesterolemia and hyperlipidemia.

16. Method according to one of claims 11 to 15, wherein the pathology related to obesity is type 2 diabetes.

17. Kit for determining if a subject is at increased risk to develop obesity or pathology related to obesity, comprising at least one pair of primers or a set of a pair of primers capable of amplifying a fragment of genomic DNA or RNA encoding the protein ENPPI and susceptible of containing the haplotype comprising the three following SNPs:

- K121Q,
- IVS20 delT-11, and
- a SNP localized in the 3'UTR domain sequence and wherein the first 1170 bp 3'UTR domain wild type sequence (including the TGA codon) has the sequence SEQ ID No. 15;

and wherein said at least set of a pair of primers is:

- a) 5'-TCATACTCAGGAAGACAGCAA-3' (forward primer SEQ ID No. 2) and 5'-CAATAGCCATGACTCCTAA-3' (reverse primer SEQ ID No. 3) for "K121Q" SNP;
- b) 5'-AGCATTTTTACTTTAGTT-3' (forward primer SEQ ID No. 4) and 5'-ATAATGTATGGTAGGATTT-3' (reverse primer SEQ ID No. 5) for IVS20 del T - 11 SNP; and
- c) 5'-ATATTCCTATCCTGCTCACT-3' (forward primer SEQ ID No. 6) and 5'-TGCAGCTGGCCCTTAGGCCG-3' (reverse primer SEQ ID No. 7) for A>G +1044 TGA SNP.

18. Kit for determining if a subject is at increased risk to develop obesity or pathology related to obesity, comprising at least one pair of primers or a set of a pair of primers capable of amplifying a fragment of genomic DNA or RNA encoding the protein ENPPI and susceptible of containing the haplotype comprising the three following SNPs:

- K121Q,
- IVS20 delT-11, and
- a SNP localized in the 3'UTR domain sequence and wherein the first 1170 bp 3'UTR domain wild type sequence (including the TGA codon) has the sequence SEQ ID No. 15;

and wherein said at least set of a pair of primers is:

- a) - 5'-CTTTCCCAATCACTACAGCATTGTCA-3' (forward primer, exon 7, SEQ ID No. 8), and
- 5'-TTTCAGACCATCCATCAGCATACCAAC-3' (reverse primer exon 12, SEQ ID No. 9); and for the 3'UTR SNP;
- b) - 5'-GTCCTGTGTTTGACTTTGATTATGA-3' (forward primer exon 23, SEQ ID No. 10), and
- 5'-CCCTTAGCCGTTGAAGAATGGTCA-3' (reverse primer 3'UTR, SEQ ID No. 11).

19. Isolated variant nucleic sequence, or complementary sequence thereof, of a mammal, preferably human, genomic or RNA sequence of the gene coding for ENPP1 protein, **characterized in that** said isolated variant nucleic sequence comprises at least the presence of the haplotype comprising the three following SNPs as defined in claims 1 and 4 to 6.

20. Cloning and/or expression vector containing a nucleic acid sequence according to claim 19.

21. Host cell transformed by a vector according to claim 20.

22. Transgenic mammal, except human, **characterized in that** it comprises a transformed cell according to claim 21

or a cell containing a recombinant functional ENPPI gene comprising at least the haplotype comprising the three following SNPs as defined in claims 1 and 4 to 6.

23. Use of cell according to claim 21 or a transgenic mammal according to claim 22, for screening chemical or biochemical compounds for the treatment or the prevention of obesity or pathology related to obesity.

24. Method for selecting a chemical or biochemical compound for the treatment or the prevention of obesity or pathology related to obesity, **characterized in that** it comprises the step of:

- a) bringing into contact a cell according to claim 21 or a transgenic mammal according to claim 26 with said chemical or biochemical compound to be tested; and
- b) observing whether the expression product of the ENPP1 gene is decreased or not in said cell or said mammal.

25. Method for selecting a chemical or biochemical compound for the treatment or the prevention of obesity or pathology related to obesity according to claim 24, **characterized in that**:

- in step a), the chemical or biochemical compound to be tested is administered to a transgenic mammal according to claim 22 or to a transgenic mammal overexpressing the ENPP1 protein; and
- in step b), it is observed whether the serum or plasma ENPP1 protein concentration is decreased compared to a reference,

said chemical or biochemical compound being selected if a decrease of said concentration is observed.

Patentansprüche

1. Verfahren zum Feststellen, ob eine Person ein erhöhtes Risiko aufweist, Fettsucht oder die mit der Fettsucht assoziierten Pathologie zu entwickeln, anhand einer biologischen Probe von der genannten Person, enthaltend genomisches DNA oder RNA, wobei das genannte Verfahren die nachfolgenden Schritte umfasst:

a) Feststellen wenigstens eines Gen-Allels oder einer RNA-Sequenz, die das Protein ENPP1 kodiert, falls die genannte Sequenz oder Fragment davon den Haplotyp, umfassend die drei nachfolgenden SNPs enthält:

- K121Q,
- IVS20 delT-11 und
- ein SNP, das in der 3'UTR-Domänensequenz lokalisiert ist und wobei die erste 1170 bp 3'UTR-Domänensequenz des Wildtyps (einschließlich des TGA-Kodons) die Sequenz SEQ ID Nr. 15 aufweist; und

b) Beobachten, ob die Person ein erhöhtes Risiko aufweist, Fettsucht oder die mit der Fettsucht assoziierten Pathologie zu entwickeln, anhand der Beobachtung ob die genannte Sequenz oder Fragmente davon das genannte Haplotyp enthält, wobei die Anwesenheit des genannten Haplotyps darauf hindeutet, dass die genannte Person ein erhöhtes Risiko aufweist, Fettsucht oder die mit der Fettsucht assoziierten Pathologie zu entwickeln.

2. Verfahren nach Anspruch 1, wobei die mit Fettsucht assoziierte Pathologie ausgewählt ist aus der Gruppe bestehend aus Typ 2 Diabetes, Bluthochdruck, Cholesterinspiegelerhöhung und Fettspiegelerhöhung.

3. Verfahren nach Anspruch 1 oder 2, wobei die mit der Fettsucht assoziierte Pathologie Typ 2 Diabetes ist.

4. Verfahren nach einem der Ansprüche 1 bis 3, wobei das genannte SNP, das in der 3'UTR-Domäne lokalisiert ist, ausgewählt ist aus der Gruppe von SNPs, bestehend aus A>G+1044 TGA, A>C+1092 TGA und C>T+1157 TGA.

5. Verfahren nach Anspruch 4, wobei das genannte SNP, das in der 3'UTR-Domäne lokalisiert ist, ausgewählt ist aus der Gruppe von SNPs, bestehend aus A>G+1044 TGA und A>C+1092 TGA.

6. Verfahren nach Anspruch 5, wobei das genannte SNP, das in der 3'UTR-Domäne lokalisiert ist, A>G+1044 TGA ist.

7. Verfahren zum Feststellen, ob eine Person ein erhöhtes Risiko aufweist, Fettsucht oder die mit der Fettsucht as-

EP 1 891 237 B9

soziierten Pathologie zu entwickeln, **dadurch gekennzeichnet, dass** das genannte Verfahren die nachfolgenden Schritte umfasst: a) und b) des Verfahrens nach einem der Ansprüche 1 bis 6, wobei:

- in Schritt a) festgestellt wird, ob die Person für den genannten Haplotyp umfassend die genannten drei SNPs erbrein ist; und
- in Schritt b) beobachtet wird, ob die Person ein erhöhtes Risiko aufweist, Fettsucht oder die mit der Fettsucht assoziierten Pathologie zu entwickeln, in dem beobachtet wird, ob die Person für den genannten Haplotyp erbrein ist, wobei eine Erbreinheit darauf hinweist, dass die genannte Person ein erhöhtes Risiko aufweist, Fettsucht oder die mit der Fettsucht assoziierten Pathologie zu entwickeln.

8. Verfahren nach einem der Ansprüche 1 bis 7, wobei der Schritt a) zum Feststellen der Anwesenheit oder Abwesenheit der genannten drei Allel-Risikohaplotyp durch ein Verfahren umfassend einen Schritt zur Verstärkung der genomischen DNA- oder RNA-Sequenz oder Fragmenten davon durchgeführt wird, die dafür anfällig sind, die genannten Haplotyp-SNPs durch eine Polymerase-Kettenreaktion (PCR) zu enthalten.

9. Verfahren nach Anspruch 7, wobei der Schritt von PCR aus genomischem DNA, mRNA oder cDNA schließlich nach einem Schritt einer reversen Transkription oder Fragmenten davon unter Anwendung der nachfolgenden Parametersätze durchgeführt wird:

- a) 5'-TCATACTCAGGAAGACAGCAA-3' (Vorwärtsprimer SEQ ID Nr. 2) und 5'-CAATAGCCATGACTCCTAA-3' (Rückwärtsprimer SEQ ID No. 3) für "K121Q" SNP;
- b) 5'-AGCATTTTTACTTTAGTT-3' (Vorwärtsprimer SEQ ID Nr. 4) und 5'-ATAATGTATGGTAGGATTT-3' (Rückwärtsprimer SEQ ID Nr. 5) für IVS20 del T-11 SNP; und
- c) 5'-ATATTCCTATCCTGCTCACT-3' (Vorwärtsprimer SEQ ID Nr. 6) und 5'-TGCAGCTGGCCCTTAGGCCG-3' (Rückwärtsprimer SEQ ID Nr. 7) für A>G+1044 TGA SNP.

10. Verfahren nach einem der Ansprüche 1 bis 9, wobei die genannte biologische Probe eine Probe der Leber, Adipozyten oder Pankreas-Betazellen ist, falls die Anwesenheit des genannten Haplotyps an der RNA-Sequenz oder Fragmenten davon bestimmt ist.

11. Verfahren zum Feststellen, Verfahren zum Feststellen, ob eine Person ein erhöhtes Risiko aufweist, Fettsucht oder die mit der Fettsucht assoziierten Pathologie zu entwickeln, wobei das genannte Verfahren die nachfolgenden Schritte umfasst:

- a) Feststellen der Konzentration des Serums oder Plasma-ENPP1-Proteins und ferner der Anwesenheit des Haplotyps, umfassend die drei in Anspruch 1 und 4 bis 6 definierten SNPs; und
- b) Beobachten, ob die Person ein erhöhtes Risiko aufweist, Fettsucht oder die mit der Fettsucht assoziierten Pathologie zu entwickeln, indem beobachtet wird, ob die genannte Konzentration erheblich höher liegt, als der Normalwert, wobei die Anwesenheit der genannten erheblich erhöhten Konzentration darauf hinweist, dass die Person ein erhöhtes Risiko aufweist, Fettsucht oder die mit der Fettsucht assoziierten Pathologie zu entwickeln.

12. Verfahren nach Anspruch 11, wobei die genannte erheblich erhöhte Konzentration des Serums oder Plasma-ENPP1-Proteins wenigstens mehr als 10 %, bevorzugt mehr als 15 % der normalen Serum- oder Plasmakonzentration entspricht.

13. Verfahren nach Anspruch 11 oder 12, wobei die genannte Konzentration des Serums oder des Plasma-ENPP1-Proteins durch Anwendung von Antikörpern bestimmt wird, die in der Lage sind, das ENPP1-Protein spezifisch zu erkennen.

14. Verfahren nach Anspruch 13, **dadurch gekennzeichnet, dass** die genannten Antikörper in Kontakt mit dem zu testenden Serum oder Plasma gebracht werden, und zwar unter Bedingungen, die die Formation eines spezifischen immunologischen Komplexes zwischen dem ENPP1-Protein und dem genannten Antikörper ermöglichen und indem die gebildeten immunologischen Komplexe quantifiziert werden.

15. Verfahren nach einem der Ansprüche 11 bis 14, wobei die mit der Fettsucht assoziierte Pathologie ausgewählt ist aus der Gruppe bestehend aus Typ 2 Diabetes, Bluthochdruck, Cholesterinspiegelerhöhung und Fettspiegelerhöhung.

EP 1 891 237 B9

16. Verfahren nach einem der Ansprüche 11 bis 15, wobei die mit der Fettsucht assoziierte Pathologie Typ 2 Diabetes ist.

17. Satz zum Feststellen, ob eine Person ein erhöhtes Risiko aufweist, Fettsucht oder die mit der Fettsucht assoziierten Pathologie zu entwickeln, umfassend wenigstens ein Paar Primer oder einen Satz eines Primer-Paars, die in der Lage sind, ein Fragment eines genomischen DNAs oder RNAs zu verstärken, welches das Protein ENPP1 kodiert und dafür anfällig ist, den Haplotyp umfassend die drei nachfolgenden SNPs zu enthalten:

- K121Q,
- IVS20 delT-11 und
- ein SNP, das in der 3'UTR-Domänensequenz lokalisiert ist und wobei die erste 1170 bp 3'UTR-Domänensequenz des Wildtyps (einschließlich des TGA-Kodons) die Sequenz SEQ ID Nr. 15 aufweist;

und wobei der genannte wenigstens eine Satz an Primern

- a) 5'-TCATACTCAGGAAGACAGCAA-3' (Vorwärtsprimer SEQ ID Nr. 2) und 5'-CAATAGCCATGACTCCTAA-3' (Rückwärtsprimer SEQ ID No. 3) für "K121Q" SNP;
- b) 5'-AGCATTTTTACTTTAGTT-3' (Vorwärtsprimer SEQ ID Nr. 4) und 5'-ATAATGTATGGTAGGATTT-3' (Rückwärtsprimer SEQ ID Nr. 5) für IVS20 del T-11 SNP; und
- c) 5'-ATATTCCTATCCTGCTCACT-3' (Vorwärtsprimer SEQ ID Nr. 6) und 5'-TGCAGCTGGCCCTTAGGCCG-3' (Rückwärtsprimer SEQ ID Nr. 7) für A>G+1044 TGA SNP ist.

18. Satz zum Feststellen, ob eine Person ein erhöhtes Risiko aufweist, Fettsucht oder die mit der Fettsucht assoziierten Pathologie zu entwickeln, umfassend wenigstens ein Paar Primer oder einen Satz eines Primer-Paars, die in der Lage sind, ein Fragment eines genomischen DNAs oder RNAs zu verstärken, welches das Protein ENPP1 kodiert und dafür anfällig ist, den Haplotyp umfassend die drei nachfolgenden SNPs zu enthalten:

- K121Q,
- IVS20 delT-11 und
- ein SNP, das in der 3'UTR-Domänensequenz lokalisiert ist und wobei die erste 1170 bp 3'UTR-Domänensequenz des Wildtyps (einschließlich des TGA-Kodons) die Sequenz SEQ ID Nr. 15 aufweist;

und wobei der genannte wenigstens eine Satz an Primern

- a) 5'-CTTTCCCAATCACTACAGCATTGTCA-3' (Vorwärtsprimer, Exon 7, SEQ ID Nr. 8) und 5'-TTTCAGACCATCCATCAGCATACCAAC-3' (Rückwärtsprimer, Exon 12, SEQ ID No. 9); und für das 3'UTR SNP;
- b) 5'-GTCCTGTGTTTGACTTTGATTATGA-3' (Vorwärtsprimer, Exon 23, SEQ ID Nr. 10) und 5'-CCCTTAGGCCGTTGAAGAATGGTCA-3' (Rückwärtsprimer 3'UTR, SEQ ID Nr. 11) ist.

19. Isolierte Nukleinsäuresequenzvariante, oder komplementäre Sequenz davon, eines Säugetiers, bevorzugt eines Menschen, genomische oder RNA-Sequenz des für ENPP1-Protein kodierenden Gens, **dadurch gekennzeichnet, dass** die genannte isolierte Nukleinsäuresequenzvariante wenigstens die Anwesenheit des Haplotyps umfasst, der die drei nachfolgenden wie in Anspruch 1 und 4 bis 6 definierten SNPs umfasst.

20. Klonungs- und/oder Expressionsvektor, enthaltend eine Nukleinsäuresequenz nach Anspruch 19.

21. Wirtszelle, die durch einen Vektor nach Anspruch 20 transformiert ist.

22. Transgenes Säugetier, mit Ausnahme eines Menschen, **dadurch gekennzeichnet, dass** es eine transformierte Zelle nach Anspruch 21 oder eine Zelle umfasst, die ein rekombinantes, funktionelles ENPP1-Gen enthält, das wenigstens den Haplotyp umfasst, der die drei nachfolgenden in Anspruch 1 und 4 bis 6 definierten SNPs umfasst.

23. Anwendung der Zelle nach Anspruch 21 oder eines transgenen Säugetiers nach Anspruch 22, zum Screening chemischer oder biochemischer Verbindungen zur Behandlung oder zur Vorbeugung von Fettsucht oder der mit Fettsucht assoziierten Pathologie.

24. Verfahren zur Auswahl einer chemischen oder biochemischen Verbindung zur Behandlung oder zur Vorbeugung von Fettsucht oder der mit Fettsucht assoziierten Pathologie, **dadurch gekennzeichnet, dass** es die nachfolgenden

Schritte umfasst:

- 5 a) Inkontaktbringen einer Zelle nach Anspruch 21 oder eines transgenen Säugetiers nach Anspruch 26 mit der zu testenden chemischen oder biochemischen Verbindung; und
b) Beobachten, ob das Expressionsprodukt des ENPP1-Gens in der genannten Zelle oder dem genannten Säugetier erhöht ist.

10 25. Verfahren zur Auswahl einer chemischen oder biochemischen Verbindung zur Behandlung oder zur Vorbeugung von Fettsucht oder der mit Fettsucht assoziierten Pathologie nach Anspruch 24, **dadurch gekennzeichnet, dass:**

- 15 - in Schritt a) die zu testende chemische oder biochemische Verbindung an ein transgenes Säugetier nach Anspruch 22 oder an ein transgenes Säugetier verabreicht wird, das das ENPP1-Protein übermäßig exprimiert; und
- in Schritt b) beobachtet wird, ob die Serum- oder Plasma-ENPP1-Proteinkonzentration im Vergleich zu einer Referenz erhöht ist,

wobei die chemische oder biochemische Verbindung ausgewählt ist, falls eine Erhöhung der genannten Konzentration beobachtet wird.

20 Revendications

25 1. Procédé pour déterminer si un sujet est exposé à un risque augmenté de développer une obésité ou une pathologie associée à l'obésité à partir d'un échantillon biologique dudit sujet contenant de l'ADN génomique ou de l'ARN, ledit procédé comprenant les étapes consistant à :

a) déterminer sur au moins une séquence d'allèle d'un gène ou d'ARN codant pour la protéine ENPP1, si ladite séquence, ou des fragments de celle-ci, contient l'haplotype comprenant les trois SNP suivants :

- 30 - K121Q
- IVS20 delT-11, et
- un SNP localisé dans la séquence du domaine UTR 3' et où la première séquence de type sauvage du domaine UTR 3' de 1170 pb (comprenant le codon TGA) possède la séquence SEQ ID NO: 15 ; et

35 b) observer si le sujet est exposé ou non à un risque augmenté de développer une obésité ou une pathologie associée à l'obésité si ladite séquence, ou des fragments de celle-ci, contient ledit haplotype, la présence dudit haplotype indiquant que ledit sujet est exposé à un risque augmenté de développer une obésité ou une pathologie associée à l'obésité.

40 2. Procédé selon la revendication 1, où la pathologie associée à l'obésité est sélectionnée dans le groupe consistant en le diabète de type 2, l'hypertension, l'hypercholestérolémie et l'hyperlipidémie.

3. Procédé selon la revendication 1 ou 2, où la pathologie associée à l'obésité est le diabète de type 2.

45 4. Procédé selon l'une des revendications 1 à 3, où ledit SNP localisé dans le domaine UTR 3' est sélectionné dans le groupe de SNP consistant en A>G +1044 TGA, A>C +1092 TGA et C>T +1157 TGA.

50 5. Procédé selon la revendication 4, où ledit SNP localisé dans le domaine UTR 3' est sélectionné dans le groupe consistant en A>G +1044 TGA et A>C +1092 TGA.

6. Procédé selon la revendication 5, où ledit SNP localisé dans le domaine UTR 3' est A>G +1044 TGA.

55 7. Procédé pour déterminer si un sujet est exposé à un risque augmenté de développer une obésité ou une pathologie associée à l'obésité, **caractérisé en ce que** ledit procédé comprend l'étape a) et b) du procédé selon l'une des revendications 1 à 6, dans lequel :

- à l'étape a), on détermine si le sujet est homozygote ou non pour ledit haplotype comprenant lesdits trois SNP ; et
- à l'étape b), on observe si le sujet est exposé ou non à un risque augmenté de développer une obésité ou

EP 1 891 237 B9

une pathologie associée à l'obésité en observant si le sujet est homozygote pour ledit haplotype, le caractère homozygote indiquant que ledit sujet est exposé à un risque augmenté de développer une obésité ou une pathologie associée à l'obésité.

- 5 **8.** Procédé selon l'une des revendications 1 à 7, dans lequel l'étape a), qui consiste à déterminer la présence ou l'absence dudit haplotype à risque à trois allèles, est réalisée par un procédé comprenant une étape d'amplification de la séquence d'ADN génomique ou d'ARN, ou de fragments de celle-ci, susceptible de contenir lesdits SNP du haplotype par une réaction en chaîne par polymérase (PCR).
- 10 **9.** Procédé selon la revendication 7, dans lequel l'étape de PCR est réalisée à partir d'ADN génomique, d'ARNm ou d'ADNc, éventuellement après une étape de transcription inverse, ou de fragments de ceux-ci, en utilisant les jeux d'amorces suivants :
- 15 a) 5'-TCATACTCAGGAAGACAGCAA-3' (amorce sens SEQ ID NO: 2) et 5'-CAATAGCCATGACTCCTAA-3' (amorce inverse SEQ ID NO: 3) pour le SNP « K121Q » ;
b) 5'-AGCATTTTTACTTTAGTT-3' (amorce sens SEQ ID NO: 4) et 5'-ATAATGTATGGTAGGATTT-3' (amorce antisens SEQ ID NO: 5) pour le SNP IVS20 delT-11 ; et
c) 5'-ATATTCCTATCCTGCTCACT-3' (amorce sens SEQ ID NO: 6) et 5'-TGCAGCTGGCCCTTAGGCCG-3' (amorce antisens SEQ ID NO: 7) pour le SNP A>G +1044 TGA.
- 20 **10.** Procédé selon l'une des revendications 1 à 9, dans lequel ledit échantillon biologique est un échantillon de foie, des adipocytes ou des cellules bêta pancréatiques lorsque la présence dudit haplotype est déterminée sur une séquence d'ARN ou des fragments de celle-ci.
- 25 **11.** Procédé pour déterminer si un sujet est exposé à un risque augmenté de développer une obésité ou une pathologie associée à l'obésité à partir d'un échantillon de plasma ou de sérum dudit sujet, ledit procédé comprenant les étapes consistant à :
- 30 a) déterminer la concentration de la protéine ENPP1 dans le sérum ou le plasma et en outre la présence de l'haplotype comprenant les trois SNP, tel que défini dans les revendications 1 et 4 à 6 ; et
b) observer si le sujet est exposé ou non à un risque augmenté de développer une obésité ou une pathologie associée à l'obésité en observant si ladite concentration est significativement plus élevée que le taux normal, la présence de ladite concentration significativement plus élevée indiquant que le sujet est exposé à un risque augmenté de développer une obésité ou une pathologie associée à l'obésité.
- 35 **12.** Procédé selon la revendication 11, dans lequel ladite concentration significativement plus élevée de la protéine ENPP1 dans le sérum ou le plasma correspond à au moins plus de 10 %, de préférence plus de 15 % de la concentration normale dans le sérum ou le plasma.
- 40 **13.** Procédé selon la revendication 11 ou 12, dans lequel ladite concentration de la protéine ENPP1 dans le sérum ou le plasma est déterminée en utilisant des anticorps capables de spécifiquement reconnaître la protéine ENPP1.
- 45 **14.** Procédé selon la revendication 13, **caractérisé en ce que** lesdits anticorps sont mis en contact avec le sérum ou le plasma à tester, dans des conditions permettant la formation d'un complexe immunologique spécifique entre la protéine ENPP1 et ledit anticorps, et **en ce que** les complexes immunologiques formés sont quantifiés.
- 50 **15.** Procédé selon l'une des revendications 11 à 14, dans lequel la pathologie associée à l'obésité est sélectionnée dans le groupe consistant en le diabète de type 2, l'hypertension, l'hypercholestérolémie et l'hyperlipidémie.
- 55 **16.** Procédé selon l'une quelconque des revendications 11 à 15, dans lequel la pathologie associée à l'obésité est le diabète de type 2.
- 17.** Kit pour déterminer si un sujet est exposé à un risque augmenté de développer une obésité ou une pathologie associée à l'obésité, comprenant au moins une paire d'amorces ou un jeu d'une paire d'amorces capable d'amplifier un fragment d'ADN génomique ou d'ARN codant pour la protéine ENPP1 et susceptible de contenir l'haplotype comprenant les trois SNP suivants :

- K121Q,

EP 1 891 237 B9

- IVS20 delT-11, et
- un SNP localisé dans la séquence du domaine UTR 3' et où la première séquence de type sauvage du domaine UTR 3' de 1170 pb (comprenant le codon TGA) possède la séquence SEQ ID NO: 15 ;

5 et dans lequel ledit au moins un jeu d'une paire d'amorces est :

- a) 5'-TCATACTCAGGAAGACAGCAA-3' (amorce sens SEQ ID NO: 2) et 5'-CAATAGCCATGACTCCTAA-3' (amorce inverse SEQ ID NO: 3) pour le SNP « K121Q » ;
- 10 b) 5'-AGCATT TTTACTTTAGTT-3' (amorce sens SEQ ID NO: 4) et 5'-ATAATGTATGGTAGGATTT-3' (amorce antisens SEQ ID NO: 5) pour le SNP IVS20 del-T11 ; et
- c) 5'-ATATTCCTATCCTGCTCACT-3' (amorce sens SEQ ID NO: 6) et 5'-TGCAGCTGGCCCTTAGGCCG-3' (amorce antisens SEQ ID NO: 7) pour le SNP A>G +1044 TGA.

15 **18.** Kit pour déterminer si un sujet est exposé à un risque augmenté de développer une obésité ou une pathologie associée à l'obésité, comprenant au moins une paire d'amorces ou un jeu d'une paire d'amorces capable d'amplifier un fragment d'ADN génomique ou d'ARN codant pour la protéine ENPP1 et susceptible de contenir l'haplotype comprenant les trois SNP suivants :

- K121Q,
- 20 - IVS20 delT-11, et
- un SNP localisé dans la séquence du domaine UTR 3' et où la première séquence de type sauvage du domaine UTR 3' de 1170 pb (comprenant le codon TGA) possède la séquence SEQ ID NO: 15 ;

25 et dans lequel ledit au moins un jeu d'une paire d'amorces est :

- a) -5'-CTTCCCAATCACTACAGCATTGTCA-3' (amorce sens, exon 7, SEQ ID NO: 8), et
- 5'-TTTCAGACCATCCATCAGCATACCAAC-3' (amorce antisens, exon 12, SEQ ID NO: 9) ; et pour le SNP d'UTR 3' ;
- 30 b) -5'-GTCCTGTGTTTGACTTTGATTATGA-3' (amorce sens, exon 23, SEQ ID NO: 10), et
- 5'-CCCTTAGGCCGTTGAAGAATGGTCA-3' (amorce antisens UTR 3', SEQ ID NO: 11).

35 **19.** Séquence d'acide nucléique variante isolée, ou séquence complémentaire de celle-ci, d'un mammifère, de préférence un humain, séquence génomique ou d'ARN du gène codant pour la protéine ENPP1, **caractérisée en ce que** ladite séquence d'acide nucléique variante isolée comprend au moins la présence de l'haplotype comprenant les trois SNP suivants, tel que défini dans les revendications 1 et 4 à 6.

20. Vecteur de clonage et/ou d'expression contenant une séquence d'acide nucléique selon la revendication 19.

21. Cellule hôte transformée par un vecteur selon la revendication 20.

40 **22.** Mammifère transgénique, à l'exception d'un humain, **caractérisé en ce qu'il** comprend une cellule transformée selon la revendication 21 ou une cellule contenant un gène ENPP1 fonctionnel recombinant comprenant au moins l'haplotype comprenant les trois SNP suivants, tel que défini dans les revendications 1 et 4 à 6.

45 **23.** Utilisation d'une cellule selon la revendication 21 ou d'un mammifère transgénique selon la revendication 22, pour cribler des composés chimiques ou biochimiques pour le traitement ou la prévention de l'obésité ou d'une pathologie associée à l'obésité.

50 **24.** Procédé pour sélectionner un composé chimique ou biochimique pour le traitement ou la prévention de l'obésité ou d'une pathologie associée à l'obésité, **caractérisé en ce qu'il** comprend l'étape consistant à :

- a) mettre en contact une cellule selon la revendication 21 ou un mammifère transgénique selon la revendication 26 avec ledit composé chimique ou biochimique à tester ; et
- 55 b) observer si le produit d'expression du gène ENPP1 est réduit ou non dans ladite cellule ou ledit mammifère.

25. Procédé pour sélectionner un composé chimique ou biochimique pour le traitement ou la prévention de l'obésité ou d'une pathologie associée à l'obésité selon la revendication 24, **caractérisé en ce que** :

EP 1 891 237 B9

- à l'étape a), le composé chimique ou biochimique à tester est administré à un mammifère transgénique selon la revendication 22 ou à un mammifère transgénique surexprimant la protéine ENPP1 i et
- à l'étape b), on observe si la concentration de la protéine ENPP1 dans le sérum ou le plasma est réduite par comparaison à une référence,

5

ledit composé chimique ou biochimique étant sélectionné si une réduction de ladite concentration est observée.

10

15

20

25

30

35

40

45

50

55

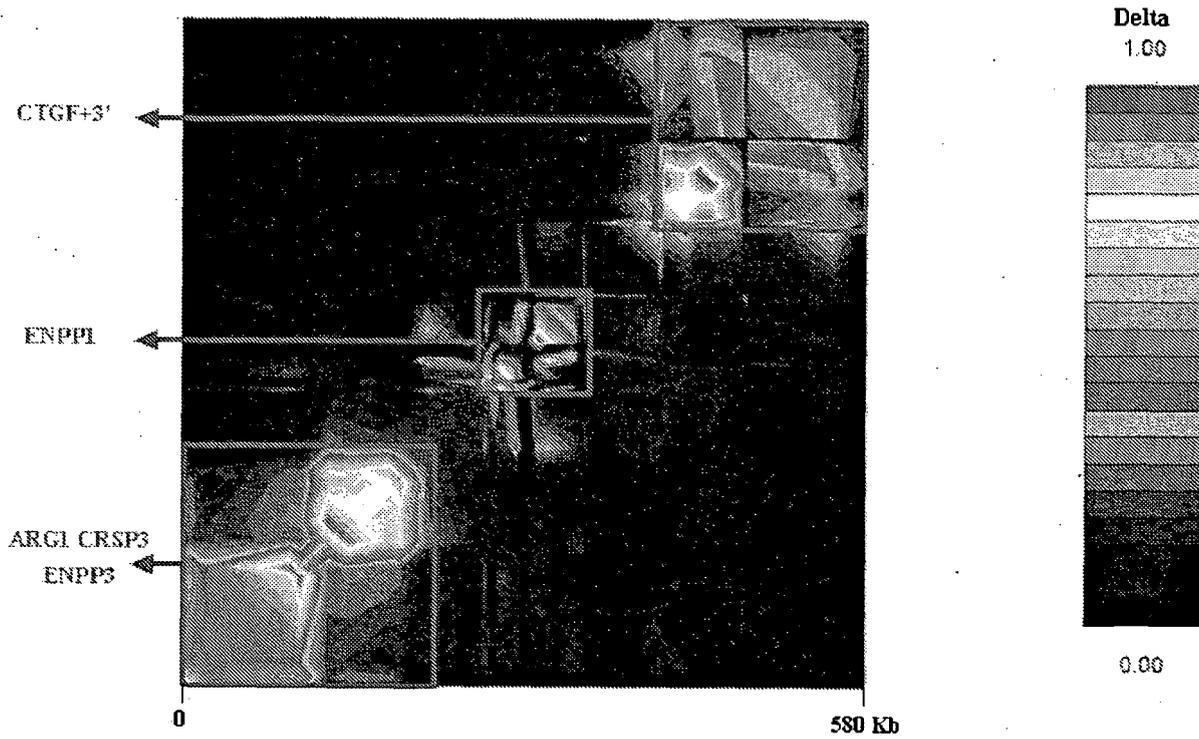


FIGURE 1

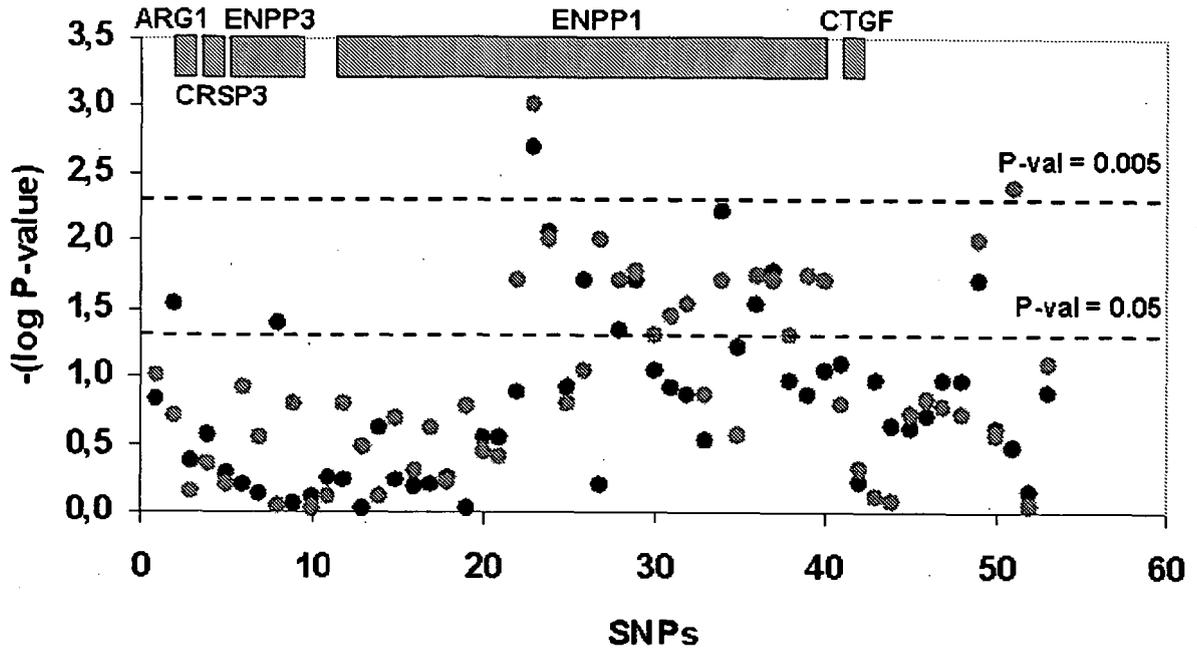


FIGURE 2

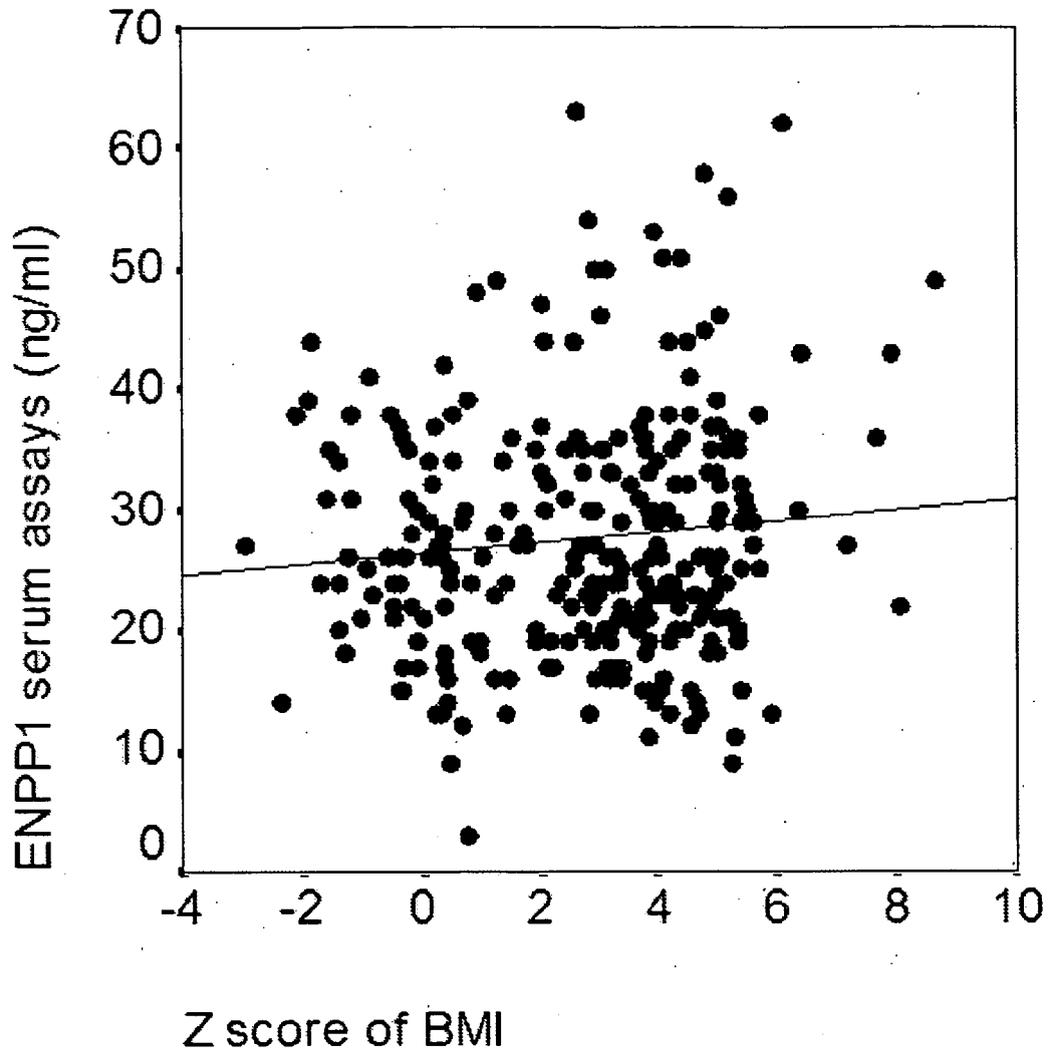


FIGURE 3

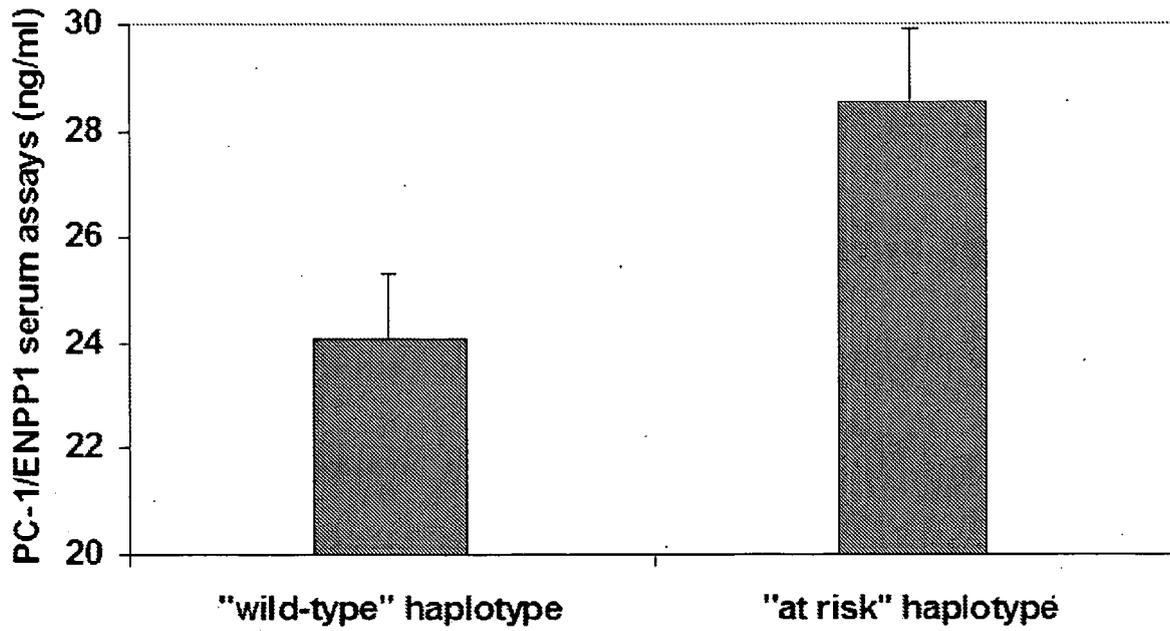


FIGURE 4

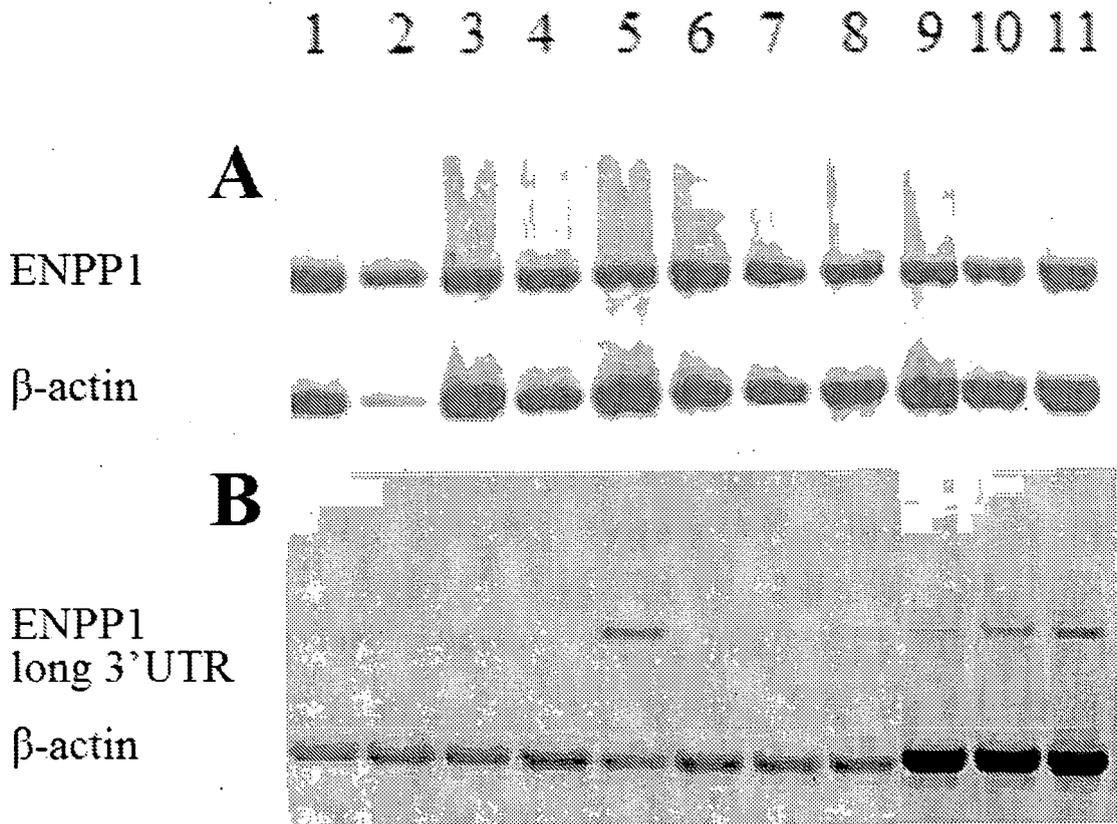


FIGURE 5A - FIGURE 5B

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 6465185 B1 [0006]
- EP 05291298 A [0102]

Non-patent literature cited in the description

- **Sambrook et al.** Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, 1989 [0056]
- **Meyre, D. et al.** A genome-wide scan for childhood obesity-associated traits in French families shows significant linkage on chromosome 6q22.31-q23.2. *Diabetes*, 2004, vol. 53, 803-11 [0101]
- **Atwood, L.D. et al.** Genomewide Linkage Analysis of Body Mass Index across 28 Years of the Framingham Heart Study. *Am. J. Hum. Genet.*, 2002, vol. 71, 1044-50 [0101]
- **Duggirala, R. et al.** A major locus for fasting insulin concentrations and insulin resistance on chromosome 6q with strong pleiotropic effects on obesity-related phenotypes in nondiabetic Mexican Americans. *Am. J. Hum. Genet.*, 2001, vol. 68, 1149-64 [0101]
- **Abney, M. ; Ober, C. ; McPeck, M.S.** Quantitative-trait homozygosity and association mapping and empirical genomewide significance in large, complex pedigrees: fasting serum- insulin level in the Hutterites. *Am. J. Hum. Genet.*, 2002, vol. 70, 920-34 [0101]
- **Gelder Ehm, M.G. et al.** Genomewide search for type 2 diabetes susceptibility genes in four American populations. *Am. J. Hum. Genet.*, 2000, vol. 66, 1871-81 [0101]
- **Ghosh, S. et al.** The Finland-United States investigation of non-insulin-dependent diabetes mellitus genetics (FUSION) study. I. An autosomal genome scan for genes that predispose to type 2 diabetes. *Am. J. Hum. Genet.*, 2000, vol. 67, 1174-85 [0101]
- **Demenais, F. et al.** A meta-analysis of four European genome screens (GIFT Consortium) shows evidence for a novel region on chromosome 17p11.2-q22 linked to type 2 diabetes. *Hum Mol Genet*, 2003, vol. 12, 1865-73 [0101]
- **Xiang, K. et al.** Genome-Wide Search for Type 2 Diabetes/Impaired Glucose Homeostasis Susceptibility Genes in the Chinese: Significant Linkage to Chromosome 6q21-q23 and Chromosome 1q21-q24. *Diabetes*, 2004, vol. 53, 228-34 [0101]
- **Maddux, B.A. et al.** Membrane, glycoprotein PC-1 and insulin resistance in non-insulin- dependent diabetes mellitus. *Nature*, 1995, vol. 373, 448-51 [0101]
- **Dong, H. et al.** Increased Hepatic Levels of the Insulin Receptor Inhibitor, PC-1/NPP1, Induce Insulin Resistance and Glucose Intolerance. *Diabetes*, 2005, vol. 54, 367-72 [0101]
- **Costanzo, B.V. et al.** The Q allele variant (GLN121) of membrane glycoprotein PC-1 interacts with the insulin receptor and inhibits insulin signaling more effectively than the common K allele variant (LYS121). *Diabetes*, 2001, vol. 50, 831-6 [0101]
- **Rolland-Cachera, M.F. et al.** Body Mass Index variations: centiles from birth to 87 years. *Eur J Clin Nutr*, 1991, vol. 45, 13-21 [0101]
- **Maddux, B.A. ; Goldfine, I.D.** Membrane glycoprotein PC-1 inhibition of insulin receptor function occurs via direct interaction with the receptor alpha-subunit. *Diabetes*, 2000, vol. 49, 13-19 [0101]
- **Pizzuti, A. et al.** A polymorphism (K121Q) of the human glycoprotein PC-1 gene coding region is strongly associated with insulin resistance. *Diabetes*, 1999, vol. 48, 1881-4 [0101]
- **Gu, H.F. et al.** Association between the human glycoprotein PC-1 gene and elevated glucose and insulin levels in a paired-sibling analysis. *Diabetes*, 2000, vol. 49, 1601-3 [0101]
- **Kubaszek, A. ; Pihlajamaki, J. ; Karhapaa, P. ; Vauhkonen, I. ; Laakso, M.** The K121Q polymorphism of the PC-1 gene is associated with insulin resistance but not with dyslipidemia. *Diabetes Care*, 2003, vol. 26, 464-7 [0101]
- **Abate, N. et al.** Genetic polymorphism PC-1 K121Q and ethnic susceptibility to insulin resistance. *J Clin Endocrinol Metab*, 2003, vol. 88, 5927-34 [0101]
- **Hamaguchi, K. et al.** The PC-1 Q121 allele is exceptionally prevalent in the Dominican Republic and is associated with type 2 diabetes. *J Clin Endocrinol Metab*, 2004, vol. 89, 1359-64 [0101]
- **Frittitta, L. et al.** A cluster of three single nucleotide polymorphisms in the 3'-untranslated region of human glycoprotein PC-1 gene stabilizes PC-1 mRNA and is associated with increased PC-1 protein content and insulin resistance-related abnormalities. *Diabetes*, 2001, vol. 50, 1952-5 [0101]

- **Hager, J. et al.** A genome-wide scan for human obesity genes reveals a major susceptibility locus on chromosome 10. *Nat Genet*, 1998, vol. 20, 304-8 [0101]
- **Deng, H.W. ; Chen, W.M. ; Recker, R.R.** QTL fine mapping by measuring and testing for Hardy-Weinberg and linkage disequilibrium at a series of linked marker loci in extreme samples of populations. *Am J Hum Genet*, 2000, vol. 66, 1027-45 [0101]
- **Eaves, I.A. et al.** Transmission ratio distortion at the INS-IGF2 VNTR. *Nat Genet*, 1999, vol. 22, 324-5 [0101]
- **Li, C. ; Scott, L.J. ; Boehnke, M.** Assessing whether an allele can account in part for a linkage signal: the Genotype-IBD Sharing Test (GIST). *Am J Hum Genet*, 2004, vol. 74, 418-31 [0101]
- **Rutsch, F. et al.** PC-1 nucleoside triphosphate pyrophosphohydrolase deficiency in idiopathic infantile arterial calcification. *Am J Pathol*, 2001, vol. 158, 543-54 [0101]
- **Weill, J. ; Vanderbecken, S. ; Froguel, P.** Understanding the rising incidence of type 2 diabetes in adolescence. *Arch Dis Child*, 2004, vol. 89, 502-4 [0101]
- **Jenkinson, C.P. et al.** *Comprehensive analysis of snps in ENPP1: association with diabetes*, 2004 [0101]
- **Ohlsson, H. ; Karlsson, K. ; Edlund, T.** IPF 1-, a homeodomain-containing transactivator of the insulin gene. *Embo J*, 1993, vol. 12, 4251-9 [0101]
- **Hugot, J.P. et al.** Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature*, 2001, vol. 411, 599-603 [0101]
- **Hampe, J. et al.** Evidence for a NOD2-independent susceptibility locus for inflammatory bowel disease on chromosome 16p. *Proc Natl Acad Sci USA*, 2002, vol. 99, 321-6 [0101]
- **Holder, J.L., Jr. ; Butte, N.F. ; Zinn, A.R.** Profound obesity associated with a balanced translocation that disrupts the SIM1 gene. *Hum Mol Genet*, 2000, vol. 9, 101-8 [0101]
- **Michaud, J.L. et al.** Sim1 haploinsufficiency causes hyperphagia, obesity and reduction of the paraventricular nucleus of the hypothalamus. *Hum Mol Genet*, 2001, vol. 10, 1465-73 [0101]
- **Tregouet, D.A. ; Escolano, S. ; Tiret, L. ; Mallet, A. ; Golmard, J.L.** A new algorithm for haplotype-based association analysis: the Stochastic-EM algorithm. *Ann Hum Genet*, 2004, vol. 68, 165-77 [0101]
- **Schwartz, M.W.** Progress in the search for neuronal mechanisms coupling type 2 diabetes to obesity. *J Clin Invest*, 2001, vol. 108, 963-4 [0101]
- **Bruning, J.C. et al.** Role of brain insulin receptor in control of body weight and reproduction. *Science*, 2000, vol. 289, 2122-5 [0101]
- **Kim, J.K. et al.** Redistribution of substrates to adipose tissue promotes obesity in mice with selective insulin resistance in muscle. *J Clin Invest*, 2000, vol. 105, 1791-7 [0101]
- **Michael, M.D. et al.** Loss of insulin signaling in hepatocytes leads to severe insulin resistance and progressive hepatic dysfunction. *Mol Cell*, 2000, vol. 6, 87-97 [0101]
- **Kulkarni, R.N. et al.** Tissue-specific knockout of the insulin receptor in pancreatic beta cells creates an insulin secretory defect similar to that in type 2 diabetes. *Cell*, 1999, vol. 96, 329-39 [0101]
- **Um, S.H. et al.** Absence of S6K1 protects against age- and diet-induced obesity while enhancing insulin sensitivity. *Nature*, 2004, vol. 431, 200-205 [0101]
- **Barroso, I. et al.** Candidate Gene Association Study in Type 2 Diabetes Indicates a Role for Genes Involved in beta-Cell Function as Well as Insulin Action. *PLoS Biol* 1, 2003, E20 [0101]
- **Odeleye, O.E. ; de Courten, M. ; Pettitt, D.J. ; Ravussin, E.** Fasting hyperinsulinemia is a predictor of increased body weight gain and obesity in Pima Indian children. *Diabetes*, 1997, vol. 46, 1341-5 [0101]
- **Johnson, M.S. ; Figueroa-Colon, R. ; Huang, T.T. ; Dwyer, J.H. ; Goran, M.I.** Longitudinal changes in body fat in African American and Caucasian children: influence of fasting insulin and insulin sensitivity. *J Clin Endocrinol Metab*, 2001, vol. 86, 3182-7 [0101]
- **Poskitt, E.M.** Defining childhood obesity: the relative body mass index (BMI). European Childhood Obesity group. *Acta Paediatr*, 1995, vol. 84, 961-3 [0101]
- **Hercberg, S. et al.** A primary prevention trial using nutritional doses of antioxidant vitamins and minerals in cardiovascular diseases and cancers in a general population: the SU.VI.MAX study-design, methods, and participant characteristics. SUpplementation en Vitamines et Mineraux Antioxydants. *Control Clin Trials*, 1998, vol. 19, 336-51 [0101]
- **Lafay, L. et al.** Determinants and nature of dietary underreporting in a free-living population: the Fleurbaix Laventie Ville Sante (FLVS) Study. *Int J Obes Relat Metab Disord*, 1997, vol. 21, 567-73 [0101]
- **Oberkofler, H. et al.** Complex haplotypes of the PGC-lalpha gene are associated with carbohydrate metabolism and type 2 diabetes. *Diabetes*, 2004, vol. 53, 1385-93 [0101]
- **Cole, T.J.** The LMS method for constructing normalized growth standards. *Eur J Clin Nutr*, 1990, vol. 44, 45-60 [0101]
- **Seltzer, H.S. ; Allen, E.W. ; Herron, A.L., Jr. ; Brennan, M.T.** Insulin secretion in response to glycemic stimulus: relation of delayed initial release to carbohydrate intolerance in mild diabetes mellitus. *J Clin Invest*, 1967, vol. 46, 323-35 [0101]
- **Risch, N.** Linkage strategies for genetically complex traits. I. Multilocus models. *Am J Hum Genet*, 1990, vol. 46, 222-228 [0101]

EP 1 891 237 B9

- **Kruglyak, L. ; Lander, E.S.** Complete multipoint sib-pair analysis of qualitative and quantitative traits. *Am J Hum Genet*, 1995, vol. 57, 439-54 **[0101]**
- **Sham, P.C. ; Curtis, D.** Monte Carlo tests for associations between disease and alleles at highly polymorphic loci. *Ann Hum Genet*, 1995, vol. 59 (1), 97-105 **[0101]**
- **Dudbridge, F. ; Koeleman, B.P. ; Todd, J.A. ; Clayton, D.G.** Unbiased application of the transmission/disequilibrium test to multilocus haplotypes. *Am J Hum Genet*, 2000, vol. 66, 2009-12 **[0101]**
- **Lukowiak, B. et al.** Identification and purification of functional human beta-cells by a new specific zinc-fluorescent probe. *J Histochem Cytochem*, 2001, vol. 49, 519-28 **[0101]**