

**(12) STANDARD PATENT  
(19) AUSTRALIAN PATENT OFFICE**

**(11) Application No. AU 2015238195 C1**

(54) Title  
**Mitochondrial markers of neurodegenerative diseases**

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

(21) Application No: **2015238195** (22) Date of Filing: **2015.03.27**

(87) WIPO No: **WO15/144964**

(30) Priority Data

(31) Number **P201430444** (32) Date **2014.03.28** (33) Country **ES**

(43) Publication Date: **2015.10.01**  
(44) Accepted Journal Date: **2019.07.18**  
(44) Amended Journal Date: **2022.03.10**

(71) Applicant(s)  
**Fundacio Institut D'Investigacio Biomedicia De Bellvitge (Idibell);Universitat de Barcelona**

(72) Inventor(s)  
**Barrachina Castillo, Marta;Ferrer Abizanda, Isidre;Blanch Lozano, Marta**

(74) Agent / Attorney  
**Spruson & Ferguson, GPO Box 3898, Sydney, NSW, 2001, AU**

(56) Related Art  
**TAKASAKI, S., "Mitochondrial SNPs associated with Japanese centenarians, Alzheimer's patients, and Parkinson's patients", COMPUTATIONAL BIOLOGY AND CHEMISTRY, (2008), vol. 32, no. 5, ISSN 1476-9271, pages 332 - 337**  
**TAKASAKI, S., "Japanese Alzheimer's disease and other complex disorders diagnosis based on mitochondrial SNP haplogroups", INTERNATIONAL JOURNAL OF ALZHEIMER'S DISEASE, (2012), vol. 2012, ISSN 2090-0252, page 245038**

## (12) SOLICITUD INTERNACIONAL PUBLICADA EN VIRTUD DEL TRATADO DE COOPERACIÓN EN MATERIA DE PATENTES (PCT)

(19) Organización Mundial de la Propiedad Intelectual Oficina internacional



WIPO | PCT



(10) Número de Publicación Internacional

WO 2015/144964 A3

(43) Fecha de publicación internacional  
1 de octubre de 2015 (01.10.2015)(51) Clasificación Internacional de Patentes:  
*C12Q 1/68* (2006.01)(21) Número de la solicitud internacional:  
PCT/ES2015/070230(22) Fecha de presentación internacional:  
27 de marzo de 2015 (27.03.2015)

(25) Idioma de presentación: español

(26) Idioma de publicación: español

(30) Datos relativos a la prioridad:  
P201430444 28 de marzo de 2014 (28.03.2014) ES(71) Solicitantes: INSTITUT D'INVESTIGACIÓ BIOMÈDICA DE BELLVITGE (IDIBELL) [ES/ES]; Hospital Duran i Reynals, 3<sup>a</sup> planta, Gran Via de l'Hospitalet, 199, E-08908 Hospitalet de Llobregat (ES). UNIVERSITAT DE BARCELONA [ES/ES]; Centro de Patentes UB, Baldiri Reixac, 4, Torre D, E-08028 Barcelona (ES).

(72) Inventores: BARRACHINA CASTILLO, Marta; Hospital Duran i Reynals, 3a planta, Gran Via de l'Hospitalet, 199, E-08907 Hospitalet de Llobregat-Barcelona (ES). FERRER ABIZANDA, Isidre; Servicio de Anatomía Patológica, Hospital Universitari de Bellvitge, Feixa Llarga, s/n, E- 08907 Hospitalet de Llobregat-Barcelona (ES). BLANCH LOZANO, Marta; Hospital Duran i Reynals, 3a planta, Gran Via de l'Hospitalet, 199, E-08907 Hospitalet de Llobregat-Barcelona (ES).

(74) Mandatario: ARIAS SANZ, Juan; ABG Patentes, S.L, Edificio Euromor, Avenida de Burgos 16D, E-28036 Madrid (ES).

(81) Estados designados (a menos que se indique otra cosa, para toda clase de protección nacional admisible): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Estados designados (a menos que se indique otra cosa, para toda clase de protección regional admisible): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), euroasiática (AM, AZ, BY, KG, KZ, RU, TJ, TM), europea (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

## Publicada:

- con informe de búsqueda internacional (Art. 21(3))
- antes de la expiración del plazo para modificar las reivindicaciones y para ser republicada si se reciben modificaciones (Regla 48.2(h))
- con la parte de lista de secuencias de la descripción (Regla 5.2(a))

(88) Fecha de publicación del informe de búsqueda internacional:

17 de diciembre de 2015

(54) Title: MITOCHONDRIAL MARKERS OF NEURODEGENERATIVE DISEASES

(54) Título : MARCADORES MITOCONDRIALES DE ENFERMEDADES NEURODEGENERATIVAS

(57) Abstract: The invention relates to an *in vitro* method for diagnosing or determining the risk of a subject developing a neurodegenerative disease based on the determination of the methylation pattern in certain regions of the mitochondrial DNA of said subject or based on the determination of the nucleotide in the polymorphic position 16519 of the mitochondrial DNA of said subject. The invention further relates to nucleic acids suitable for implementing the invention.(57) Resumen: La presente invención se refiere a un método *in vitro* para diagnosticar o para determinar el riesgo de desarrollar una enfermedad neurodegenerativa en un sujeto basado en la determinación del patrón de metilación en ciertas regiones del ADN mitocondrial de dicho sujeto o en la determinación del nucleótido en la posición polimórfica 16519 del ADN mitocondrial de dicho sujeto. Finalmente, la presente invención se refiere a ácidos nucleicos adecuados para poner en práctica la invención.

**DESCRIPTION**

Mitochondrial markers of neurodegenerative diseases

**5 Field of the invention**

The present invention is classified as part of the diagnosis methods of neurological diseases.

**Background of the invention**

0 A considerable amount of neurodegenerative diseases are caused by or associated with altered mitochondrial function.

5 Alzheimer's disease (AD) and Parkinson's disease (PD) are within this group. The pathophysiological characteristics of AD and PD are related to deposits of aggregated protein. Specifically, AD is associated with the formation of intracellular phosphorylated tau aggregates in neurofibrillary tangles and extracellular aggregates of  $\beta$ -amyloid peptide in the senile plaques and PD is associated with the formation of abnormal  $\alpha$ -synuclein aggregates that constitute the main component of the so called Lewy bodies and Lewy neurites.

:0 25 It is accepted that Alzheimer's patients show decreased levels of the subunit ND4 in their brain tissue and that Parkinson's patients show reduced levels of ND6 in the substantia nigra. Furthermore, genetic studies have identified mutations in several COX genes and in the D-loop region as well as deletions of mtDNA in the brains of subjects with AD and in the substantial nigra of subjects with PD.

30 Different methods and strategies have been developed in the technology for diagnosis, prediction of onset and the development of neurodegenerative diseases, and specifically of AD and PD. In that way, diagnosis methods have been described for neurodegenerative diseases based on the identification of mutations in mitochondrial DNA through the employment of the RFLP (restriction fragment length polymorphism) technique or of other related techniques. The WO98038334 document describes a method of AD diagnosis based on the identification of mutations in COX genes. It has also proposed a method of PD diagnosis in a subject through the identification of single nucleotide polymorphisms in mitochondrial DNA samples of a subject (WO 2000063441). Other documents of the prior art describe diagnosis methods of Alzheimer's or Parkinson's disease based on the identification of polymorphisms

in nuclear protein encoding genes that control the process of mitochondrial transcription. Despite the efforts that have been made to date, there remains a need for reliable methods of diagnosis for neurodegenerative diseases such as AD and PD, as well for the diagnosis of the stage of such diseases and to prognosticate the evolution of the same.

5

Any reference to publications cited in this specification is not an admission that the disclosures constitute common general knowledge.

0

In a first aspect, the invention relates to an *in vitro* method to diagnose or determine the risk of development of a neurodegenerative disease selected from Alzheimer's disease and Parkinson's disease in a subject comprising to determine in a sample from said subject comprising mitochondrial DNA, methylation pattern in the D-loop region and/or in the ND1 gene, wherein the methylation pattern is determined in at least one site selected from the group consisting of:

- i. the CpG sites in the D-loop region shown in Table 1
- ii. the CpG sites in the ND1 gene shown in Table 2,
- iii. the CHG sites in the D-loop region shown in Table 3,
- iv. the CHG sites in the ND1 gene shown in Table 4, and/or
- v. the CHH sites in the D-loop region shown in Table 5,

wherein a hypermethylation in at least one of said CpG sites in the D-loop region, a hypermethylation in at least one of said CHG sites in the D-loop region, a hypermethylation in at least one of said CHH sites in the D-loop region, a hypomethylation in at least one of said

25 CpG sites in the ND1 gene and/or a hypomethylation in at least one of said CHG sites in the ND1 gene it is indicative that the subject suffers from Alzheimer's disease or that the subject is at a high risk of developing Alzheimer's disease; or

wherein a hypomethylation in at least one of said CpG sites in the D-loop region, a hypomethylation in at least one of said CHG sites in the D-loop region and/or a 30 hypomethylation in at least one of said CHH sites in the D-loop region it is indicative that the subject suffers from Parkinson's disease or that the subject at an elevated risk of developing Parkinson's disease.

In a second aspect, the invention relates to an *in vitro* method to select a subject for submission to treatment to prevent the progression of a neurodegenerative disease selected from Alzheimer's disease or Parkinson's disease in a subject comprising determining in a sample

from said subject containing mitochondrial DNA methylation pattern in the D-loop region and/or in the ND1 gene, where the methylation pattern is determined in at least one site selected from the group formed by:

- 5 i. the CpG sites in the D-loop region shown in Table 1,
- ii. the CpG sites in the ND1 gene shown in Table 2;
- iii. the CHG sites in the D-loop region shown in Table 3,
- iv. the CHG sites in the ND1 gene shown in Table 4, and/or
- v. the CHH sites in the D-loop region shown in Table 5

wherein a hypermethylation in at least one of said CpG sites in the D-loop region, a 0 hypermethylation in at least one of said CHG sites in the D-loop region, a hypermethylation in at least one of said CHH sites in the D-loop region, a hypomethylation in at least one of said CpG sites in the ND1 gene and/or a hypomethylation in at least one of said CHG sites in the ND1 gene it is indicative that the subject is a candidate to receive treatment aimed at preventing Alzheimer's disease; or

5 wherein a hypomethylation in at least one of said CpG sites in the D-loop region, a hypomethylation in at least one of said CHG sites of the D-loop region and/or a hypomethylation in at least one of said CHH sites in the D-loop region is indicative that the subject is a candidate to receive treatment aimed at preventing Parkinson's disease.

:0 In a third aspect, the invention relates to a method to monitor the progression of a neurodegenerative disease selected from Alzheimer's disease or Parkinson's disease in a subject comprising:

25 a) determining in a sample from said subject containing mitochondrial DNA the methylation pattern in the D-loop region, and/or in the ND1 gene, wherein the methylation pattern is determined in at least one site from the group formed by:

- i. the CpG sites in the D-loop region shown in Table 1,
- ii. the CpG sites in the ND1 gene shown in Table 2;
- iii. the CHG sites in the D-loop region shown in Table 3,
- iv. the CHG sites in the ND1 gene shown in Table 4, and/or the CHH sites in 30 the D-loop region shown in Table 5,

b) comparing the methylation pattern determined in step a) with said methylation pattern obtained in an earlier stage of the disease,

35 wherein a hypermethylation in at least one of said CpG sites in the D-loop region, a hypermethylation in at least one of the said CHG sites in the D-loop region, a hypermethylation in at least one of said CHH sites in the D-loop region, a hypomethylation in at least one of said CpG sites in the ND1 gene and/or a

hypomethylation in at least one of said CHG sites in the ND1 gene with respect to said methylation pattern determined in an earlier stage of disease is indicative of the progression of Alzheimer's disease; and

5 wherein a hypomethylation in at least one of said CpG sites in the D-loop region, a hypomethylation in at least one of said CHG sites in the D-loop region and/or a hypomethylation in at least one of the said CHH sites in the D-loop region with respect to said determined methylation pattern at an earlier stage of the disease is indicative of the progression of Parkinson's disease.

0 In a fourth aspect, the invention relates to an *in vitro* method to diagnose or determine the risk of development of Alzheimer's disease in a subject that contains in a sample comprising mitochondrial DNA from said subject, the nucleotide at the polymorphic position 16519 according to the sequence defined under the accession number NC\_012920 in the NCBI database, wherein the detection of nucleotide C at said polymorphic position or presence of 5 the nucleotide C at said polymorphic position in at least 60% of the molecules of mitochondrial DNA of the subject is indicative that the subject suffers from said disease or that the subject has an elevated risk of developing the disease.

In a fifth aspect, the invention relates to an *in vitro* method to select a subject to be submitted 10 to preventive treatment for Alzheimer's disease, that involves determining in a sample containing mitochondrial DNA from said subject, the nucleotide at the polymorphic position 16519 according to the sequence defined under accession number NC\_012920 in the NCBI database, where the detection of nucleotide C at said polymorphic position or the presence of 25 the nucleotide C at said polymorphic position in at least 60% of the mitochondrial DNA molecules from said subject is indicative that the subject is eligible to receive treatment aimed at preventing Alzheimer's disease.

In a sixth aspect, the invention relates to a nucleic acid selected from the group formed by:

30 (i) a nucleic acid comprising at least 9 contiguous nucleotides in a region of Mitochondrial DNA where said region comprises at least one methylation site selected from the group formed by:

35 a) the CpG sites in the D-loop region shown in Table 1,  
b) the CpG sites in the ND1 gene shown in Table 2;  
c) the CHG sites in the D-loop region shown in Table 3,  
d) the CHG sites in the ND1 gene shown in Table 4, and  
e) the CHH sites in the D-loop region shown in table 5,

(ii) a nucleic acid comprising at least 9 contiguous nucleotides in a region of mitochondrial DNA region where said region comprises at least one methylation site selected from the group formed by:

- 5 a) the CpG sites in the D-loop region shown in Table 1,
- b) the CpG sites in the ND1 gene shown in Table 2;
- c) the CHG sites in the D-loop region shown in Table 3,
- d) the CHG sites in the ND1 gene shown in Table 4, and
- e) the CHH sites in the D-loop region shown in Table 5.

where the position corresponding to the cytosine in the CpG, CHG or CHH site is uracil; and

0 (iii) a polynucleotide that hybridizes specifically with the nucleic acids of (i) or (ii)

In a seventh aspect, the invention relates to a kit comprising at least one oligonucleotide capable of specifically hybridizing and in a methylation-dependent manner with a mitochondrial DNA sequence comprising a methylation site selected from the group formed by:

- 5 i. the CpG sites in the D-loop region shown in Table 1,
- ii. the CpG sites in the ND1 gene shown in Table 2;
- iii. the CHG sites in the D-loop region shown in Table 3,
- iv. the CHG sites in the ND1 gene shown in Table 4, and/or
- v. the CHH sites in the D-loop region shown in Table 5

!0 In an eighth aspect, the invention relates to a kit comprising at least one oligonucleotide capable of specifically hybridizing at position 5' or at position 3' with respect to a methylation site in the selected mitochondrial DNA from the group formed by:

- 25 (i) the CpG sites in the D-loop region shown in Table 1,
- (ii) the CpG sites in the ND1 gene shown in Table 2;
- (iii) the CHG sites in the D-loop region shown in Table 3,
- (iv) the CHG sites in the ND1 gene shown in Table 4, and/or
- (v) the CHH sites in the D-loop region shown in Table 5

wherein the methylated cytosine in said position has converted to uracil or to another base that  
30 is distinguishable from cytosine in its hybridization properties.

In a ninth aspect, the invention refers to the use of the kits defined in the seventh and eight inventive aspects to determine the methylation pattern of mitochondrial DNA and to determine the diagnosis neurodegenerative disease in a subject selected from Alzheimer's disease and  
35 Parkinson's disease.

In a tenth aspect, the invention relates to use of a kit to determine the diagnosis of a neurodegenerative disease in a subject selected from Alzheimer's disease and Parkinson's disease, wherein the kit comprises at least one oligonucleotide capable of specifically and in a methylation-dependent manner hybridizing with a mitochondrial DNA sequence comprising

5 a methylation site selected from the group formed by:

- (i) the CpG sites in the D-loop region shown in Table 1,
- (ii) the CpG sites in the ND1 gene shown in Table 2,
- (iii) the CHG sites in the D-loop region shown in Table 3,
- (iv) the CHG sites in the ND1 gene shown in Table 4, and/or
- 0 (v) the CHH sites in the D-loop region shown in Table 5.

In an eleventh aspect, the invention relates to use of a kit to determine the diagnosis of a neurodegenerative disease in a subject selected from Alzheimer's disease and Parkinson's disease, wherein the kit comprises at least one oligonucleotide capable of specifically

5 hybridizing with a region 5' or 3' position with respect to a methylation site in mitochondrial DNA selected from the group consisting of:

- (i) the CpG sites in the D-loop region shown in Table 1,
- (ii) the CpG sites in the ND1 gene shown in Table 2,
- (iii) the CHG sites in the D-loop region shown in Table 3,
- (iv) the CHG sites in the ND1 gene shown in Table 4, and/or
- (v) the CHH sites in the D-loop region shown in Table 5,

wherein the methylated cytosine in said position has converted to uracil or another base that is distinguishable from cytosine in its hybridization properties.

25 The term 'comprise' and variants of the term such as 'comprises' or 'comprising' are used herein to denote the inclusion of a stated integer or stated integers but not to exclude any other integer or any other integers, unless in the context or usage an exclusive interpretation of the term is required.

30 **Brief description of the drawings**

**Figure 1:** Graphics Log2 (OR) for CpG (A) CHG (B) and CHH (C) sites in the amplicon D-Loop in the entorhinal cortex in AD pathology related cases. Methylation sites of 5' to 3' are represented in the x-axis. The sites marked with a diamond are differentially methylated sites (FDR < 0.05). Dots are OR estimate values, one for each site, and the band is the band of union of all the confidence intervals of 95%. C: samples control, AD: Alzheimer's Disease.

**Figure 2:** Grafics Log2(OR) for CpG (A) CHG (B) and CHH (C) sites in the ND1 amplicon in the entorhinal cortex of AD pathology related cases. In Methylation sites from 5' to 3' are represented in the x-axis. The places marked with a diamond are differentially methylated sites of union of all the confidence intervals of 95%. C: control samples, AD: Alzheimer's disease.

**Figure 3:** Graphics Log2 (OR) for CpG and non-CpG (CHG and CHH) sites in the amplicon D-loop in the substantia nigra in patients with PD. In the x-axis are Methylation sites from 5' to 3' are represented in the x-axis. The sites marked with a diamond are differentially methylated sites (FDR <0.05). Dots are estimated OR values, one for each site, and the band is the band of union of all the confidence intervals of 95%. C: control samples, PD: Parkinson's disease.

**Figure 4:** Graphics log2 (OR) for CpG (A) and CHG (B) sites in the amplicon D-Loop in the frontal cortex of APP/PS1 mice and wild mice (WT) of six months. methylation sites sorted by 5' to 3' are represented in the x-axis. The sites marked with a diamond are differentially methylated sites (FDR < 0.05). Dots are estimated OR values, one for each site, and the band is the band of union of all the confidence intervals of 95%. C: control samples, WT: wild, TG: Transgenic.

**Figure 5:** Graphics Log2 (OR) for CG (A) CHG (B) and CHH (C) sites in the amplicon D-Loop in the frontal cortex of APP/PS1 of mice three, six and twelve months. Methylation sites sorted by 5' to 3' are represented in the x-axis. The sites marked with a diamond are differentially methylated sites (FDR <0.05). Dots are estimated OR values, one for each site, and the band is the band of union of all the confidence intervals of 95%. C: control samples, TG: Transgenic.

### **Detailed description of the invention**

The authors of the present invention have developed a method for diagnosing neurodegenerative diseases based on the determination of the methylation pattern in a sample of mitochondrial DNA from a subject. The inventors have found that, surprisingly, there are variations in the methylation pattern in the D-loop region and the ND1 gene in subjects suffering from AD or PD when compared with healthy subjects such as demonstrated in the examples. Furthermore, the inventors have discovered differential methylation patterns associated with the development of these diseases.

### **First method of the invention**

The first feature of the invention relates to an *in vitro* method to diagnose or determine the risk of developing a neurodegenerative disease selected from Alzheimer's disease and Parkinson's disease in a subject (hereinafter, first method of the invention) that involves determining in a

sample from said subject comprising mitochondrial DNA, the methylation pattern in the D-loop region and/or in the ND1 gene, where the methylation pattern is determined in at least one site selected from the group formed by:

- (i) the CpG sites in the D-loop region shown in Table 1,
- (ii) the CpG sites in the ND1 gene shown in Table 2;
- (iii) the CHG sites in the D-loop region shown in Table 3,
- (iv) the CHG sites in the ND1 gene shown in Table 4, and/or
- (v) the CHH sites in the D-loop region shown in Table 5;

where there is hypermethylation in at least one of said CpG sites in the D-loop region, hypermethylation in at least one of said CHG sites in the D-loop region, hypermethylation in at least one of said CHH sites in the D-loop region, hypomethylation in at least one of said CpG sites in the ND1 gene and/or hypomethylation in at least one of said CHG sites in the ND1 gene, it is indicative that the subject suffers from Alzheimer's disease or that the subject is at an elevated risk of developing Alzheimer's disease or where there is hypomethylation in at least one of said CpG sites in the D-loop region, hypomethylation in at least one of said CHG sites in the D-loop region and/or hypomethylation in at least one of the said CHH sites in the D-loop region, it is indicative that the subject suffers from Parkinson's disease or that the subject is at an elevated risk of developing Parkinson's disease.

The term "diagnosis" as used in this document, refers to both the process of trying to determine and/or identify a possible disease in a subject, that is to say the diagnostic procedure, as well as the opinion reached through this process, that is to say, the diagnostic opinion. As such, it can also be seen as an attempt to classify the status of an individual in separate and distinct categories that allow medical decisions about treatment and prognosis to be taken. As will be understood by the person skilled in the art, such diagnosis may not be correct for 100% of the subjects to be diagnosed with, although it is preferred that it is.

However, the term requires that a statistically significant portion of subjects can be identified as suffering from a disease, in particular a neurodegenerative disease selected from Alzheimer's disease and Parkinson's disease in the context of the invention, or a predisposition thereto. The person skilled in the art may determine whether a part is statistically significant using different well known statistical evaluation tools, for example, by determining confidence intervals, determining the value of p, Student's t-test, the Mann-Whitney test, etc. (See Dowdy Wearden, 1983). Preferred confidence intervals are at least 50%, at least 60%, at least 70%, at least 80%, at least 90% or at least 95%. P values are preferably 0.05, 0.025, 0.001 or lower.

The expression "risk of developing a neurodegenerative disease" as used herein, refers to the predisposition, susceptibility or propensity of a subject to develop a neurodegenerative disease. The risk of developing a neurodegenerative disease generally implies that there is a high or low risk or higher or lower risk. Like that, a subject has a high risk of developing a neurodegenerative disease, particularly Alzheimer's disease or Parkinson's disease, has a likelihood of developing this disease of at least 50%, or at least 60%, or at least 70%, or at least 80%, or at least 90, or at least 95%, or at least 97%, or at least 98%, or at least 99%, or at least 100%. Similarly, a subject at a low risk of developing a neurodegenerative disease, particularly Alzheimer's disease or Parkinson's disease, is a subject having at least one chance of developing the disease from at least 0%, or at least one 1%, or at least 2%, or at least 3%, or at least 5%, or at least 10%, or at least 20%, or at least 30%, or at least 40%, or at least 49%.

In general, the expression "predict risk", "risk prediction", or similar, refers to the risk that a patient has of developing a neurodegenerative disease selected from Alzheimer's disease or Parkinson's disease, either high or low. As will be understood by the person skilled in the art, the prediction (or risk), albeit preferred, does not need to be correct for 100% of the subjects to be evaluated, but it is preferable that it is. The term, however, requires that a statistically significant portion of subjects can be identified with a higher probability to have a particular result. The person skilled in the art can determine without difficulty whether a part is statistically significant using several well-known statistical tools for assessment, for example, determination of confidence intervals, determination of p-value, cross-validation classification indices, etc. (more details in " Statistics for research" Dowdy and Wearden, John Wiley & Sons, New York, 1983). Preferred confidence intervals are at least 50%, at least 60%, at least 70%, at least 80%, at least 90% or at least 95%. P-values are, preferably, 0.1, 0.05, 0.02, 0.01 or lower.

The term "neurodegenerative disease" as used herein, includes chronic and progressive processes which are characterized by the selective and symmetric loss of neurons in motor, sensory and cognitive systems.

The term "Alzheimer's disease" or "senile dementia" or AD refers to a mental impairment associated with a specific degenerative brain disease characterized by the appearance of senile plaques, neuritic tangles and progressive neuronal loss that is clinically manifested in progressive deficiencies of memory, confusion, behavioral problems, inability to care for oneself, gradual physical deterioration and, ultimately, death. In preferred embodiments, the

Alzheimer's disease at any stage according to Braak staging:

- Stages I-II: the brain area affected by the presence of neurofibrillary tangles corresponding to the transentorhinal region of the brain
- Stages III-IV: the affected brain area also extends to areas of the limbic region such as the hippocampus
- Stages V-VI: the affected brain area also involves the neocortical region

This classification by neuropathological stages correlates with the clinical evolution of the existing disease and there is a parallel between the decline in memory with the neurofibrillary changes and the formation of neuritic plaques in the entorhinal cortex and the hippocampus (stages I to IV). Also, the isocortical presence of these changes (stages V and VI) correlates with clinically severe alterations. The transentorhinal state (I-II) corresponds to clinically silent periods of disease. The limbic state (III-IV) corresponds to a clinically incipient AD. The neocortical state corresponds to a fully developed AD.

The term "Parkinson's disease" or "idiopathic parkinsonism" or "paralysis agitans" or PD as used herein, refers to a chronic, degenerative disease that involves problems of movement control, tremor, rigidity, bradykinesia in all kinds of movements such as walking, sitting, eating, talking, etc., and well as postural instability. Symptoms of the disease are clearly associated with the selective degeneration of dopaminergic neurons in the substantia nigra. The dopaminergic deficit induces a consequent loss of striatal neurons causing a variety of cytological changes including  $\alpha$ -synuclein aggregation in so-called Lewy bodies. The "substantia nigra" is a nucleus of the basal ganglia located in the upper portions of the midbrain, under the thalamus and takes its color from the neuromelanin. In preferred embodiments, the PD is in any of the stages according to the Braak staging:

- Stage I: the affected area is the dorsal motor nucleus and/or intermediate reticular zone.
- Stage II: the affected area extends to coreuleus locus and to the nucleus raphe
- Stage III: the affected area extends to the midbrain, in particular the substantia nigra pars compacta.
- Stage IV: The affected area extends to the transentorhinal region of the anteromedial temporal mesocortex and allocortex.
- Stage V: The affected area extends to the insular cortex, the cingulate cortex and the temporal gyrus.

- Stage VI: The affected area extends to frontal and parietal area of the cortex.

The term "subject" as used herein, refers to a person, such as a human being, non-human primate (e.g., chimpanzees and other apes and monkey species), farm animals, such as birds, 5 fish, cattle, sheep, pigs, goats and horses, Pets such as dogs and cats; mammals, laboratory animals including rodents, such as mice, rats and guinea pigs. The term does not denote a particular age or sex. In a specific embodiment of the invention, the subject is a mammal. In a preferred embodiment of the invention, the subject is a human.

0 The expression "sample comprising mitochondrial DNA" as used herein refers to any sample that can be obtained from a subject in which there is genetic material from the mitochondria suitable for detecting the methylation pattern.

The term "mitochondrial DNA" or "mtDNA" as used herein, refers to the genetic material 5 located in the mitochondria of living organisms. It is a closed, circular double-stranded molecule. In humans it consists of 16,569 base pairs, containing a small number of genes, distributed between the H chain and L chain. Mitochondrial DNA encodes 37 genes: two ribosomal RNA, 22 transfer RNA and 13 proteins that participate in oxidative phosphorylation.

:0 In a specific embodiment of the invention, the sample comprising mitochondrial DNA is selected from a solid tissue biopsy or biofluid. The samples can be obtained by conventional methods known to the person skilled in the art.

In an even more specific embodiment, the biofluid is selected from peripheral blood or 25 cerebrospinal fluid.

In an even more specific embodiment, said solid tissue is brain tissue.

In a preferred embodiment of the invention, if desired to diagnose a subject or if desired to 30 determine the risk of developing PD, said sample is a brain tissue sample obtained from the substantia nigra.

If the material where it is desired to determine the methylation pattern according to the present 35 method, that is to say the mtDNA, is in a solid tissue or biofluid preferably, it proceeds a prior nucleic acid extraction from the sample using any suitable technique for this. In a preferred embodiment of the invention, the DNA fraction suitable for the implementation of the invention

is 35 total DNA. DNA extraction may be carried out using method known to the persons skilled in the art (Sambrook et al., 2001. "Molecular Cloning: A Laboratory Manual", 3rd ed, Cold Spring Harbor Laboratory Press, NY, Vol. 1-3) including, without limitation, density gradient centrifugation, extraction in two stages using aqueous phenol or chloroform with ethanol, 5 column chromatography, methods based on the ability of DNA to bind itself on glass and/or silicates, such as diatomaceous earth as preparations or crystal beds, using commercial kits, for example, "Q-Biogene fast DNA @ kits " or "QIAamp® (R) DNA Blood Mini Kit" (Qiagen, Hilden, Germany) the "G-Spin IIp" (Intron Biotechnology, Korea) or the "Bio 101 Fast Prep System" (Qbiogene®, Madrid, Spain) or the methods described in US5.057.426, US4.923.978 0 and European Patent Application EP0512767A1.

If desired, the present method can be carried out in samples in which the mitochondrial fraction has previously been isolated and subsequently the DNA thereof has been isolated. The isolation of the mitochondrial fraction can be carried out using any known method of cell 5 fractionation. Such methods comprise the previous cell disruption by techniques including physical disruption of the membranes, application of ultrasounds, pressure application or enzymatic techniques, followed by differential centrifugation by applying density gradients (such as those of Ficoll as or Percoll). They can also use commercial kits, for example, Qproteome "Mitochondrial isolation kit" (Qiagen, Hilden, Germany) or "Mitochondrial isolation 10 kit for cultured cells" (Thermo Scientific, USA). These kits are based on the same basic principle, namely cell lysis and differential centrifugation to isolate or enrich the mitochondrial fraction.

The first method of the invention involves determining the methylation pattern in a sample from 25 a subject comprising mitochondrial DNA. The term "DNA methylation" As used herein, refers to a biochemical process involving the addition of a methyl group (-CH<sub>3</sub>) to DNA cytosine nucleotides (C) or adenine (A). DNA methylation at the 5 position of cytosine has the specific effect of gene repression and has been found in all of the vertebrates examined.

30 The term "methylation pattern" as used herein refers but is not limited to the presence or absence of methylation of one or more nucleotides. In this way, said one or more nucleotides are comprised in a single nucleic acid molecule.

35 Said one or more nucleotides are capable of being methylated or not. The term "methylation status" can also be used when only considered a single nucleotide. A methylation pattern can be quantified; in the case it is considered more than one nucleic acid molecule.

The term "D-loop" or "control region" as used herein, refers to a region of non-coding mtDNA containing approximately 1100 base pairs, visible under electron microscopy, which is generated during H chain replication for the synthesis of a short segment of the heavy strand,

5 7S DNA.

The term "ND1" or " NADH dehydrogenase 1" or "ND1mt ", as used herein, refers to the gene localized in the mitochondrial genome that encodes the protein NADH dehydrogenase 1 or ND1. The human ND1 gene sequence is deposited in the GenBank database (version of 0 January 2, 2014) under the accession number NC\_012920. SEQ ID NO 1. The ND1 protein is part of the enzyme complex called complex I which is active in the mitochondria and is involved in the process of oxidative phosphorylation.

The term "CpG site" as used herein, refers to DNA regions, particularly mitochondrial DNA regions, where a cytosine nucleotide is followed by a guanine nucleotide in linear sequence of bases along its length. "CpG" is an abbreviation for "C-phosphate-G", i.e., cytosine and guanine separated by only a phosphate; phosphate binds together any two nucleosides in the DNA. The term "CpG" is used to distinguish this linear sequence of CG bases pairing of guanine and cytosine. Cytosine in the CpG dinucleotides may be methylated to form 5-methylcytosine.

The term "CHG site" as used herein, refers to DNA regions, particularly mitochondrial DNA regions, where a cytosine nucleotide and a guanine nucleotide are separated by a variable nucleotide (H) which can be adenine, cytosine or thymine. The cytosine of the CHG site can 25 be methylated to form 5-methylcytosine.

The term "CHH site" as used herein, refers to DNA regions, particularly regions of mitochondrial DNA, where a cytosine nucleotide is followed by a first and a second variable nucleotide (H) which can be adenine, cytosine or thymine. The cytosine of the CHG site can 30 be methylated to form 5- methylcytosine.

In a specific embodiment, the first method of the invention comprises determining in a sample of a subject comprising mitochondrial DNA, the methylation pattern in at least one site selected from the CpG sites of the D-loop region, selected from the sites shown in Table 1.

35 **Table 1:** List of CpG positions 16386 and 256 in the D-loop region.

CpG site	Position (bp)
<b>CpG 2</b>	16427
<b>CpG 3</b>	16449
<b>CpG 4</b>	16454
<b>CpG 5</b>	16495
<b>CpG 6</b>	16542
<b>CpG 7</b>	16565
<b>CpG 8</b>	33
<b>CpG 9</b>	61
<b>CpG 10</b>	78
<b>CpG 11</b>	80
<b>CpG 12</b>	91
<b>CpG 13</b>	96
<b>CpG 14</b>	105
<b>CpG 15</b>	120
<b>CpG 16</b>	162
<b>CpG 17</b>	170
<b>CpG 18</b>	186

The term "determination of the methylation pattern in a CpG site" as used herein, refers to the determination of the methylation status of a particular CpG site. The determination of the methylation pattern of a CpG site can be performed by multiple processes known to the person skilled in the art.

In a specific embodiment, the first method of the invention involves determining the methylation pattern of at least one CpG site in the D-loop region selected from the sites shown in Table 1. In another specific embodiment, the first method of the invention involves determining the methylation pattern in at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15 or at least 16 CpG sites selected from Table 1.

In a still more specific and preferred embodiment of the invention, the first method of the invention comprises determining the methylation pattern of all of the CpG sites in the D-loop region shown in Table 1.

In another specific embodiment, the first method of the invention comprises determining in a sample of a subject containing mitochondrial DNA, the methylation pattern in at least one site selected from CpG sites of the ND1 gene shown in Table 2.

**Table 2:** List of CpG sites in the ND1 gene between the positions 3313 and 3686.

CpG site	Position (bp)
CpG 1	3351
CpG 2	3375
CpG 3	3379
CpG 4	3406
CpG 7	3453
CpG 12	3549
CpG 13	3642

5 In another particular embodiment, the first method of the invention involves determining the methylation pattern in a CpG site selected from the ND1 gene shown in Table 2. In another specific embodiment, the first method of the invention involves determining the methylation pattern in at least 2, at least 3, at least 4, at least 5 or at least 6 CpG sites selected from Table 2.

0

In a still more particular and preferred embodiment of the invention, the first method of the invention involves determining the methylation pattern in all of the CpG sites of the ND1 gene shown in Table 2.

5 In another particular embodiment, the first method of the invention comprises determining in a sample of a subject comprising mitochondrial DNA, the methylation pattern in at least one site selected from the CHG sites in the D-loop region, shown in Table 3.

20

**Table 3:** List of the CHG sites between positions 16386 and 256 of the D-loop region

CHG site	Position (bp)
CHG 2	16426
CHG 3	16453
CHG 4	16459
CHG 5	16466
CHG 6	16479

CHG site	Position (bp)
CHG 7	16514
CHG 8	6
CHG 9	33
CHG 10	64
CHG 11	104
CHG 12	122
CHG 13	128
CHG 14	141
CHG 16	253

The term "determination of the methylation pattern in a CHG site" as used herein, refers to the determination of the methylation status of a particular CHG site.

5 The determination of the methylation pattern of a CHG site can be performed by multiple processes known to the person skilled in the art.

In a specific embodiment, the first method of the invention involves determining the methylation pattern of at least one CHG site of the D-loop region selected from the sites shown in Table 3.

0 In another specific embodiment, the first method of the invention involves determining the methylation pattern in at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13 or at least 14 or at least 15 CHG sites selected from Table 3.

15 In a still more specific and preferred embodiment of the invention, the first method of the invention involves determining the methylation pattern in all of the CHG sites of the D-loop region shown in Table 3.

In another specific embodiment, the first method of the invention comprises determining

20

in a sample of a subject comprising mitochondrial DNA, the methylation pattern in at least one site selected from CHG sites of the ND1 gene shown in Table 4.

**Table 4:** List of CHG sites in the ND1 gene between the positions 3313 and 3686.

25

CHG site	Position (bp)
CHG 1	3374
CHG 2	3435

CHG site	Position (bp)
CHG 4	3524
CHG 5	3529
CHG 6	3589
CHG 7	3641
CHG 8	3657

In another specific embodiment, the first method of the invention involves determining the methylation pattern in at least one CHG site in the ND1 gene selected from the sites shown in Table 4. In another specific embodiment, the first method of the invention comprises 5 determining the methylation pattern in at least 2, at least 3, at least 4, at least 5 or at least 6 CHG sites selected from Table 4.

In a still more specific and preferred embodiment of the invention, the first method of the invention involves determining the methylation pattern in all of the CHG sites of the ND1 gene 0 shown in Table 4.

In another specific embodiment, the first method of the invention comprises determining in a sample of a subject comprising mitochondrial DNA, the methylation pattern in at least one site selected from the CHH sites of the D-loop region, selected from the CHH sites shown in table 5. 5.

**Table 5:** List of CHH sites between positions 16386 and 256 positions of the D-loop region.

CHH site	Position (bp)
CHH 5	16419
CHH 6	16425
CHH 7	16429
CHH 8	16439
CHH 9	16442
CHH 10	16446
CHH 11	16451
CHH 12	16458
CHH 13	16465
CHH 14	16478
CHH 15	16498
CHH 16	16507
CHH 17	16511
CHH 18	16520

CHH site	Position (bp)
<b>CHH 19</b>	16527
<b>CHH 20</b>	16536
<b>CHH 21</b>	16540
<b>CHH 22</b>	16546
<b>CHH 23</b>	16549
<b>CHH 24</b>	16560
<b>CHH 25</b>	16563
<b>CHH 26</b>	4
<b>CHH 27</b>	11
<b>CHH 28</b>	15
<b>CHH 29</b>	18
<b>CHH 30</b>	26
<b>CHH 31</b>	29
<b>CHH 32</b>	39
<b>CHH 33</b>	43
<b>CHH 34</b>	48
<b>CHH 35</b>	76
<b>CHH 36</b>	86
<b>CHH 37</b>	110
<b>CHH 38</b>	113
<b>CHH 39</b>	132
<b>CHH 40</b>	140
<b>CHH 41</b>	144
<b>CHH 42</b>	147
<b>CHH 43</b>	150
<b>CHH 44</b>	164
<b>CHH 45</b>	167
<b>CHH 46</b>	190
<b>CHH 47</b>	194
<b>CHH 49</b>	198

The term "determination of a methylation pattern in a CHH site", as used herein, refers to determining the methylation status of a particular CHH site. The determination of the methylation pattern of a CHG site you can be performed by multiple processes known to the person skilled in the art.

In another specific embodiment, the first method of the invention comprises determining the methylation pattern in at least one CHH site of the D-loop region selected from those shown in Table 5. In another specific embodiment, the first method of the invention involves determining

the methylation pattern in at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28 at least 29, at least 30, at least 31, at least 32, at least 33, at least 34, at least 35, at least 36, at least 37, at least 38, at least 39, at least 40, at least 41, at least 42, or at least 43 CHH sites selected from Table 5.

In a still more specific and preferred embodiment of the invention, the first method of the invention involves determining the methylation pattern in all of the CHH sites in the D-loop region shown on table 5.

In preferred methods of embodiment, the first method of the invention includes:

- (i) determination of the pattern of methylation in all of the CpG sites of the D-loop region shown in Table 1 and all CpG sites of the ND1 gene shown in Table 2
- (ii) determination of the methylation pattern in all of the CpG sites in the D-loop region shown in Table 1 and in all of the CHG sites of the D-loop region shown in Table 3
- (iii) determination of the methylation pattern in all of the CpG sites of the D-loop region shown in Table 1 and all CHG sites of the ND1 gene shown in Table 4
- (iv) determination of the methylation pattern in all CpG sites in the D-loop region shown in Table 1 and in all CHH sites in the D-loop region shown in Table 5
- (v) determination of the methylation pattern in all of the CpG sites in the ND1 gene shown in Table 2 and in all of the CHG sites in the D-loop region shown on Table 3
- (vi) determination of the methylation pattern in all of the CpG sites in the NDI gene shown in Table 2 and in all of the CHG sites in the NDI gene shown in Table 4,
- (vii) determination of the methylation pattern in all of the CpG sites in the ND1 gene shown in Table 2 and in all of the CHH sites of the D-loop region shown in Table 5
- (viii) determination of the methylation pattern in all of the CHG sites in the D-loop region shown in Table 3 and all CHG sites of the ND1 gene shown in Table 4
- (ix) determination of the methylation pattern in all CHG sites in the D-loop region shown in Table 3 and/or
- (x) determination of the methylation pattern in all of the CHG sites in the ND1 gene shown in Table 4 and in all of the CHH sites in the D-loop region in Table 5.

In some embodiments, the determination of the methylation pattern in at least one CpG site, at least one CHG site and/or at least one CHH site according to the first method of the invention is carried out in a whole blood sample, in which case the determination can be made directly.

In other embodiments, the sample that contains mitochondrial DNA, preferably a sample of total DNA, is extracted from cells present in a biological fluid (e.g., whole blood, cerebrospinal fluid) as an initial stage and in such cases, the total nucleic acid extracted from these samples represent the suitable work material for the later analysis. The Isolation of total DNA or 5 mitochondrial DNA can be performed by conventional methods known to the person skilled in the art (cited supra). After isolating and amplifying (if necessary) the nucleic acid that contains mitochondrial DNA, the methylation pattern of one or more CpG sites, one or more CHG sites and/or one or more CHH sites is determined. The person skilled in the art will easily recognize that the analysis of the methylation pattern present in one or several of the CpG, CHG and/or 0 CHH sites described herein in the mitochondrial DNA of a subject, can be carried out by any method or technique capable of measuring the methylation pattern present in such sites.

In another specific embodiment, the first method of the invention involves determining the methylation pattern in said CpG, CHG and/or CHH sites through a technique selected from the 5 group consisting of Methylation Specific PCR (MSP), a method based on enrichment. (e.g. MeDIP, MBD-seq and MethylCap), bisulfite sequencing and through a bisulphite based method (e.g. RRBs, Infinium, GoldenGate, Cobra, MSP, MethyLight) and a method by restriction of digestion (e.g., MRE-seq, or HELP trial), pyrosequencing, or differential conversion, differential restriction, differential weight of the methylated DNA CpG, CHG and/or 0 CHH sites.

In a specific and preferred embodiment of the invention the methylation pattern of one or more CpG, CHG and/or CHH sites in the D-loop region and/or one or more CpG and/or CHG sites in the ND1 gene is determined through pyrosequencing. Briefly, this technique is based on the 25 principle of sequencing by synthesis and detection of released pyrophosphate (PPi) during DNA synthesis. This technique employs a series of four enzymes to detect nucleic acid sequences during the synthesis process; DNA polymerase, ATP sulfurylase, luciferase, apyrase and adenosine 5' fosfatosulphate (APS) and luciferin used as substrates.

30 To determine the methylation pattern in mitochondrial DNA, it is necessary to chemically treat said sample so that all cytosine unmethylated bases are modified at uracil bases, or another base which differs from cytosine in terms of base pairing behavior, while the bases of 5 methylcytosine remain unchanged. The term "modify" as used herein means the conversion of an unmethylated cytosine to another nucleotide that will distinguish the unmethylated 35 cytosine from the methylated cytosine. The conversion of unmethylated cytosine bases, but not methylated, in the sample containing mitochondrial DNA is carried out with a conversion

agent. The term "conversion agent" or "conversion reagent" as used herein, refers to a reagent capable of converting an unmethylated cytosine to uracil or another base that is differentially detectable to cytosine in terms of hybridization properties. The conversion agent is preferably a bisulfate such as bisulfites or hydrogen sulfite. However, other agents that similarly modify unmethylated cytosine, but not methylated cytosine can also be used in this method of the invention, such as hydrogen sulfite. The reaction is performed according to standard procedures (Frommer et al., 1992, Proc. Natl. Acad. Sci. USA 89: 1827-1831; Olek, 1996, Nucleic Acids Res. 24: 5064-6; EP 1394172). It is also possible to carry out the conversion enzymatically, e.g. using cytidine deaminases specific methylation.

In a preferred embodiment of the first method of the invention, the sample containing mitochondrial DNA has been treated with a reagent capable of converting an unmethylated cytosine to uracil or another base that is detectably different from cytosine in terms of hybridization properties. In a more preferable embodiment, the sample comprising mitochondrial DNA is treated with bisulfite using an appropriate commercial kit, for example "EZ Methylation Kit" (Zymo Research, Ecogen; Barcelona, Spain).

Once the sample containing mitochondrial DNA has been treated with a bisulfite, the D-loop region and/or the ND1 gene containing one or more CpG, CHG and/or CHH sites shown in Tables 1 to 5 can be amplified using primers that distinguish unmethylated sequence (in which the cytosine of the CpG site is converted into uracil) from the methylated sequence (in which the cytosine in the CpG site remains as cytosine). Many amplification methods rely on an enzymatic chain reaction such as, for example, a polymerase chain reaction (PCR), ligase chain reaction (LCR), ligase chain reaction Polymerase 35, Gap- LCR, repair chain reaction, 3SR and NASBA. Furthermore, there is strand displacement amplification (SDA), transcription mediated amplification (TMA), and Q $\beta$ -amplification, etc, this being merely an illustrative list. Methods of nucleic acid amplification are described in Sambrook et al., 2001 (cited supra). Other amplification methods include PCR method specific methylation (MSP) described in US document 5,786,146 which combines bisulfite treatment and allele-specific PCR (see, e.g., US Patent 5,137,806, US 5,595,890, US 5,639,611). Uracil is recognized as a thymine by Taq polymerase and therefore, according to PCR, the resultant product contains cytosine only at the position where 5-methylcytosine DNA exist in the starting template.

In a preferred embodiment of the invention, once the sample comprising mitochondrial DNA, preferably a sample of total DNA, has been treated with bisulfite, the region containing one or more CpG, CHG and/or CHH site(s) can be amplified using primers that are not specific to the

methylated sequence. For example, the preferred sequence of the primers not corresponding to a nucleotide sequence comprising a CpG dinucleotide.

Amplification products are detected according to standard procedures in prior art. The

5 amplified nucleic acid can be determined through methods known to person skilled in the art and are described in e.g., Sambrook et al., 2001 (cited et supra). There may also be additional purification steps before the target nucleic acid is detected, for example a precipitation step. The detection methods may include but may not limited to binding or interleaving of specific dyes such as ethidium bromide which intercalate into double stranded DNA and changes its 0 fluorescence thereafter. The purified nucleic acids may also be separated by electrophoretic methods optionally after a restriction of digestion and be displayed later. There are also probe-based assays which exploit the oligonucleotide hybridization to specific sequences and subsequent detection of the hybrid. It is also possible to sequence the target nucleic acid after further steps known to the person skilled in the art. Other methods use various nucleic acid 5 sequences with a silicon chip to which specific probes are bound and produce a signal when a complementary sequence binds.

In a preferred embodiment of the invention, after amplification of the region of interest where it is desired to determine the methylation pattern (eg in the D-loop region or in the ND1 gene)

:0 pyrosequencing is used to determine in said sequence CpG, CHG and/or modified CHH sites after treatment with bisulphite. The reason for cytosine/thymine in each of the sites can be quantitatively determined based on the amount of cytosine and thymine incorporated during the extension step of the sequence.

25 Alternatively, the methylation pattern of at least one CpG, CHG and/or CHH sites in the D-loop region or in at least one CpG and/or CHG site of the ND1 gene of the site can be confirmed by restriction enzyme digestion and Southern blot analysis. Examples of endonucleases 30 methylation sensitive restriction that may be used include *SmaI*, *SacII*, *EagI*, *MspI*, *HpaII*, *BstUI* Y *BssHII*, for example.

35 The term "hypermethylation" as used herein, refers to an altered methylation pattern where one or more nucleotides, preferably cytosines from CpG, CHG and/or CHH sites are methylated compared to a reference sample. Said reference sample is preferably a sample that contains mitochondrial DNA obtained from a subject not suffering from a neurodegenerative disease selected from AD or PD. In particular, the term refers to a number larger than 5-methylcytosines in one or more CpG sites in the D-loop region shown in Table 1,

in one or more CpG sites in the ND1 gene shown in Table 2, in one or more CHG sites in the D-loop region shown in Table 3, in one or more CHG sites in the ND1 gene shown in Table 4 and/or one or more CHH sites in the D-loop region shown in Table 5, in a sequence of mitochondrial DNA when compared with the relative amount of 5-methylcytosines present in  
5 said one or more sites in a reference sample.

The term "hypomethylation" as used herein, refers to an altered methylation pattern where one or more nucleotides, preferably cytosines in the CpG, CHG and/or CHH sites, are unmethylated compared to a reference sample. The term "reference sample" refers to a  
0 sample containing mitochondrial DNA obtained from a subject not suffering from a neurodegenerative disease selected from AD or PD. In particular, said term refers to a small number of 5-methylcytosines in one or more CpG sites in the D-loop region shown in Table 1, in one or more CpG sites of the ND1 gene shown in Table 2, in one or more sites CHG sites in the D-loop region shown in Table 3, in one or more CHG sites in the ND1 gene shown in  
5 Table 4 and/or one or more CHH sites in the D-loop region shown in Table 5 in a sequence of mitochondrial DNA as compared to the relative amount of 5-methylcytosines present in said one or more CpG sites, one or more CHG sites and/or one or more CHH sites in a reference sample.

:0 In a preferred embodiment of the invention, said reference sample containing mitochondrial DNA is selected from tissue samples, or biofluids, preferably blood samples or cerebrospinal fluid of subjects. In a preferred embodiment, said reference sample is total DNA. Methods to obtain these samples as well as methods of isolating total DNA or mitochondrial DNA of a sample have been detailed above. In a still more preferred embodiment, the reference sample  
25 is a sample containing mitochondrial DNA from age-matched subjects.

In this first method, the invention provides some specific CpG, CHH and CHG sites that are related to the diagnosis or risk of developing a neurodegenerative disease selected from Alzheimer's disease and Parkinson's disease. Therefore:

30

- hypermethylation in at least one of the CpG sites in the D-loop region shown in Table 1,
- hypermethylation in at least one of the CHG sites in the D-loop region shown in Table 3,
- 35 • hypermethylation in at least one of the CHH sites in the D-loop region shown in Table 5,

- hypomethylation in at least one of the CpG sites in the ND1 gene shown in Table 2, and/or
- hypomethylation in at least one of the CHG sites in the ND1 gene shown in Table 4, is indicative that the subject suffers from Alzheimer's disease or that the subject is at a high risk of developing Alzheimer's disease; or
  - hypomethylation in at least one of the CpG sites in the D-loop region shown in Table 1,
  - hypomethylation in at least one of the CHG sites in the D-loop region shown in Table 3, and/or
  - hypomethylation in at least one of the CHH sites in the D-loop region shown in Table 5 is indicative that the subject suffers from Parkinson's or that the subject is at a high risk of developing Parkinson's disease.

In a specific embodiment, the first method of the invention involves determining the methylation pattern of all the CpG sites, all CHG sites and all CHH sites in the D-loop region shown in Tables 1, 3 and 5, and the methylation pattern of all the CpG sites all of all CHG sites of the ND1 gene shown in Tables 2 and 4.

The authors of the present invention have found that the degree of methylation in the CpG, CHG and CHH sites in the D-loop region is greater in subjects suffering from Alzheimer's in stages I-II than in subjects suffering from the disease in stages III-IV.

In a specific embodiment, if the methylation pattern with respect to the reference pattern is observed in a sample containing mitochondrial DNA from a subject diagnosed with Alzheimer's disease stage I-II, then hypomethylation in at least one of the CpG sites in the D-loop region shown in Table 1 or hypomethylation in one of said CHG sites in the D-loop region shown in Table 3, indicates that the subject suffers from Alzheimer's disease stage III-IV.

### **Second method of the invention**

In a second aspect, the invention relates to an *in vitro* method to select a subject to be submitted to preventive treatment of a neurodegenerative disease selected from Alzheimer's disease and Parkinson's disease in a subject (hereinafter, second method of the invention) that involves determining in a sample of said subject containing mitochondrial DNA the methylation pattern in the D-loop region, and/or in the ND1 gene, where the methylation pattern is determined in at least one site selected from the group formed by:

- i. the CpG sites in the D-loop region shown in Table 1,
- ii. the CpG sites in the ND1 gene shown in Table 2;
- iii. the CHG sites in the D-loop region shown in Table 3,
- iv. the CHG sites in the ND1 gene shown in Table 4, and/or
- v. the CHH sites in the D-loop region shown in Table 5

wherein a hypermethylation in at least one of said CpG sites in the D-loop region, a hypermethylation in at least one of said CHG sites in the D-loop region, a hypermethylation in at least one of said CHH sites in the D-loop region, a hypomethylation in at least one of said CpG sites in the ND1 gene and/or a hypomethylation in at least one of said CHG sites in the ND1 gene is indicative that the subject is a candidate to receive treatment aimed at preventing Alzheimer's disease or

wherein a hypomethylation in at least one of said CpG sites in the D-loop region, a hypomethylation in at least one of said CHG sites in the D-loop region and/or a hypomethylation in at least one of the said CHH sites in the D-loop region it is indicative that the subject is a candidate to receive treatment aimed at preventing Parkinson's disease.

The term "preventive treatment", as used herein, refers to the prevention or conjunction of prophylactic measures to prevent disease to prevent or delay the onset of symptoms thereof as well as reducing or alleviating clinical symptoms thereof. In particular, the term refers to the prevention or set of measures to prevent the occurrence, to delay or to relieve the clinical symptoms associated with a neurodegenerative disease selected from Alzheimer's and Parkinson's disease. Desired clinical outcomes associated with the administration of the treatment to a subject include but are not limited to, stabilization of the pathological stage of the disease, delay in the progression of the disease or improvement in the physiological state of the subject.

Suitable preventative treatments aimed at preventing or delaying the onset of the symptoms of Alzheimer's disease include but are not limited to, cholinesterase inhibitors such as donepezil hydrochloride (Aricept), rivastigmine (Exelon) and galantamine (Reminyl) or antagonists N-methyl-D-aspartate (NMDA). Treatments aimed at preventing or delaying the onset of the symptoms of Parkinson's disease include but are not limited to, L-dopa, inhibitors of catechol-o-methyl transferase (COMT) such as tolcapone (Tasmar) and entacapone (Comtan), monoamine oxidase B (MAOB) such as selegiline (Eldepryl) and rasagiline (Azilect) and dopamine agonists such as pramipexole, rotigotine and ropinirole.

The term "to select" as used herein, refers to the action of choosing a subject for submission to preventive treatment of a neurodegenerative disease selected from Alzheimer's disease and Parkinson's disease.

5

The terms "subject", "neurodegenerative disease", "Alzheimer's disease", "Parkinson's disease", "sample", "mitochondrial DNA", "D-loop region", "ND1 gene", "CpG site", "CHG site", CHH site", "methylation pattern", "hypermethylation" and "hypomethylation" have been described in detail in the context of the first method of the invention and used with the same meaning in the second method of the invention.

0

In a specific embodiment of the second method of the invention, the sample comprising mitochondrial DNA is selected from a biopsy of a solid tissue or biofluid. The samples can be obtained by conventional methods known to the person skilled in the art.

5

In an even more specific embodiment, the biofluid is selected from peripheral blood or cerebrospinal fluid.

10

In an even more specific embodiment, said solid tissue is brain tissue. In a preferred embodiment of the invention, if desired to select a subject to be submitted to preventive treatment of Parkinson's disease, said sample is a brain tissue sample obtained from the substantia nigra.

25

In a specific embodiment, the second method of the invention involves determining in a sample of a subject containing mitochondrial DNA, the methylation pattern in at least one site selected from CpG sites in the D-loop region, shown in Table 1.

30

In a specific embodiment, the second method of the invention involves the determination of the methylation pattern in at least one CpG site in the D-loop region selected from those shown in Table 1. In another specific embodiment, the first method of the invention involves determining the methylation pattern in at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15 or at least 16 CpG sites selected from Table 1.

35

In a still more specific and preferred embodiment of the invention, the second method of the invention involves determining the methylation pattern in all of the CpG sites in the D-loop

region shown in Table 1.

In another specific embodiment, the second method of the invention involves determining in a sample of a subject comprising mitochondrial DNA, the methylation pattern in at least one site selected from the CpG sites in the ND1 gene, shown in Table 2.

In another specific embodiment, the second method of the invention involves determining the methylation pattern in a CpG site in the ND1 gene selected from the sites shown in Table 2. In another specific embodiment, the second method of the invention involves determining the methylation pattern in at least 2, at least 3, at least 4, at least 5 or at least 6 CpG sites selected from Table 2.

In a still more specific and preferred embodiment of the invention, the second method of the invention involves determining the methylation pattern of all of the CpG sites in the ND1 gene shown in Table 2.

In another particular embodiment, the second method of the invention involves determining in a sample of a subject containing mitochondrial DNA, the methylation pattern in at least one site selected from the CHG sites in the D-loop region, shown in Table 3.

In a particular embodiment, the second method of the invention involves determining the methylation pattern in at least one CHG site in the D-loop region selected from the sites shown in Table 3. In another specific embodiment, the second method of the invention involves determining the methylation pattern in at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13 or at least 14 or at least 15 CHG sites selected from Table 3.

In a still more specific and preferred embodiment of the invention, the second method of the invention involves determining the methylation pattern of all the CHG sites in the D-loop region shown in Table 3.

In another specific embodiment, the second method of the invention involves the determination in a sample of a subject that contains mitochondrial DNA, the methylation pattern in at least one site selected from the CHG sites in ND1 gene, shown in Table 4.

In another specific embodiment, the second method of the invention involves determining the

methylation pattern in at least one CHG site in the ND1 gene selected from the sites shown in Table 4. In another specific embodiment, the second method of the invention involves determining the methylation pattern in at least 2, at least 3, at least 4, at least 5 or at least 6 CpG sites selected from Table 4.

5

In a still more particular and preferred embodiment of the invention, the second method of the invention involves determining the methylation pattern in all of the CHG sites of the ND1 gene shown in Table 4.

0

In another specific embodiment, the second method of the invention involves determining in a sample of a subject containing mitochondrial DNA, the methylation pattern in at least one site selected from the CHH sites in the D-loop region, shown in Table 5.

5

In another specific embodiment, the second method of the invention involves determining the methylation pattern in at least one CHH site in the D-loop region selected from those shown in Table 5. In another specific embodiment, the second method of the invention involves determining the methylation pattern in at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, at least 30, at least 31, at least 32, at least 33, at least 34, at least 35, at least 36, at least 37, at least 38, at least 39 at least 40, at least 41, at least 42 or at least 43 CHH sites selected from Table 5.

10

In a still more specific and preferred embodiment of the invention, the second method of the invention involves determining the methylation pattern in all of the CHH sites in the D-loop region shown in Table 5.

25

In preferred methods of embodiment, the second method of the invention includes:

30

- (i) determination of the methylation pattern in all of the CpG sites in the D-loop region shown in Table 1 and in all of the CpG sites in the ND1 gene shown in Table 2
- (ii) determination of the methylation pattern in all of the CpG sites in the D-loop region shown in Table 1 and in all of the CHG sites in the D-loop region shown in Table 3
- (iii) determination of the methylation pattern in all of the CpG sites in the D-loop region shown in Table 1 and in all of the CHG sites in the ND1 gene shown in Table 4
- (iv) determination of the methylation pattern in all CpG sites in the D-loop region shown

in Table 1 and in all CHH sites in the D-loop region shown in Table 5  
(v) determination of the methylation pattern in all of the CpG sites in the ND1 gene shown in Table 2 and in all CHG sites in the D-loop region shown in Table 3  
(vi) determination of the methylation pattern in all of the CpG sites in the ND1 gene shown in Table 2 and in all of the CHG sites in the ND1 gene shown in Table 4  
5 (vii) determination of the methylation pattern in all CpG sites in the ND1 gene shown in Table 2 and in all the CHH sites in the D-loop region shown in Table 5  
(viii) determination of the methylation pattern in all of the CHG sites in the D-loop region shown in Table 3 and in all of the CHG sites in the ND1 gene shown in Table 4  
0 (ix) determination of the methylation pattern in all of the CHG sites in the D-loop region shown in Table 3 and/or  
(x) determination of the methylation pattern in all of the CHG sites in the ND1 gene shown in Table 4 and in all of the CHH sites in the D-loop region shown in Table 5.

5 Suitable methods for determining the methylation pattern in a sample of a subject containing mitochondrial DNA have been described in detail in the context of the first method of the invention.

In a specific embodiment, the second method of the invention involves determining the 10 methylation pattern in said CpG, CHG and/or CHH sites through a technique selected from the group consisting of Methylation-Specific PCR -based enrichment method (e.g. MeDIP, MBD-seq and MethylCap), sequencing bisulfite and based on bisulfite method (e.g. RRBs, Infinium, GoldenGate, Cobra, MSP, MethylLight) and a method for restriction digestion (e.g., MRE-seq, or HELP trial), pyrosequencing, assay chIP-on-chip, or differential conversion, differential 25 restriction, differential weight of the CpG, CHG and/or CHH sites methylated DNA.

In a specific and preferred embodiment of the invention the methylation pattern of one or more CpG, CHG and/or CHH sites in the D-loop region and/or one or more CpG or CHH sites in the ND1 gene, according to the second method of the invention is determined by pyrosequencing.

30 According to the second method of the invention:

- hypermethylation in at least one of the CpG sites in the D-loop region shown in Table 1,
- hypermethylation in at least one of CHG sites in the D-loop region shown in Table 3,
- 35 • hypermethylation in at least one of the CHH sites in the D-loop region shown in Table 5

- hypomethylation in at least one of the CpG sites in the ND1 gene shown in Table 2
- and/or hypomethylation in at least one of the CHG sites in the ND1 gene shown in Table 4, is indicative that the subject is a candidate to receive treatment aimed at preventing Alzheimer's disease; or

5           • hypomethylation in at least one of the CpG sites in the D-loop region shown in Table 1

- hypomethylation in at least one of the CHG sites in the D-loop region shown in Table 3,
- and/or hypomethylation in at least one of the CHH sites in the D-loop region shown in Table 5

0           is indicative that the subject is a candidate to receive treatment aimed at preventing Parkinson's disease.

In a specific embodiment, the second method of the invention involves determining the methylation pattern of all the CpG sites, all CHG sites and all of the CHH sites in the D-loop 5 region shown in Tables 1, 3 and 5, and the methylation pattern of all the CpG sites and all of the CHG sites in the ND1 gene shown in Tables 2 and 4.

### **Third method of the invention**

:0           In the third feature, the invention relates to an *in vitro* method to monitor the progression of a neurodegenerative disease selected from Alzheimer's or Parkinson's disease in a subject (hereinafter, third method of the invention) Involving:

25           (a) Determination in a sample from said subject containing mitochondrial DNA, the methylation pattern in the D-loop region, and/or in the ND1 gene, where the methylation pattern is determined at least one site selected from the group formed by:

                 (i) the CpG sites in the D-loop region shown in Table 1,

                 (ii) the CpG sites in the ND1 gene shown in Table 2;

                 (iii) the CHG sites in the D-loop region shown in Table 3,

30           (iv) the CHG sites in the ND1 gene shown in Table 4, and/or

                 (v) the CHH sites in the D-loop region shown in Table 5 and

                 (b) comparing the methylation pattern determined in step a) with said methylation pattern obtained in an earlier stage of the disease, in the case of hypermethylation in at least one of said CpG sites in the D-loop region, hypermethylation in at least one of said CHG sites in the D-loop region, hypermethylation in at least one of the said CHH 35 sites in the D-loop region, hypomethylation in at least one of said CpG sites in the ND1

5 gene and/or hypomethylation in at least one of said CHG sites in the ND1 gene relative to said methylation pattern determined at an earlier stage of disease, is indicative of the advance of Alzheimer's disease; and where there is hypomethylation in at least one of said CpG sites in the D-loop region, hypomethylation in at least one of said CHG sites in the D-loop region and/or hypomethylation in at least one of said CHH sites in the D-loop region with respect to said methylation pattern determined at an earlier stage of the disease is indicative of the advance of Parkinson's disease.

0 The term "monitor the progression" which is equivalent to "determine the prognosis" refers to 0 determining the progression of a disease in a subject diagnosed with said disease. Particularly, the term refers to determining the progression of a neurodegenerative disease selected from Alzheimer's disease and Parkinson's disease in a subject diagnosed with said disease. As the person skilled in the art knows, there are several suitable parameters to determine the evolution of a disease in a subject, for example, the evolution of a neurodegenerative disease 5 selected from AD and PD, can be determined by, determination of overall survival.

In a specific embodiment, the subject under study has been diagnosed with AD stage I-II.

10 In another specific embodiment, the subject under study has been diagnosed with PD in stages III-V.

The term "overall survival", as used herein, refers to the percentage of patients who survive from the time of diagnosis or treatment of a neurodegenerative disease selected from Alzheimer's and Parkinson's disease, after a defined period of time.

25 The terms "subject", "neurodegenerative disease", "Alzheimer's disease", "Parkinson's disease", "sample", "mitochondrial DNA", "D-loop region", "ND1 gene", "CpG site", "CHG site", "CHH site", "methylation pattern", "hypermethylation" and "hypomethylation" have been described in detail in the context of the first method of the invention and are used with the 30 same meaning in the third method of the invention.

According to the third method of the invention, the first step of determining the prognosis of a neurodegenerative disease selected from AD and PD involves determining in a sample containing mitochondrial DNA from a subject diagnosed with said neurodegenerative disease, 35 the methylation pattern in the D-loop region and/or in the ND1 gene in at least one site selected from sites shown in Tables 1 to 5.

In a second step, the third method of the invention involves comparing the methylation pattern obtained in said first step with said methylation pattern obtained in an earlier stage of disease. Therefore, the third method of the invention involves determining the methylation pattern in a sample containing mitochondrial DNA (first sample) from a subject diagnosed with said neurodegenerative disease, the methylation pattern in the D-loop region and/or in the ND1 gene in at least one site selected from the CpG, CHG and/or CHH sites shown in Tables 1 to 5 and, after a suitable period of time, determining in a sample containing mitochondrial DNA (second sample) of said subject diagnosed with said neurodegenerative disease, the methylation pattern at these sites. Said second sample can be obtained a period of one month, two months, three months, four months, five months, six months, one year, two years, three years, four years, five years, ten years or more after obtaining the first sample.

In a particular embodiment, said first sample is obtained from a subject who is not receiving any proper treatment for said neurodegenerative disease selected from AD and PD and said second sample is obtained after a period of time where treatment of the disease is taking place. In another specific embodiment, said first sample is obtained at the beginning of appropriate treatment for said neurodegenerative disease and the second sample is obtained at one or more points during the course of the treatment.

According to the third method of the invention:

- hypermethylation in at least one of the CpG sites in the D-loop region shown in Table 1,
- hypermethylation in at least one of the CHG sites in the D-loop region shown in Table 3,
- hypermethylation in at least one of the CHH sites in the D-loop region shown in Table 5,
- hypomethylation in at least one of the CpG sites in the ND1 gene shown in Table 2,
- and/or hypomethylation in at least one of the CHG sites in the ND1 gene shown in Table 4, with respect to said methylation pattern determined at an earlier stage of disease is indicative of the advance of Alzheimer's disease; or
- hypomethylation in at least one of the CpG sites in the D-loop region shown in Table 1,
- hypomethylation in at least, one of the CHG sites in the D-loop region shown in Table 3,
- and/or hypomethylation in at least one of the CHH sites in the D-loop region shown in

Table 5,

with respect to said methylation pattern determined at an earlier stage of disease is indicative of the advance of Parkinson's disease.

5 The expression "advance of Alzheimer's disease" as used herein, refers to the subject being at a more evolved stage of the disease with respect to the stage where the subject was diagnosed. That is, if the subject was classified as stage I-II of the AD (according to the brain impairment and/or symptoms or clinical manifestations of the disease present in that subject) it is considered that the subject is at a more advanced stage of the disease if the subject goes  
0 from being classified at stage III-IV or a stage V-VI or if the subject goes from being classified as stage II I-IV to being classified as stage V-VI stage of AD.

The expression "advancement of Parkinson's disease" as used herein, refers to the subject being at a more evolved stage of the disease with respect to the stage where the subject was  
5 diagnosed. That is, if the subject was classified as stage I-II of PD (according to brain impairment and/or symptoms or clinical manifestations of disease present in that subject) it is considered that the subject is in a more advanced stage of the disease if the subject goes to being classified as stage III-IV or at stage V-VI or if the subject goes from being classified as stage III-IV to being classified as stage V-VI of PD.

:0  
In a specific embodiment of the third method of the invention, the sample comprising mitochondrial DNA is selected from a biopsy of a solid tissue or biofluid. The samples can be obtained by conventional methods known to the persons skilled in the art  
25 In an even more specific embodiment, the biofluid is selected from peripheral blood or cerebrospinal fluid.

In an even more specific embodiment, said solid tissue is brain tissue.

30 In a preferred embodiment of the invention, if the progression of Parkinson's disease is monitored, said brain tissue sample is obtained from the substantia nigra.

In a particular embodiment, the third method of the invention involves determining in a sample of a subject containing mitochondrial DNA, the methylation pattern in at least one site selected  
35 from the CpG sites in the D-loop region, shown in table 1.

In a specific embodiment, the third method of the invention involves determining the methylation pattern of at least one CpG site in the D-loop region selected from the sites shown in Table 1. In another specific embodiment, the first method of the invention involves determining the methylation pattern in at least 2, at least 3, at least 4, at least 5, at least 6, at 5 least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15 or at least 16 CpG sites selected from Table 1.

In a still more specific and preferred embodiment of the invention, the third method of the invention comprises determining the methylation pattern of all of the CpG sites in the D-loop 0 region shown in Table 1.

In another specific embodiment, the third method of the invention comprises determining in a sample of a subject containing mitochondrial DNA, the methylation pattern in at least one site selected from the CpG sites in the ND1 gene shown in Table 2.

In another specific embodiment, the third method of the invention involves determining the methylation pattern in a CpG site in the ND1 gene selected from the sites shown in Table 2. In another specific embodiment, the third method of the invention involves determining the methylation pattern in at least 2, at least 3, at least 4, at least 5 or at least 6 CpG sites selected 0 from Table 2.

In a still more specific and preferred embodiment of the invention, the third method of the invention involves determining the methylation pattern in all of the CpG sites in the ND1 gene shown in Table 2.

In a specific embodiment, the third method of the invention involves determining in a sample 25 of a subject containing mitochondrial DNA, the methylation pattern in at least one site selected from the CHG sites in the D-loop region, shown in Table 3.

30 In a specific embodiment, the third method of the invention involves determining the methylation pattern in at least one CHG site in the D-loop region selected from the sites shown in Table 3. In another specific embodiment, the third method of the invention involves determining the methylation pattern in at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13 or at least 14 or 15 35 CHG sites selected from Table 3.

In a still more specific and preferred embodiment of the invention, the third method of the invention involves determining the methylation pattern in all of the CHG sites in the D-loop region shown in Table 3.

5 In another specific embodiment, the third method of the invention involves determining in a sample of a subject containing mitochondrial DNA, the methylation pattern in at least one site selected from the CHG sites in the ND1 gene, shown in Table 4.

0 In another specific embodiment, the third method of the invention involves determining the 0 methylation pattern in at least one CHG site in the ND1 gene selected from the sites shown in Table 4. In another specific embodiment, the third method of the invention involves determining the methylation pattern in at least 2, at least 3, at least 4, at least 5, or at least 6 CpG sites selected from Table 4.

5 In a still more specific and preferred embodiment of the invention, the third method of the invention involves determining the methylation pattern in all of the CHG sites of the ND1 gene shown in Table 4.

10 In another specific embodiment, the third method of the invention involves determining in a 10 sample of a subject containing mitochondrial DNA, the methylation pattern in at least one site selected from the CHH sites in the D-loop region, shown in Table 5.

25 In another specific embodiment, the third method of the invention involves determining the methylation pattern in at least one CHH site in the D-loop region selected from the sites shown in Table 5. In another specific embodiment, the third method of the invention involves determining the methylation pattern in at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 30, at least 31, at least 32, at least 33, at least 34, at least 35, at least 36, at least 37, at least 38, at least 39, at least 40, at least 41, at least 42, or at least 43 CHH sites selected from Table 5.

35 In a still more specific and preferred embodiment of the invention, the third method of the invention involves determining the methylation pattern in all of the CHH sites in the D-loop region shown in Table 5.

In preferred methods of embodiment, the third method of the invention includes:

- i. determination of the methylation pattern in all of the CpG sites in the D-loop region shown in Table 1 and in all of the CpG sites in the ND1 gene shown in Table 2
- ii. determination of the methylation pattern in all of the CpG sites in the D-loop region shown in Table 1 and in all CHG sites of the D-loop region shown in Table 3,
- 5 iii. determination of the methylation pattern in all of the CpG sites in the D-loop region shown in Table 1 and in all of the CHG sites in the ND1 gene shown in Table 4,
- iv. determination of the methylation pattern in all CpG sites in the D-loop region shown in Table 1 and in all CHH sites in the D-loop region shown in Table 5,
- 0 v. determination of the methylation pattern in all CpG sites of the ND1 gene shown in Table 2 and in all the CHG sites in the D-loop region shown in Table 3,
- vi. determination of the methylation pattern in all of the CpG sites in the ND1 gene shown in Table 2 and in all of the CHG sites in the ND1 gene shown in Table 4,
- 7 vii. determination of the methylation pattern in all CpG sites in the ND1 gene shown in Table 2 and in all the CHH sites in the D-loop region shown in Table 5.

:0

viii. determination of the methylation pattern in all of the CHG sites in the D-loop region shown in Table 3 and all of the CHG sites in the ND1 gene shown in Table 4 and/or

ix. determination of the methylation pattern in all of the CHG sites in the D-loop region shown in Table 3

25 x. determination of the methylation pattern in all CHG sites in the ND1 gene shown in Table 4 and in all the CHH sites in the D-loop region shown in Table 5.

30 Suitable methods to determine the methylation pattern in a sample of a subject containing mitochondrial DNA methods have been described in detail in the context of the first method of the invention.

35 In another specific embodiment, the third method of the invention involves determining the methylation pattern in said CpG, CHG and/or CHH sites through a technique selected from the group consisting of Methylation Specific PCR (MSP), a method based on enrichment. (e.g. MeDIP, MBD-seq and MethylCap), bisulfite sequencing and through a bisulphite based

method (eg. RRBs, Infinium, GoldenGate, Cobra, MSP, MethyLight) and a method by restriction of digestion (e.g., MRE-seq, or HELP trial), pyrosequencing, or differential conversion, differential restriction, differential weight of the methylated DNA CpG, CHG and/or CHH sites.

5

In a specific and preferred embodiment of the invention the methylation pattern of one or more CpG, CHG and/or CHH sites in the D-loop region and/or one or more CpG or CHH sites in the ND1 gene, according to the third method of the invention, is determined by pyrosequencing.

0

In another specific embodiment, the third method of the invention involves determining the methylation pattern of all CpG sites, all CHG sites and all CHH sites in the D-loop region shown in Tables 1, 3 and 5, and the methylation pattern of all CpG sites and all of all the CHG sites in the ND1 gene shown in Tables 2 and 4.

5

#### Fourth method of the invention

The authors of the present invention have discovered a single nucleotide polymorphism (SNP) in the D-loop region of mitochondrial DNA that is statistically associated with the development of Alzheimer's disease if the SNP is in at least 60% of the mtDNA molecules of a subject.

10

Therefore, in the fourth feature, the invention relates to an *in vitro* method to diagnose or determine the risk of developing Alzheimer's disease in a subject (hereinafter, the fourth method of the invention) that involves determining in a sample that contains mitochondrial DNA of said subject the nucleotide at the polymorphic position 16519 according to the sequence defined under accession number NC\_012920 (SEQ ID NO: 1) in the NCBI database, where detection of the nucleotide C at said polymorphic position or the presence of the nucleotide C at said polymorphic position in at least 60% of the mitochondrial DNA molecules of said subject is indicative that the subject suffers from said disease or that the subject has an elevated risk of developing the disease.

25

The terms "diagnosis", "determine the risk", "sample" and "mitochondrial DNA" have been defined in the context of the first, second and third method of the invention and are used with the same meaning in the fourth method of the invention.

30

The presence of a specific nucleotide at a polymorphic position can be defined as the percentage of DNA molecules having such nucleotide at said polymorphic position with respect

to the total of DNA molecules present in the sample. According to the fourth method of the invention it is considered that the subject is suffering from Alzheimer's disease or that the subject has an elevated risk of developing said disease if the mtDNA of said subject presents the nucleotide C at the polymorphic position 16519 according to the sequence defined under 5 the accession number NC\_012920 when the percentage of molecules that present said nucleotide at that position is at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90% or more.

In a specific embodiment, it is considered that the subject has Alzheimer's disease or that the 0 subject has an elevated risk of developing the disease if the MtDNA of said subject presents the nucleotide C at the polymorphic position 16519 according to the sequence defined under the accession number NC\_012920 in 60% of mtDNA molecules. In a preferred embodiment of the invention it is considered that the subject has Alzheimer's disease or that the subject has an elevated risk of developing said disease if the MtDNA of said subject presents the 5 nucleotide C at the polymorphic position 16519 according to the sequence defined under the accession number NC\_012920 in 71% of mtDNA molecules.

In another preferred embodiment of the invention, it is considered that the subject has Alzheimer's disease or the subject is at a high risk of developing the disease if the mtDNA from 10 said subject presents the nucleotide C at the polymorphic position 16519 according to the sequence defined under accession number NC\_012920 in 74% of mtDNA molecules. In a preferred embodiment of the invention it is considered that the subject has Alzheimer's disease or that the subject is at a high risk of developing said disease if the MtDNA of said subject presents the nucleotide C at the polymorphic position 16519 according to the sequence 25 defined under the accession number NC\_012920 in 78% of mtDNA molecules.

As the person skilled in the art knows, a sample containing mitochondrial DNA can be homoplasmic or heteroplasmic. The term "heteroplasmy" or "heteroplasmic mitochondrial DNA" as used herein, refers to mitochondrial DNA of a subject that consists of a mixture of 30 DNA from at least two different genotypes of mitochondria. The term "homoplasmia" or "homoplasmic mitochondrial DNA" as used herein refers to mitochondrial DNA of a subject that is formed by a single genotype of DNA from a single mitochondrial genotype.

In a specific embodiment, in the context of the present invention the sample of mitochondrial 35 DNA is homoplasmic, in which case all mitochondria present in the subject contains identical genetic material so all mitochondria of a subject present in their genome nucleotide C at the

polymorphic position 16519 according to the sequence defined under the accession number NC\_012920. In those embodiments wherein the sample of mitochondrial DNA is homoplasmic, the fourth method of the invention allows to diagnose or determine the risk of developing Alzheimer's disease in a subject by detecting the C nucleotide in the polymorphic position

5 16519 according to the sequence defined under accession number NC\_012920 in said mitochondrial DNA sample; that is, the detection of the nucleotide C at the polymorphic position 16519 according to sequence defined under the accession number NC\_012920 in a sample of homoplasmic mitochondrial DNA from a subject is indicative that the subject suffers from Alzheimer's disease or that the subject has an elevated risk of developing the disease.

0 Conversely, the detection of the nucleotide T at the polymorphic position 16519 according to the sequence as defined under the accession number NC\_012920 in a sample of homoplasmic mitochondrial DNA, indicates that the subject is not suffering from Alzheimer's disease or that the subject has a low risk of developing the disease.

5 In another specific and preferred embodiment of the invention, the mitochondrial DNA sample is heteroplasmic, i.e. mtDNA subject is from two mitochondrial populations whose genetic material are not identical to each other. According to this invention, heteroplasmy refers to the fact that a percentage of said mitochondria of said subject present in its genome the oligonucleotide C at the polymorphic position 16519 according to the sequence defined under accession number NC\_012920 and the remaining percentage of mitochondria of said subject presents the oligonucleotide T in its in its genome at the polymorphic position 16519 according to the sequence defined under accession number NC\_012920.

25 The person skilled in the art will understand that the identification of the presence of polymorphism in at least 60% of mtDNA molecules is equally useful in order to select a subject for submission to preventive treatment of Alzheimer's disease, in cases where the mtDNA of the subject present heteroplasmy, as in the cases where the mtDNA of the subject present homoplasmy, the totality of molecules will present the nucleotide T or the nucleotide C at position 16519, in which case the percentage of mtDNA molecules with polymorphisms

30 indicative that the patient is a candidate to receive preventative treatment for Alzheimer's disease is either 100% or 0%. Like that, in the case of subjects whose mtDNA is heteroplasmic, there will be a population of MtDNA molecules that present the T nucleotide in position 16519 and a second population of molecules that present the C nucleotide in position 16519. In that case, it is considered that the patient is candidate for a preventive treatment of Alzheimer's

35 disease when the percentage of mtDNA molecules that present nucleotide C at position 16519 is equal to or greater than 60%.

The determination of homoplasmy and heteroplasmy and the percentage of heteroplasmy or the "degree of heteroplasmy" can be determined by any technique known by a person skilled in the art non-limiting illustrative examples of techniques to allow to determine whether a sample of mitochondrial DNA is heteroplasic include but are not limited to southern-blot, or PCR-RFLP of mtDNA sequencing. Briefly, the PCR-RFLP technique is based on the fact that normally the presence of a SNP in a sample is associated with the creation or destruction of specific sequences or target of one or more restriction enzymes. Heteroplasmy detection by PCR-RFLP technique is a first stage in the amplification of the genetic material region containing the polymorphism to be detected, by using specific oligonucleotides, followed by a second step where the amplified fragments are subjected to enzymatic digestion reaction in the presence of appropriate restriction enzyme. Since the presence or absence of the polymorphism in the sample is associated with the presence or absence of target specific restriction pattern obtained fragment sizes determine whether the sample is formed by a unique pattern of bands, in which case the sample is homoplasmic. If, however, the analysis determines the presence of two banding patterns, corresponding to two different mitochondrial DNA populations, then the mtDNA sample is heteroplasmic.

The term "single nucleotide polymorphism" or "single nucleotide polymorphism "or "SNP", as used herein, refers to a variation in the nucleotide sequence of a nucleic acid that occurs in a single nucleotide (A, C, T or G), where every possible sequence is present in a proportion equal to or greater than 1% of the population. These polymorphisms appear when a single nucleotide in the genome is altered (for example by substitution, addition or deletion). Each version of the sequence with respect to the polymorphic site is referred to as an allele of the polymorphic site. SNPs tend to be evolutionarily stable from generation to generation and, as such, can be used to study specific genetic abnormalities in a population.

The polymorphic variant of the invention is the position 16519 based on the numbering defined by the number NC\_012920 in the NCBI database. The polymorphic variant contains a C in said position.

The terms "sequence determination of a SNP "or "detecting a SNP" are used interchangeably herein, and refer to the determination of a sequence of a particular SNP in the subject under study. Determining the sequence of the SNP can be performed by various processes known to the person skilled in the art.

In some specific embodiments of the invention, the sample comprising mitochondrial DNA is selected from a biopsy of a solid tissue or biofluid. The samples can be obtained by conventional methods known to the person skilled in the art.

5 In an even more specific embodiment, the biofluid is selected from peripheral blood or cerebrospinal fluid.

In an even more specific embodiment, said solid tissue is brain tissue.

0 If the material in which it is desired to determine said SNP according the present method is a solid tissue or biofluid, preferably there is a prior nucleic acid extraction from the sample using any suitable technique for this.

5 In a preferred embodiment of the invention, the DNA fraction suitable for the implementation of the invention is the total DNA. DNA extraction may be carried out using any method known to the person skilled in the art as has been detailed in the first method of the invention.

0 If desired, the present method can be carried out in samples which previously the mitochondrial fraction has been isolated and subsequently the DNA thereof has been isolated. The isolation of the mitochondrial fraction can be performed using any known method of cell fractionation and have been detailed in the first method of the present invention.

25 After isolating and amplifying (if necessary) the nucleic acid, the SNP sequence of the invention is detected by any method or technique capable of determining nucleotides present in a SNP or polymorphism. For example, a SNP can be detected by performing sequencing, mini-sequencing, hybridization, restriction fragment analysis, oligonucleotide ligation assay, allele-specific PCR, or a combination thereof. As such, the systems and methods limited to, nucleic acid sequencing, hybridization methods and array technology (e.g. technology available 30 BioSciences Aclara, Affymetrix, Agilent Technologies, Inc. Illumina, etc.); can also be used in techniques based on mobility shift of nucleic acid fragments amplified, for example *Single Stranded Conformational Polymorphism*(SSCP), *denaturing gradient gel electrophoresis* (DGGE), *Chemical Mismatch Cleavage* (CMC), *Restriction Fragment Polymorphisms*(RFLP), PCR-RFLP, *WAVE analysis* and the like (Methods Mol. Med. 2004; 108: 173-88). Of course, 35 this list is merely illustrative and in no way limiting. The experts in the field may use any method appropriate to achieve such detection.

In another specific embodiment, determining the sequence of said SNP is performed by PCR-RFLP.

In a specific embodiment, the fourth method of the invention corresponds to an *in vitro* method to diagnose early stages of AD. The term "early stages of Alzheimer's disease" as used herein, refers to Alzheimer's disease stage I-II according to Braak scale defined in the context of the first method of the invention.

Fifth method of the invention

In the fifth feature, the invention corresponds to an *in vitro* method to select a subject for submission to a preventive treatment of Alzheimer's disease that involves determining in a sample containing mitochondrial DNA from said subject, the nucleotide at the polymorphic position 16519 according to the sequence with accession number NC\_012920 in the NCBI database, wherein the detection of the C nucleotide at said polymorphic position or the presence of the nucleotide C at said polymorphic position in at least 60% of the mitochondrial DNA molecules indicative that the subject is a candidate to receive treatment aimed at preventing Alzheimer's disease.

:0 The terms "Alzheimer", "treatment", "sample", "mitochondrial DNA" and "polymorphism" as well as methods for obtaining the sample and detecting a polymorphism have been detailed in the context of the first, second, third and fourth method of the invention and are used here with the same meaning.

25 According to the fifth method of the invention it is considered that the subject is a candidate to be submitted to treatment aimed at preventing Alzheimer's disease if the mtDNA of the subject has nucleotide C at the polymorphic position 16519 according to the sequence defined under access number NC\_012920 when the percentage of mtDNA molecules that have said nucleotide in said position is at least 60%, at least 65%, at least 70%, at least 75%, at least 30 80%, at least 85%, at least 90% or more. In a specific embodiment, it is considered that the subject is eligible to receive preventive treatment of Alzheimer's disease if the mtDNA of said subject presents the nucleotide C at the polymorphic position 16,519 according the sequence defined under accession number NC\_012920 in 60% of mtDNA molecules.

35 The person skilled in the art will understand that the identification of the presence of the polymorphism in at least 60% of mtDNA molecules is equally useful to select a subject to be

submitted to preventive treatment of Alzheimer's disease in cases where mtDNA of the subject present heteroplasmic as in the cases which it presents homoplasmic. Like that, as in the case of subjects whose mtDNA is homoplasmic, the totality of the molecules will present the nucleotide T or the nucleotide C at position 16519, in which case the percentage of mtDNA molecules with the polymorphisms indicative that the patient is a candidate to receive preventative treatment for Alzheimer's disease is either 100% or 0%. Like that, in the case of subjects whose mtDNA is heteroplasmic, there will be a population of mtDNA molecules which present the T nucleotide at position 16519 and a second population of molecules will present the C nucleotide at the position 16519. In that case, it will be considered that the patient is a candidate for preventive treatment of Alzheimer's disease when the percentage of molecules of mtDNA present the nucleotide C in position 16519 is less than or greater than 60%.

In some specific embodiments of the invention, the sample comprising mitochondrial DNA is selected from a biopsy of a solid tissue or biofluid, samples can be obtained by conventional methods known to the person skilled in the art.

In a specific embodiment, in the context of the present invention the sample of mitochondrial DNA is homoplasmic, in which case all mitochondria present in the subject contains identical genetic material so all mitochondria of a subject present nucleotide C in their genome in the polymorphic position 16519 according to the sequence under the accession number NC\_012920. In those embodiments where the mitochondrial DNA sample is homoplasmic, the fifth method of the invention allows selecting a patient to be submitted to preventive treatment of Alzheimer's disease through the detection of the nucleotide C at the polymorphic position 16519 according to the sequence defined under accession number NC\_012920 in said mitochondrial DNA sample; i.e. detection of nucleotide C at the polymorphic position 16519 according to the sequence defined under accession number NC\_012920 in a sample of homoplasmic mitochondrial DNA from a patient is indicative that the subject is eligible to receive preventive treatment of Alzheimer's disease. Conversely, detection of nucleotide T in at polymorphic position 16519 according to the sequence defined under accession number NC\_012920 in a sample of homoplasmic mitochondrial DNA from a patient is indicative that said patient is not a candidate to receive preventive treatment for Alzheimer's disease.

In another specific and preferred embodiment of the invention, the mitochondrial DNA sample is heteroplasmic, i.e. the mtDNA of the subject is of two mitochondrial populations whose genetic material are not identical to each other. According to the present invention heteroplasmy refers to a percentage of mitochondria of said subject that present in its genome

the oligonucleotide C in the polymorphic position 16519 according to the sequence defined under the accession number NC\_012920 and the remaining percentage of mitochondria of said subject present oligonucleotide T in the genome in the polymorphic position 16519 according to the sequence defined under accession number NC\_012920.

5

Methods for determining whether a sample is homoplasmic or heteroplasmic as well as to determine the degree of heteroplasmy of a sample have been detailed in the context of the fourth method of the invention and are incorporated herein by reference.

0

In an even more specific embodiment, the biofluid is selected from peripheral blood or cerebrospinal fluid.

In an even more specific embodiment, said solid tissue is brain tissue.

5

If the material in which it is desired to determine said SNP according the present method is a solid tissue or biofluid, preferably there is a prior nucleic acid extraction from the sample using any suitable technique for this. In a preferred embodiment of the invention, the DNA fraction suitable for the implementation of the invention is the total DNA. The extraction of DNA may be carried out using any appropriate method known to persons skilled in the art as has been detailed in the first method of the invention.

0

In another specific embodiment, determining the sequence of said SNP is performed by PCR-RFLP.

25

### Polynucleotides of the invention

30

In another feature, the present invention corresponds to a nucleic acid (hereinafter "first polynucleotide of the invention") comprising at least 9 contiguous polynucleotides in a mtDNA region wherein said region comprises at least one methylation site selected from the CpG sites in the D-loop region shown in Table 1.

35

The term "polynucleotide" as used herein, refers to DNA or RNA molecules of more than 13 bases in length. The polynucleotides of the invention are preferably DNA molecules of at least 14, at least 15, at least 16, at least 18, at least 20, at least 25, at least 30, at least 35, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100 or more bases in length.

In a specific embodiment, the first polynucleotide of the invention comprises at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 20, at least 25, at least 30 or more contiguous nucleotides to at least one CpG site selected from Table 1.

5 In another feature, the present invention relates to a nucleic acid (hereinafter "second polynucleotide of the invention") comprising at least 9 contiguous polynucleotide in a mitochondrial DNA region wherein said region contains at least one methylation site selected from CpG sites in the ND1 gene shown in Table 2

0 In a specific embodiment, the second polynucleotide of the invention comprises at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 20, at least 25, at least 30 or more contiguous nucleotides, to said at least one CpG site selected from Table 2.

5 In another feature, the present invention relates to a nucleic acid (hereinafter "second polynucleotide of the invention") comprising at least 9 contiguous polynucleotide in a mitochondrial DNA region wherein said region comprises at least one methylation site selected from the CHG sites in the D-loop region shown in Table 3.

10 In a specific embodiment, the third polynucleotide of the invention comprises at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 20, at least 25, at least 30 or more contiguous nucleotides, to said at least one CHG site selected from Table 3.

25 In another feature, the present invention relates to a nucleic acid (hereinafter "fourth polynucleotide of the invention") comprising at least 9 contiguous polynucleotide in a mitochondrial DNA region wherein said region contains at least one methylation site selected from CHG sites of the ND1 gene shown in Table 4.

30 In a specific embodiment, the fourth polynucleotide of the invention comprises at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 20, at least 25, at least 30 or more contiguous nucleotides to said at least one CHG site selected from Table 4.

35 In another aspect, the present invention refers to a nucleic acid (hereinafter 35 hereinafter "fifth polynucleotide of the invention") comprising at least 9 contiguous polynucleotide in a mtDNA region wherein said region comprises at least one methylation site selected from the CHH sites in the D-loop region shown in Table 5.

In a specific embodiment, the fifth polynucleotide of the invention comprises at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 20, at least 25, at least 30 or more contiguous nucleotides, to said at least one CHH site selected from Table 5.

5 In another feature, the present invention corresponds to a nucleic acid (hereinafter "sixth polynucleotide of the invention") Comprising at least 9 contiguous polynucleotide in a mitochondrial DNA region wherein said region comprises at least one methylation site selected from the CpG sites in the D-loop region shown in Table 1, wherein the corresponding position of cytosine in said CpG site is uracil.

0 In a particular embodiment, the sixth polynucleotide of the invention comprises at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 20, at least 25, at least 30 or more contiguous nucleotides, to said at least one CpG site selected from Table 1 wherein the corresponding position of cytosine in said CpG site is uracil.

5 In another aspect, the present invention refers to a nucleic acid (hereinafter "seventh polynucleotide of the invention") comprising at least 9 contiguous polynucleotides in a mitochondrial DNA region wherein said region comprises at least one methylation site selected from the CpG sites of the ND1 gene shown in Table 2, wherein the corresponding position of cytosine in said CpG site is uracil.

25 In a particular embodiment, the seventh polynucleotide of the invention comprises at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 20, at least 25, at least 30 or more contiguous nucleotides to said at least one CpG site selected from Table 2 wherein the corresponding position of cytosine in said CpG site is uracil.

30 In another feature, the present invention refers to a nucleic acid (hereinafter "eight polynucleotide of the invention") Comprising at least 9 contiguous polynucleotides in mitochondrial DNA region wherein said region comprises at least one methylation site selected from the CHG sites in the D-loop region shown in Table 3 wherein the corresponding position of cytosine in said CHG site is uracil.

35 In a particular embodiment, the eight polynucleotide of the invention comprises at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 20, at least 25, at least 30 or more contiguous nucleotides to said at least one CHG site selected from Table 3, wherein the corresponding position of cytosine in said CHG site is uracil.

In another feature, the present invention relates to a nucleic acid (hereinafter "ninth polynucleotide of the invention") comprising at least 9 contiguous polynucleotides in a mitochondrial DNA region wherein said region comprises at least one methylation site selected from the CHG sites of the ND1 gene shown in Table 4, wherein the corresponding position of cytosine in said CHG site is uracil.

5 In a specific embodiment, the ninth polynucleotide of the invention comprises at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 20, at least 25, at 0 least 30 or more contiguous nucleotides, that is said, at least one CHG site selected from Table 4, wherein the corresponding position of cytosine in said CHG site is uracil.

In another feature, the present invention relates to a nucleic acid (hereinafter "tenth polynucleotide of the invention") Comprising at least 9 contiguous polynucleotide in a 5 mitochondrial DNA region wherein said region comprises at least one methylation site selected from the CHH sites in the D-loop region shown in Table 5, wherein the corresponding position of cytosine in said CHH site is uracil.

10 In a specific embodiment, the tenth polynucleotide of the invention comprises at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 20, at least 25, at least 30 or more contiguous nucleotides to said at least one CHH site selected from Table 5, 0 wherein the corresponding position of cytosine in said CHH site is uracil.

25 In another feature, the invention relates to a polynucleotide that specifically hybridises with said first, second, third, fourth, fifth, sixth, seventh, eighth, ninth and tenth polynucleotides of the invention.

The expression "that specifically hybridises" or "capable of hybridising in a specific form", as used herein, refers to the ability of an oligonucleotide or of a polynucleotide of specifically 30 recognizing a sequence of a CpG, CHG or CHH site. As used herein the term "hybridization" It is the process of combining two nucleic acid molecules or single-stranded molecules with a high degree of similarity resulting in a simple double-stranded molecule by specific pairing between complementary bases. Normally hybridization occurs under very stringent conditions or moderately stringent conditions.

by comparing the nucleotide sequence of a molecule to the nucleotide sequence of a second molecule. The variants according to the present invention include nucleotide sequences that are at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% similar or identical to the sequence of said at least one CpG, CHH and CGH site selected from the sites shown in Tables 1 to 5.

5 The degree of identity between two nucleic acid molecules is determined using computer algorithms and methods that are widely known to persons skilled in the art. The identity between two amino acid sequences preferably determined by BLASTN algorithm (BLAST Manual, Altschul et al, 1990, NCBI NLM NIH Bethesda, Md. 20894, Altschul, S., et al, *J. Mol Biol.* 215: 403-10).

0 The "rigor" of hybridization reactions is readily determined by an ordinary expert in the field, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt concentration. In general, longer probes require higher temperatures for proper annealing while shorter probes need lower temperatures. Hybridization generally depends on

5 the ability of denatured DNA to reanneal when complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridisable sequence, the higher the relative temperature that can be used. As a result, it is deduced that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and 10 explanation of stringency of hybridization reactions, see Ausubel et al., *Current Protocols in Molecular Biology*, Wiley Interscience Publishers, (1995).

The term "stringent conditions" or "high stringency conditions" as used herein, typically: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium 25 chloride/0.0015 citrate /0.1% sodium dodecyl sulfate at 50°C sodium; (2) employ during hybridization a denaturing agent such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 mM buffer sodium 30 phosphate at pH 6.5 with 750 mM sodium chloride, 75 mM citrate sodium at 42°C; or (3) employ 50% formamide, 5 × SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5x Denhardt's solution, DNA sonicated 35 salmon sperm (50 mg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC (sodium chloride/sodium citrate) and 50% formamide, followed by a high-stringency wash 0.1 x SSC consisting containing EDTA at 55°C.

35 "Moderately stringent conditions" may be identified as described by Sambrook et al, *Molecular Cloning. A Laboratory Manual*, New York: Cold Spring Harbor Press, 1989, and include the

use of washing solution and hybridization conditions (e.g., temperature, ionic force and % SDS) less stringent than those described above. An example of moderately stringent conditions is overnight incubation at 37°C in a solution comprising: 20% formamide, 5 × SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10% dextran sulfate and 20 mg/ml salmon sperm DNA denatured fragmented, followed by washing the filters in 1 × SSC at about 35 37-50° C. the person skilled in the art will know how to adjust the temperature, ionic strength, etc. if necessary to accommodate factors such as probe length and the like.

0 Kits of the Invention

In another feature, the present invention relates to a kit (hereinafter "first kit of the invention") Comprising at least one oligonucleotide capable of specifically hybridizing and methylation-dependent manner with a mitochondrial DNA sequence comprising a methylation site selected 5 from the group formed by:

- (i) the CpG sites in the D-loop region shown in Table 1,
- (ii) the CpG sites in the ND1 gene shown in Table 2,
- (iii) the CHG sites in the D-loop region shown in Table 3,
- (iv) the CHG sites in the ND1 gene shown in Table 4, and/or
- (v) the CHH sites in the D-loop region shown in Table 5

The terms "CpG site", "CHG site", "CHH site", "D-loop region" and "ND1 gene" have been described in detail in the context of the first method of the invention and the term "capable of hybridising in a specific form" is defined in the context of the polynucleotides of the invention. 25 These terms are used with the same meaning in the context of the invention kits.

In a preferred embodiment, the oligonucleotides are part of the invention kit which is capable of specifically hybridizing dependently from methylation with a mitochondrial DNA sequence 30 comprising a methylation site selected from the group consisting of:

- (i) the CpG sites in the D-loop region shown in Table 1,
- (ii) the CpG sites in the ND1 gene shown in Table 2,
- (iii) the CHG sites in the D-loop region shown in Table 3,
- (iv) the CHG sites in the ND1 gene shown in Table 4, and/or
- (v) the CHH sites in the D-loop region shown in Table 5

They constitute at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90% or at least 100% of the total amount of the oligonucleotides that form the kit. In further embodiments, said oligonucleotides are at least 5 55% at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% of the total amount of oligonucleotides which form the kit.

Suitable kits include various reagents for use according to the present invention, in suitable 0 containers and packaging materials, including tubes, vials, and shrink wrap packages and blow molders. In addition, invention kits can contain instructions for the simultaneous, sequential or separate use of the various components found in the kit. Such instructions may be in the form of printed material or in the form of electronic media capable of storing instructions so that they can be read by a subject, such as electronic storage media (magnetic disks, tapes and the 5 like), optical media (CD-ROM, DVD) and the like. Additionally, or alternatively, the media can contain Internet addresses that provide such instructions.

Suitable materials for inclusion in an exemplary kit according to the present invention comprise one or more of the following: reagents capable of amplifying a specific sequence of a domain 0 either total or mtDNA DNA without the need to carry out PCR; reagents required to discriminate between the different possible alleles in the sequence domains amplified by PCR amplification or non-PCR (e.g., restriction endonuclease, hybridizing oligonucleotides preferably CpG, CHG and/or CHH sites methylated or unmethylated, including those modified to contain enzymes or fluorescent chemical groups that amplify the oligonucleotide signal and make that the 25 discrimination between CpG, CHG and/or CHH methylated or unmethylated sites CHH is more robust); or reagents required for physically separating the various products amplified regions (e.g., agarose or polyacrylamide and a buffer for use in electrophoresis, HPLC columns, SSCP gels, formamide gels or a support matrix for MALDI-TOF).

30 The term "oligonucleotide" as used herein, refers to a DNA molecule or short RNA, with up to bases in length. Oligonucleotides of the invention are preferably DNA molecules at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 15, at least 20, at least 30, at least 40, at least 45 or 50 bases of length.

35 As used in the kit of the invention, at least one oligonucleotide capable of hybridizing with at least a sequence of a CpG site in the D-loop region selected from the CpG sites shown in

Table 1, at least one oligonucleotide capable of hybridizing with at least one sequence of a CpG site in the ND1 gene shown in Table 2, at least one oligonucleotide capable of hybridizing with at least one sequence of a CHG site in the D-loop region shown in Table 3, at least one oligonucleotide capable of hybridizing with at least one sequence of a CHG site in the ND1 gene shown in Table 4 and/or at least one oligonucleotide capable of hybridizing with at least one sequence of a CHH site in the D-loop region shown in Table 5, in a methylation specific manner, it is used as a primer to amplify the region containing said CpG, CHG and/or CHH site(s) Alternatively, at least one oligonucleotide can also be used as a probe to detect such CpG, CHG or CHH methylated or unmethylated sites.

In a preferred embodiment of the first kit of the invention comprise oligonucleotides capable of specifically hybridizing with all CpG sites in the D-loop region shown in Table 1, all CpG sites of the ND1 gene shown in Table 2, with all of the CHG sites of the D-loop region shown in Table 3, with all of CHG sites in the ND1 gene shown in Table 4 and/or all CHH sites in the D-loop region shown in Table 5.

If desired, the first kit of the invention may comprise a first oligonucleotide capable of specifically hybridizing with bisulfites-treated oligonucleotide comprising at least one sequence of a CpG site in the D-loop region shown in Table 1 when said CpG site is methylated, and at

least one oligonucleotide or polynucleotide capable of specifically hybridizing with the same bisulfite-treated oligonucleotide comprising at least one sequence of a CpG site in the D-loop region when said CpG site is unmethylated; and/or the kit may comprise a first oligonucleotide capable of specifically hybridizing with bisulfites-treated oligonucleotide comprising at least one sequence of a CpG site in the ND1 gene shown in Table 2 when said CpG site is methylated, and at least one oligonucleotide or polynucleotide capable of specifically hybridizing with the same bisulfite-treated oligonucleotide comprising at least one sequence of a CpG site of the D-loop region when said CpG site is unmethylated; and/or the kit may comprise a first oligonucleotide capable of specifically hybridizing with a bisulfite-treated oligonucleotide comprising at least one sequence of a CHG site in the D-loop region shown in

Table 3 when said CHG site is methylated, and at least one oligonucleotide or polynucleotide capable of specifically hybridizing with the same bisulfite-treated oligonucleotide that comprises at least a sequence of a CHG site of the D-loop region when said CHG site is not methylated; and/or the kit may comprise a first oligonucleotide capable of specifically hybridizing with bisulfite-treated oligonucleotide comprising at least one sequence of a CHG site in the ND1 gene shown in Table 4 when said site CHG is methylated, and at least one oligonucleotide or polynucleotide capable of hybridizing specifically with the same bisulfites-

treated oligonucleotide that contains at least one sequence of a CHG site in the D-loop region when said CHG site is unmethylated; and/or the kit may comprise a first oligonucleotide capable of specifically hybridizing with bisulfites-treated oligonucleotide comprising at least a sequence of a CHH site in the D-loop region shown in Table 5 when said CHH site is 5 methylated and at least one oligonucleotide or polynucleotide capable of specifically hybridizing the same bisulfite-treated oligonucleotide comprising at least one sequence of a CHH site in the D-loop region when said CHH site is unmethylated.

For hybridization of an unmethylated CpG site, specific primers that hybridize with non- 0 methylated DNA, have, preferably, a T in the CG pair in "to distinguish it from the C retained in methylated DNA. It is preferable that primers contain relatively few Cs or Gs in the sequence since the Cs will be absent in the sense primer and the Gs absent in the antisense primer (cytosine converted into uracil, which is amplified as thymidine in the amplification product). Accordingly, for hybridization to a methylated CpG site, primers that specifically hybridize 5 methylated DNA preferably they have a C at the 3'CG pair.

In another feature, the invention relates to a kit (hereinafter "second kit of the invention") comprising at least one oligonucleotide capable of hybridizing.

:0 Specifically, with a zone in 5' to 3' position with respect to a methylation site of mitochondrial DNA selected from the group consisting of:

- i. the CpG sites in the D-loop region shown in Table 1,
- ii. the CpG sites in the ND1 gene shown in Table 2,
- 25 iii. the CHG sites in the D-loop region shown in Table 3,
- iv. the CHG sites in the ND1 gene shown in Table 4, and/or
- v. the CHH sites in the D-loop region shown in Table 5

where methylated cytosine in said position has converted to uracil or another base that is 30 distinguishable from cytosine in its hybridization properties.

In a preferred embodiment, the oligonucleotides that form part of the kit of the invention and which are capable of specifically hybridizing to a region or position 5' or position 3' with respect to a methylation site in the mitochondrial DNA selected from the group formed by:

- 35 i. the CpG sites in the D-loop region shown in Table 1,
- ii. the CpG sites in the ND1 gene shown in Table 2,

- iii. the CHG sites in the D-loop region shown in Table 3,
- iv. the CHG sites in the ND1 gene shown in Table 4, and/or
- v. the CHH sites in the D-loop region shown in Table 5

5 They constitute at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90% or at least 100% of the total amount of the oligonucleotides that form the kit. In further embodiments, said oligonucleotides constitute at least 55% at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% of the total amount of 0 oligonucleotides that make up the kit.

In a specific embodiment, the second kit of the invention further comprises one or more reagents to convert unmethylated cytosine to uracil or another base that is differentially detectable to cytosine in terms of hybridization properties.

5 In a preferred embodiment, the one or more reagents to convert unmethylated cytosine to uracil or another base that is detectably different from cytosine in terms of hybridization properties is bisulfite, preferably sodium bisulfite. The reagent capable of converting unmethylated cytosine to uracil or another base that is differentially detectable to cytosine in 0 terms of hybridization properties is metabisulfite, preferably sodium metabisulfite.

The term "conversion reagent" and its details are described in detail in context of the diagnostic method of the invention and used with the same meaning in the context of the kit according to the invention.

25 In a specific embodiment, the second kit of the invention comprises at least one oligonucleotide comprising a sequence selected from the sequences shown in SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4 and/or SEQ ID NO: 5.

30 In another feature, the invention refers to the use of the first and/or second kit of the invention to determine the pattern of mitochondrial DNA in a subject or to determine the diagnosis of a neurodegenerative disease in a subject selected from Alzheimer's disease and Parkinson's disease.

35 The invention is described by the following examples which are to be construed as merely illustrative and not limitative of the scope of the invention.

## Examples

### MATERIALS AND METHODS

5

#### Study design and subjects included

0

The study was conducted on 44 samples, including pathology related to Alzheimer's disease (AD), Parkinson's disease (PD) and control cases. The samples were processed in a plate divided into two lanes. In lane 1 amplicons D-Loop ND1 and samples were analyzed in entorhinal cortex cases related pathology EA and corresponding controls (Table 7). In lane 2 D-Loop amplicons were analyzed on samples of the substantia nigra in PD and its corresponding controls (Table 7). Each patient was identified in the primers used with MID (*multiplex Identifier*) (Tables 7 and 8). Methylation in CpG sites and non CpG sites (CHG and CHH, where H = A, T, or C) was analyzed using a pyrosequencer 454 GS FLX Titanium Roche that generated 569.684 sequences in lane 1, whose lengths varied from 40-1098 base pairs (bp) with an average length of about 417 bp. In lane 2 the number of sequences obtained was 513.579, whose length ranged from 40 to 933 bp (466 bp average length). Alignments of the sequences obtained for each MID, amplicon and lane against their respective reference sequences were noted and the percentages of identity between these were nearly 100%. The mean and median rates of bisulphite conversion for each locus and MID were analyzed.

5

10

15

20

25

30

The number of unmethylated sequences exceeded the number of methylated sequences identified in each site and some methylation sites were not present. Those sequences after alignment were found to have, at least, one not present site of the methylation pattern not, were removed from the analysis to avoid any bias at the time of quantization. This circumvents approximation analysis of putative mitochondrial pseudogenes whose amplicons presented nearly 100% of identity with mitochondrial DNA when analyzed in NCBI BLAST. Most CpG, CHG, and CHH sites analyzed were unmethylated. However, different methylation sites could be identified.

Lane	Amplicon	Site	Criteria	No. of Sites	C vs. AD I-II	C vs. AD III-IV	AD I-II vs. AD III-IV	C vs. PD
L1	D-loop	CG	FDR<0.01	18	17	14	12	-
L1	D-loop	CG	FDR<0.05	18	17	15	16	-

Lane	Amplicon	Site	Criteria	No. of Sites	C vs. AD I-II	C vs. AD III-IV	AD I-II vs. AD III-IV	C vs. PD
L1	ND1	CG	FDR<0.01	13	7	7	0	-
L1	ND1	CG	FDR<0.05	13	7	7	0	-
L1	D-loop	CHG	FDR<0.01	16	13	10	1	-
L1	D-loop	CHG	FDR<0.05	16	14	12	7	-
L1	ND1	CHG	FDR<0.01	9	6	5	0	-
L1	ND1	CHG	FDR<0.05	9	7	5	0	-
L1	D-loop	CHH	FDR<0.01	52	0	0	34	-
L1	D-loop	CHH	FDR<0.05	52	23	0	43	-
L1	ND1	CHH	FDR<0.01	72	0	0	0	-
L1	ND1	CHH	FDR<0.05	72	0	0	0	-
L2	D-loop	CG	FDR<0.01	18	-	-	-	17
L2	D-loop	CG	FDR<0.05	18	-	-	-	17
L2	D-loop	CHG	FDR<0.01	16	-	-	-	14
L2	D-loop	CHG	FDR<0.05	16	-	-	-	14
L2	D-loop	CHH	FDR<0.01	52	-	-	-	44
L2	D-loop	CHH	FDR<0.05	52	-	-	-	44

**Table 6:** Number of differentially methylated sites. FRD: p-value adjusted Benjamini method and Hochberg (1995)

MID	Lane	Region of brain	Diagnosis	Stage of Braak for AD	Stage of Braak for PD	Gender	Age	Time postmortem (h)
1	L1	entorhinal	CONTROL	0/0	0	M	56	5
2	L1	entorhinal	CONTROL	0/0	0	M	56	3.45
3	L1	entorhinal	CONTROL	0/0	0	F	55	8.3
4	L1	entorhinal	CONTROL	0/0	0	F	66	4.15
5	L1	entorhinal	CONTROL	0/0	0	F	64	5
6	L1	entorhinal	CONTROL	0/0	0	F	52	5.45
7	L1	entorhinal	CONTROL	0/0	0	M	57	5.20

2015238195

25 Jun 2019

MID	Lane	Region of brain	Diagnosis	Stage of Braak for AD	Stage of Braak for PD	Gender	Age	Time postmortem (h)
8	L1	entorhinal	CONTROL	0/0	0	M	64	3.3
9	L1	entorhinal	AD associated	I/A	0	F	57	5
10	L1	entorhinal	AD associated	I/A	0	F	64	2.15
11	L1	entorhinal	AD associated	I	0	M	59	16.3
12	L1	entorhinal	AD associated	II/A	0	F	86	4.15
13	L1	entorhinal	AD associated	I/A	0	M	67	14.4
14	L1	entorhinal	AD associated	II/A	0	M	66	4.55
15	L1	entorhinal	AD associated	I/A	0	F	63	8.05
16	L1	entorhinal	AD associated	II/A	0	F	76	5.45
17	L1	entorhinal	AD associated	III/A	0	F	71	6.45
18	L1	entorhinal	AD associated	III-A	0	F	77	11.5
19	L1	entorhinal	AD associated	III/B	0	M	86	3.1
20	L1	entorhinal	AD associated	IV/C	0	F	69	8.1
9	L1	entorhinal	AD associated	I/A	0	F	57	5
10	L1	entorhinal	AD associated	I/A	0	F	64	2.15

MID	Lane	Region of brain	Diagnosis	Stage of Braak for AD	Stage of Braak for PD	Gender	Age	Time postmortem (h)
11	L1	entorhinal	AD associated	I	0	M	59	16.3
12	L1	entorhinal	AD associated	II/A	0	F	86	4.15
21	L1	entorhinal	AD associated	III	0	F	79	3.4
22	L1	entorhinal	AD associated	IV/C	0	M	75	6.1
23	L1	entorhinal	AD associated	III/A	0	F	74	4
24	L1	entorhinal	AD associated	III/0	0	M	87	3.3
1	L2	SN	CONTROL	0/0	0	F	78	3.4
2	L2	SN	CONTROL	0/0	0	M	66	3
3	L2	SN	CONTROL	0/0	0	F	71	8.3
4	L2	SN	CONTROL	0/0	0	M	30	4.1
5	L2	SN	CONTROL	0/0	0	M	47	5
6	L2	SN	CONTROL	0/0	0	M	67	5
7	L2	SN	CONTROL	0/0	0	M	39	9.15
8	L2	SN	CONTROL	0/0	0	M	85	5.45
9	L2	SN	CONTROL	0/0	0	F	46	9.35
10	L2	SN	CONTROL	0/0	0	F	77	3.15
11	L2	SN	PD	0/0	3.4	M	68	9.2
12	L2	SN	PD	0/0	3.4	F	81	6.3
13	L2	SN	PD	0/0	3.4	M	76	12.3
14	L2	SN	PD	0/0	3.4	F	77	3.3
15	L2	SN	PD	0/0	3.4	F	78	27.3
16	L2	SN	PD	0/0	3.4	F	69	4.3

MID	Lane	Region of brain	Diagnosis	Stage of Braak for AD	Stage of Braak for PD	Gender	Age	Time postmortem (h)
5	L2	SN	CONTROL	0/0	0	M	47	5
6	L2	SN	CONTROL	0/0	0	M	67	5
7	L2	SN	CONTROL	0/0	0	M	39	9.15
8	L2	SN	CONTROL	0/0	0	M	85	5.45
9	L2	SN	CONTROL	0/0	0	F	46	9.35
10	L2	SN	CONTROL	0/0	0	F	77	3.15
11	L2	SN	PD	0/0	3.4	M	68	9.2
12	L2	SN	PD	0/0	3.4	F	81	6.3
13	L2	SN	PD	0/0	3.4	M	76	12.3
14	L2	SN	PD	0/0	3.4	F	77	3.3
15	L2	SN	PD	0/0	3.4	F	78	27.3
16	L2	SN	PD	0/0	3.4	F	69	4.3
17	L2	SN	PD	I/0	4	M	69	6
18	L2	SN	PD	0/0	4	F	84	4.3
19	L2	SN	PD	0/0	5	M	76	12
20	L2	SN	PD	0/0	5	M	80	7.3

**Table 7:** Summary of key clinical features and neuropathological human cases analyzed.

Braak stages Alzheimer's disease (AD) indicate the degree of presence of neurons with

fibrillary neurodegeneration (roman numerals) and senile plaques (letters) following the

5 classification of Braak. Braak stages for Parkinson's disease (PD) refers to the degree of presence of the protein  $\alpha$ -synuclein (Lewy bodies).

number MID	MID sequence	SEQ ID NO:
1	ACGAGTGCCT	6
2	ACGCTCGACA	7
3	AGACGCACTC	8
4	AGCACTGTAG	9

number MID	MID sequence	SEQ ID NO:
5	ATCAGACACG	10
6	ATATCGCGAG	11
7	CGTGTCTCTA	12
8	CTCGCGTGTG	13
9	TAGTATCAGC	14
10	TCTCTATGCG	15
11	TGATACGTCT	16
12	TACTGAGCTA	17
13	CATAGTAGTG	18
14	CGAGAGATAC	19
15	ATACGACGTA	20
16	TCACGTACTA	21
17	CGTCTAGTAC	22
18	TCTACGTAGC	23
19	TGTACTACTC	24
20	ACGACTACAG	25
21	CGTAGACTAG	26
22	TACGAGTATG	27
23	TACTCTCGTG	28
24	TAGAGACGAG	29

**Table 8:** MID sequence associated with each case analyzed**Human brain samples**

5 Tissue samples were provided by the Neurological Tissue Bank, University of Barcelona - Hospital de Barcelona and the Bank of Institute of Neuropathology, HUB-ICO-IDIBELL. The donation and procurement of samples was regulated by the ethics committee of both institutions. Half of each brain was maintained in buffered formalin 4% solution for morphological and histological study, while the other half was processed in coronal sections

10 to be frozen at -80C to be available for biochemical studies. Neuropathological examination in all controls and pathological cases took place in thirty standardized sections of the brain,

cerebellum and brain stem, which were stained with hematoxylin and eosin, and Klüver Barrera, or processed for immunohistochemistry for glial fibrillary acidic protein, microglial markers, beta-amyloid, phosphorylated tau (AT8 antibody)  $\alpha$ -synuclein,  $\beta$ -crystallin, ubiquitin and TDP-43. Cases with pathology related to AD and PD were classified according to the current neuropathological criteria (Braak and Braak 1991, 1999; Braak et al, 2003, 2006). Cases with mixed pathology (including vascular lesions) were excluded from this study. The brains used as control belonged to individuals without neurological manifestations and without injury in the neurological study. The medical records were reexamined for each case, and cases with pathology related to AD (stages I-IV) were reevaluated through phone calls or interviews with relatives, asking if they had evidence of any neurological or cognitive impairment. Only cases meeting these criteria were considered in this work. All cases analyzed summarized in Table 7.

The average post-mortem interval samples entorhinal cortex was  $4.98 \pm 1.57$  hours in controls, 7.51  $\pm$  5.13 hours in stages I-II, and  $5.70 \pm 2.85$  hours stage III-IV; for samples of the substantia nigra intervals were  $5.59 \pm 2.46$  hours and 9.23 in controls  $\pm$  7.07 hours in the case of PD.

### **Murine brain samples**

Transgenic mice APP/PS1 and wild strain were obtained from Jackson Laboratory (USA). The transgenic model expresses a murine/human chimera APP (Mo/HuAPP695swe: Swedish mutation APP) molecule and the human variant of presenilin 1 (PS1-DE9), both expression in neurons of the nervous system product. The animals were housed in standard conditions with light and dark cycles of 12 hours, with unlimited access to food and water. Stabbing was conducted following the ethical guidelines (European Communities Council Directive 86/609/EEC) approved by the local ethics committee.

### **Total DNA extraction**

Total DNA was isolated human samples of the entorhinal cortex and substantia nigra (Table 7) using the *DNeasy Blood and Tissue Kit* (Qiagen, Madrid, Spain) following the manufacturer's instructions. Total DNA samples murine was obtained from the frontal cortex using the same procedure.

### **Bisulfite treatment**

Three hundred nanograms of DNA were treated using bisulfite *EZ DNA kit methylation Kit* (Zymo Research, Ecogen, Barcelona, Spain) following the supplier's instructions. The bisulfite-treated DNA was resuspended in 30 $\mu$ l to reach a final concentration of 10 ng/ $\mu$ l. All samples were treated with bisulfite in parallel, using the same batch of reagent to avoid differences in the rate of bisulfite conversion between different commercial lots.

### Design of the primers for amplicon FLX-Loop D and ND1

The primers for the FLX experiment were designed following the instructions techniques for FLX sequencer Roche "Amplicon Fusion Primer Design Guidelines for GS FLX Titanium Series Lib-A Chemistry". Melting Primers for amplicons containing a directional primer GS FLX Titanium primer A or primer B (including a key sequence (Key) Four bases) in Lot 5 - oligonucleotide premium, plus a specific sequence for mold at 3 - first final. Moreover, a MID sequence was added (*multiplex Identifier*) Between the primer A (or primer B) and the specific sequence for subsequent automated sample identification software after steps grouping/multiplexing and sequencing. The primers used contained the following components: Primer *forward* (Primer A -Key -MID - specific sequence mold), 5'-CGTATGCCCTCCCTCGCGCCA (SEQ ID NO: 33) - -MID TCAG - specific template sequence 3'; Primer *reverse* (Primer B- Key-MID - specific template sequence): 5'CTATGCGCCTTGCCAGCCCGC (SEQ ID NO: 34) - -MID TCAG - specific template sequence -3'. Specific template sequences for each of the amplicons: D- loop- direct, 5'-TAGGGGTTTTGATTATTATTATT -3'(SEQ ID NO: 2) and D Loop -reverso, ACAAACATTCAATTATTATTATTATATCCT 5'-3' (SEQ ID NO: 3); ND1 - direct, 5'-ATGGTTAATTTTATTATTATTATTGTATT -3' (SEQ ID NO: 4) and ND1 -reverso, 5'-TAATTTAAATTAACTCACCTAATCAA -3'(SEQ ID NO: 5). The primers used in this study were designed to avoid CpG sites. Sequence MIDs specific for each patient shown in Table 8. amplified regions (D-Loop: 16386-256; ND1: 3313-3686) are based on nucleotide position map of human mtDNA ([www.mitomap.org](http://www.mitomap.org)).

### Preparation of Amplicon library

The PCR amplicons for D-Loop and ND1 were performed according to the manual *Amplicon Library Preparation Method Manual (GS FLX Titanium Series)* Of Roche. PCRs for twenty nanograms total bisulfite-treated DNA were used. DNA amplification bisulfite treated are carried out in a reaction volume 25  $\mu$ l. Each PCR reaction consisted of: 1x FastStart 10x Buffer # 2 0,05U/ $\mu$ l polymerase *HiFi FastStart Polymerase* (Roche), 200 nM of each dNTP, and 200

nM of each primer specific *forward* Y *reverse*. The primers were synthesized with a HPLC purification quality (Sigma - Aldrich, Madrid, Spain). Amplifications were performed in a thermocycler *Applied Biosystems Verity* ® (Applied Biosystems, Madrid, Spain) using the following conditions: 94°C for 3 min and then 36 cycles of 94°C for 15s, temperature annealing 5 (61°C ND1, and 62°C D-Loop) for 45s and 72°C for 1 min, followed by a final extension step at 72°C for 8 min and a final hold temperature at 4°C. two microliters of each PCR product were checked on an agarose gel stained with 1.5% SYBR ® Safe DNA Gel Stain (Invitrogen, Madrid, Spain).

#### 0 **Purification of PCR**

Purification of PCR products was performed using the *kit* Agencourt®AMPure®XP PCR Purification (Beckman Coulter, Madrid, Spain) following the instructions manual Roche *Amplicon Library Preparation Method Manual (GS FLX Titanium Series)*.

#### 5 **Quantification of amplicons and sequencing libraries FLX**

Quantification and quality control amplicon libraries and the rest of FLX sequencing protocol 0 was performed by the team of the Platform for Genomic Research Institute of Vall d'Hebron (VHIR, Barcelona, Spain).

#### 25 **Selection of differentially methylated sites**

Alignment and identification of CpG, CHG, and CHH sites and bisulfite conversion rates were performed using the HT Analyzer BIQ (Lutsik et al, 2011) software. Quality control of raw data and all statistical analysis.

30 They performed using the statistical language R and bioconductor software, <http://www.bioconductor.org>.

The selection of differentially methylated sites was based on the calculation of Fisher's Exact test statistic Test, considering those sites differentially methylated with a p-value set using Benjamini and Hochberg method (1995) below 0.05. The  $\alpha$ -value represented in *Heatmaps* 35 graphs is the ratio of methylated sequences with respect to the overall sum of methylated and unmethylated sequences per site (Du et al, 2010, BMC Bioinformatics, 30; 11: 587), i.e.  $\beta_{ij} =$

$M_i/(M_i + U_i)$  Where M is the number of methylated sequences at the site (i) and MID (j), and U is the number of non-methylated at the site (i) and MID (j) sequences.

**Example 1: DNA methylation is increased in the D-loop region and reduced in the ND1 gene in cases with early stages of AD-related pathology**

An increase in methylation in CpG and non CpG sites (CHG and CHH) was observed in the D-loop region in cases with AD pathology related to stages I/II and III/IV of Braak (Fig. 1).

- 0 The degree of methylation was higher in cases of pathology related to AD with respect to the control samples, and higher stage I/II versus stage III/IV, as shown in log2 graphics (OR) (Fig. 2). However, no differences in methylation of CHH sites between controls and cases with AD-related pathology at stages III/IV were found.
- 5 ND1 analysis revealed the presence of some less methylated CpG and CHG sites in cases with related AD pathology stage I/II and III/IV compared to control samples ( $\log_2 [O] > 0$ , Fig. 3). No differences were found for the CHH sites.

**Example 2: DNA methylation is reduced in the D-Loop region in the substantia nigra of PD cases.**

In contrast to what was observed in the entorhinal cortex in AD, the D-loop region showed a loss of methylation in almost all CpG and non CpG sites in the substantia nigra of PD cases with respect to the control samples (Fig. 3). However, as with AD, the percentage of DNA methylation represents a small part of the total mitochondrial DNA.

**Example 3: DNA methylation is increased in the region D-Loop in murine models of AD.**

In this study, mice APP/PS1 (murine model of AD) six months of age ( $n = 4$ ) and control mice of the same age ( $n = 4$ ) were used. As can be seen in Figure 4A, the CpG sites showed hypermethylation in samples obtained from pathology of mouse models with respect to the control samples. These results were also seen in CHG sites (see Figure 4B).

To determine the methylation pattern as advance of AD, they employed mice APP/PS1 of three, six and twelve months. At three months, the APP/APS1 mice have a low degree of accumulation of neuritic plaques, which increases in APP/APS1 mice six months old. AD

model mice six months of age show cognitive and memory failures. Finally, APP/APS1 animals of 12 months show similar symptoms to those observed in humans in advanced stages of AD.

The analysis of the degree of methylation in the D-Loop region corroborated the pattern observed in samples of human brain tissue. As shown in Figure 5, a tendency to hypomethylation of CpG sites CHG and CHH was observed in model mice in advanced stages of AD with respect to said CpG, CHG and CHH sites model mice in early stages of AD (Figure 5A, 5B and 5C).

**0 Example 4: The presence of the polymorphic position 16519 in mitochondrial DNA is associated with the development of Alzheimer's disease.**

The results shown in Table 9 show that the presence of the polymorphism T16519C is associated with Alzheimer's disease. Said results were obtained by conventional sequencing and using a chi-square test.

Type of sample Number of Samples analyzed)	presence of polymorphism T16519C (%)	p-value
Controls (n = 46)	56.5	-
AD I/II (n = 47)	76,6	0,02
AD III/IV (n=47)	74.4	0.03
AD V/VI (n = 46)	76.1	0,02

**Table 9:** Presence of polymorphism in samples obtained from individuals controls and individuals with Alzheimer's disease (AD) at different stages (Roman numerals).

The presence of heteroplasmy was observed in some cases and was performed again genotyping samples by using PCR-RFLP. The presence of the C allele creates a restriction site for the restriction enzyme HaeIII. The amplification of the sequence shown in SEQ ID NO: 30 was carried out using oligonucleotides SEQ ID NO: 31 and SEQ ID NO: 32. The following band patterns depending on the genotype were obtained:

Genotype T: 183 bp, 318 bp

Genotype C: 61 bp, 183 bp, 257

2015238195

25 Jun 2019

Heteroplasmia: 61 bp, 183 bp, 257 bp, 319 bp.

## CLAIMS

1. *in vitro* method to diagnose or determine the risk of developing a neurodegenerative disease selected from Alzheimer's disease and Parkinson's disease in a subject, comprising to determine in a sample from said subject comprising mitochondrial DNA, methylation pattern in the in the D-loop region, and/or in the ND1 gene, wherein the methylation pattern is determined in at least one site selected from the group consisting of:

- (i) the CpG sites in the D-loop region shown in Table 1,
- (ii) the CpG sites of the ND1 gene shown in Table 2,
- (iii) the CHG sites in the D-loop region shown in Table 3,
- (iv) the CHG sites in the ND1 gene shown in Table 4, and/or
- (v) the CHH sites in the D-loop region shown in Table 5,

wherein a hypermethylation in at least one of said CpG sites in the D-loop region, a hypermethylation in at least one of said CHG sites in the D-loop region, a hypermethylation in at least one of said CHH sites in the D-loop region, a hypomethylation in at least one of said CpG sites in the ND1 gene and/or a hypomethylation in at least one of said CHG sites in the ND1 gene is indicative that the subject suffers from Alzheimer's disease or that the subject is at a high risk of developing Alzheimer's disease; or

wherein a hypomethylation in at least one of said CpG sites in the D-loop region, a hypomethylation in at least one of said CHG sites in the D-loop region and/or a hypomethylation in at least one of the said CHH sites in the D-loop region it is indicative that the subject suffers from Parkinson's disease or that the subject at an elevated risk of developing Parkinson's disease.

2. The method according to claim 1, wherein a hypomethylation in at least one such CpG sites in the D-loop region or a hypomethylation in at least one of said CHG sites in the D-loop region with respect to the methylation pattern in a subject in stage I-II disease Alzheimer is indicative that the subject suffers from Alzheimer's disease in stage III-IV.

3. *in vitro* method to select a subject for submission to treatment to prevent the progression of a neurodegenerative disease selected from Alzheimer's disease or Parkinson's disease in a subject, comprising:

determining in a sample from said subject containing mitochondrial DNA methylation pattern in the D-loop region and/or in the ND1 gene, wherein the methylation pattern is determined in at least one site selected from the group formed by:

- i. the CpG sites in the D-loop region shown in Table 1,
- ii. the CpG site in the ND1 gene shown in Table 2,
- iii. the CHG sites in the D-loop region shown in Table 3,
- iv. the CHG sites in the ND1 gene shown in Table 4, and/or
- v. the CHH sites in the D-loop region shown in Table 5,

wherein a hypermethylation in at least one of said CpG sites in the D-loop region, a hypermethylation in at least one of said CHG sites in the D-loop region, a hypermethylation in at least one of said CHH sites in the D-loop region, a hypomethylation in at least one of said CpG sites in the ND1 gene and/or a hypomethylation in at least one of said CHG sites in the ND1 gene it is indicative that the subject is a candidate to receive treatment aimed at preventing Alzheimer's disease; or wherein a hypomethylation in at least one of said CpG sites in the D-loop region, a hypomethylation in at least one of said CHG sites in the D-loop region and/or a hypomethylation in at least one of the said CHH sites in the D-loop region is indicative that the subject is a candidate to receive treatment aimed at preventing Parkinson's disease.

**4. Method to monitor the progression of a neurodegenerative disease selected from Alzheimer's disease or Parkinson's disease in a subject comprising:**

a) determining in a sample from said subject comprising mitochondrial DNA the methylation pattern in the D-loop region, and/or in the ND1 gene, wherein the methylation pattern is determined at least one site from the group formed by:

- (i) the CpG sites in the D-loop region shown in Table 1,
- (ii) the CpG sites in the ND1 gene shown in Table 2,
- (iii) the CHG sites in the D-loop region shown in Table 3,
- (vi) the CHG sites in the ND1 gene shown in Table 4, and/or the CHH sites in the D-loop region shown in Table 5,

b) comparing the methylation pattern determined in step a) with said methylation pattern obtained in an earlier stage of the disease,

wherein a hypermethylation in at least one of said CpG sites in the D-loop region, a hypermethylation in at least one of said CHG sites of the D-loop region, a hypermethylation in at least one of said CHH sites in the D-loop region, a hypomethylation in at least one of said CpG sites of the ND1 gene and/or a hypomethylation in at least one of said CHG sites in the ND1 gene relative to said methylation pattern determined at an earlier stage of the disease, is indicative of the progression of Alzheimer's disease; and

wherein a hypomethylation in at least one of said CpG sites in the D-loop region, a hypomethylation in at least one of the said CHG sites in the D-loop region, and/or a hypomethylation in at least one of said CHH sites in the D-loop region with respect to said determined methylation pattern in an earlier stage of disease is indicative of the progression of Parkinson's disease.

**5.** The method according to claim 4, wherein the subject has been diagnosed with Alzheimer's disease stage I-II or wherein the subject has been diagnosed with Parkinson's disease in stage III -V.

**6.** The method according to claims 1 to 5, wherein determining the methylation pattern comprises determining methylation of all CpG sites in the D-loop region shown in Table 1.

**7.** The method according to claims 1 to 6, wherein determining of the methylation pattern comprises determining methylation in all CpG sites of the ND1 gene shown in Table 2.

**8.** The method according to claims 1 to 7, wherein the determination of the methylation pattern comprises determining the methylation of all CHG sites in the D-loop region shown in Table 3.

**9.** The method of any of claims 1 to 8, wherein the determination of the methylation pattern comprises determining the methylation pattern of all CHG sites in the ND1 gene shown in Table 4.

**10.** The method according to claims 1 to 9, wherein determining the methylation pattern comprises determining methylation of all CHH sites in the D-loop region shown in Table 5.

**11.** The method according to any of claims 1 to 10, wherein the methylation pattern is

determined by a technique selected from the group consisting of methylation specific PCR, bisulfite sequencing, techniques based on restriction-digestion, pyrosequencing, assay ChIP-on-chip, differential conversion, differential restriction and differential weight of the methylated site(s).

**12.** The method of claim 11, wherein the methylation pattern is determined by pyrosequencing.

**13.** Use of a kit to determine the diagnosis of a neurodegenerative disease in a subject selected from Alzheimer's disease and Parkinson's disease, wherein the kit comprises - at least one oligonucleotide capable of specifically and in a methylation-dependent manner hybridizing with a mitochondrial DNA sequence comprising a methylation site selected from the group formed by:

- (i) the CpG sites in the D-loop region shown in Table 1,
- (ii) the CpG sites in the ND1 gene shown in Table 2,
- (iii) the CHG sites in the D-loop region shown in Table 3,
- (iv) the CHG sites in the ND1 gene shown in Table 4, and/or
- (v) the CHH sites in the D-loop region shown in Table 5.

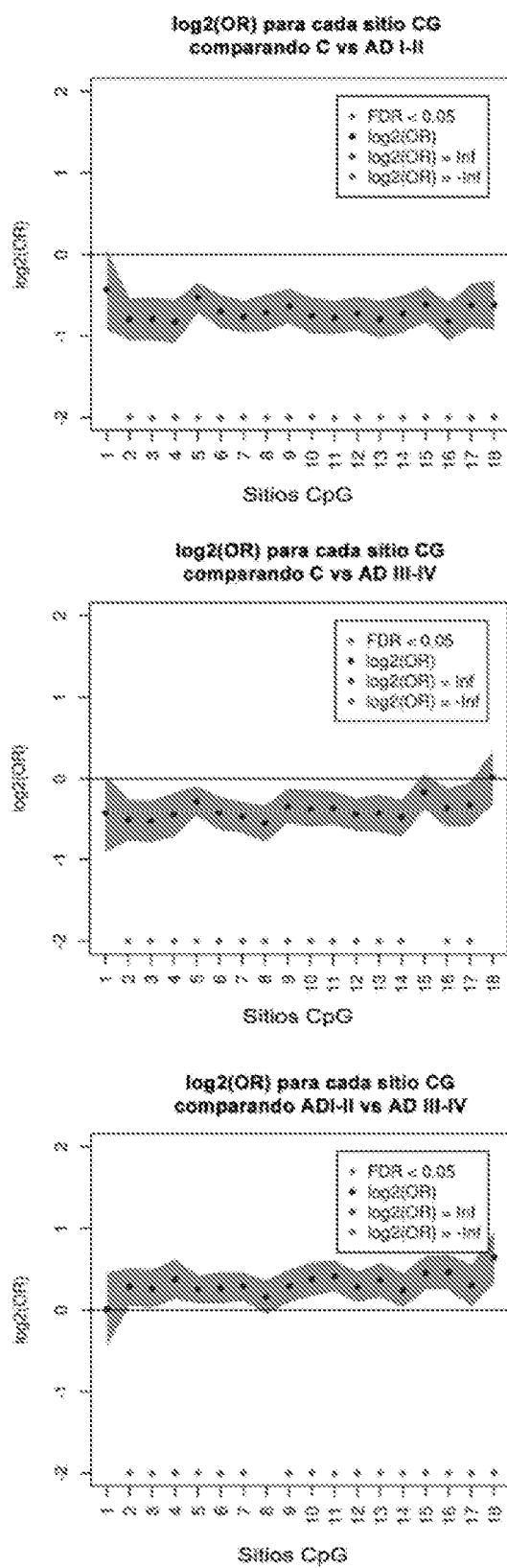
**14.** Use of a kit to determine the diagnosis of a neurodegenerative disease in a subject selected from Alzheimer's disease and Parkinson's disease, wherein the kit comprises at least one oligonucleotide capable of specifically hybridizing with a region 5' or 3' position with respect to a methylation site in mitochondrial DNA selected from the group consisting of:

- (i) the CpG sites in the D-loop region shown in Table 1,
- (ii) the CpG sites in the ND1 gene shown in Table 2,
- (iii) the CHG sites in the D-loop region shown in Table 3,
- (iv) the CHG sites in the ND1 gene shown in Table 4, and/or
- (v) the CHH sites in the D-loop region shown in Table 5,

wherein the methylated cytosine in said position has converted to uracil or another base that is distinguishable from cytosine in its hybridization properties.

**15.** The use of a kit according to claim 14, wherein said at least one oligonucleotide comprises a sequence selected from SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4 and/or SEQ ID NO: 5.

A)

**Figura 1**

B)

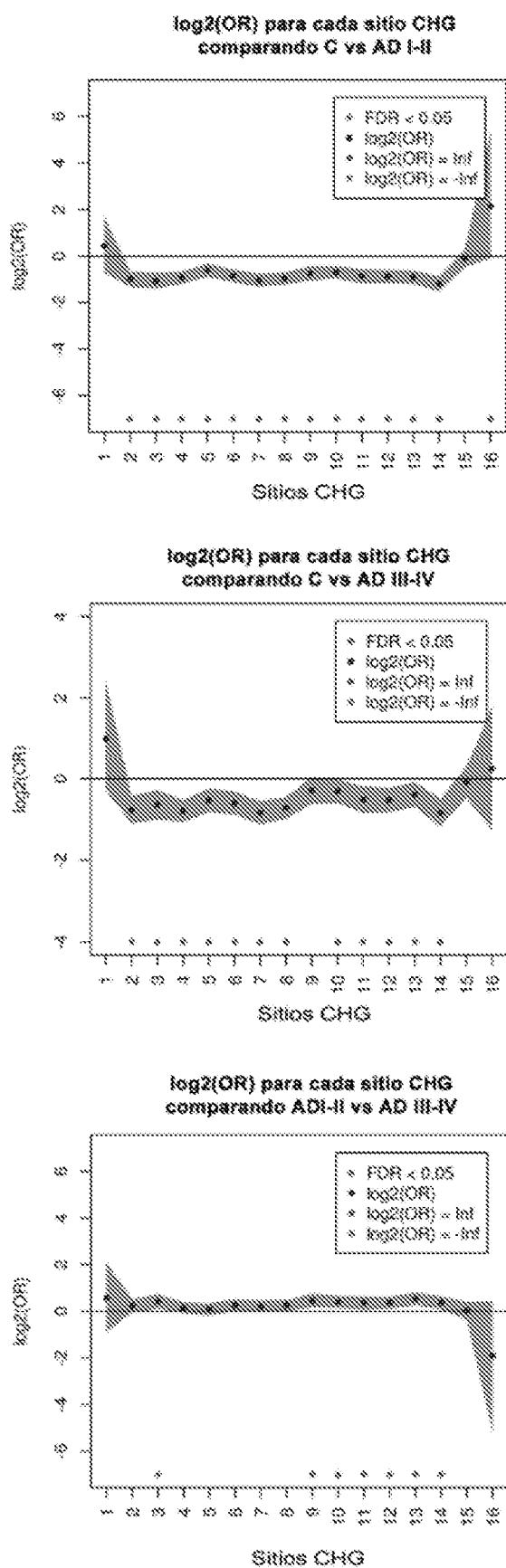


Figura 1 (cont.)

C)

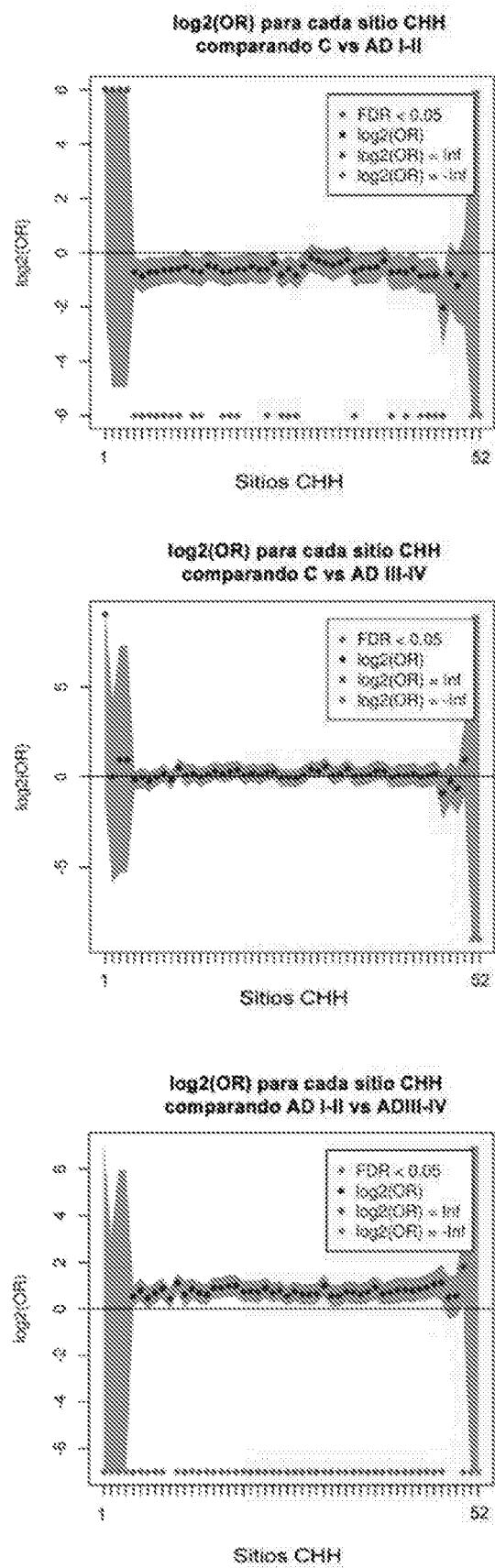
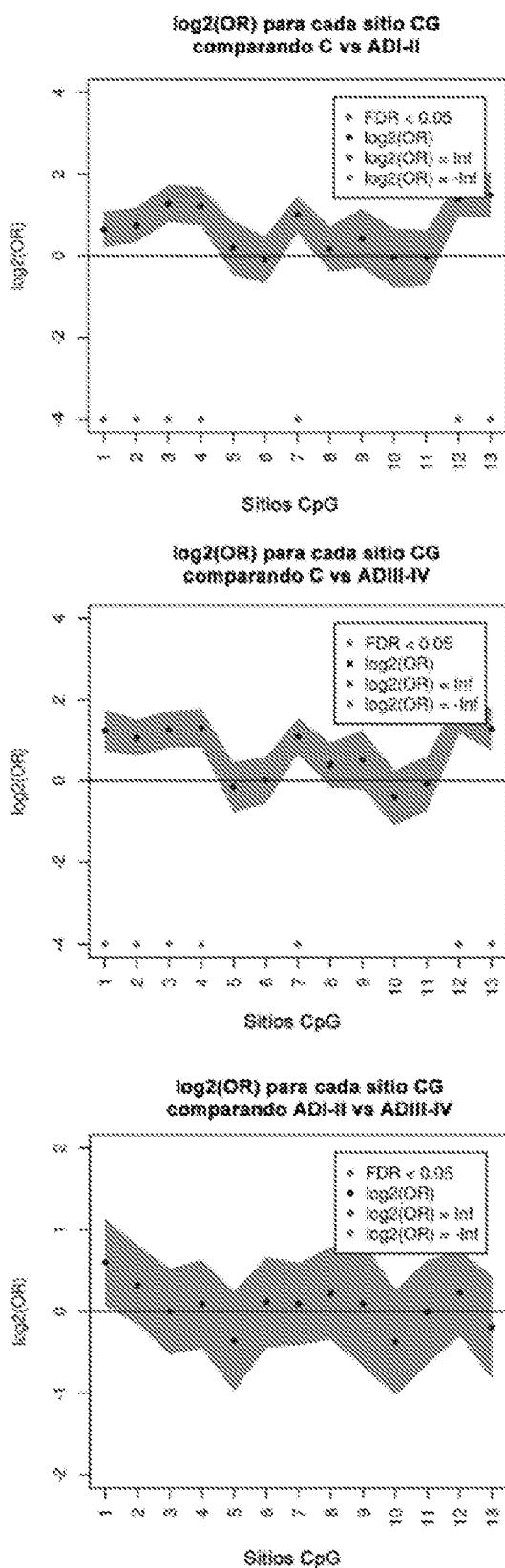


Figura 1 (cont.)

A)

**Figura 2**

B)

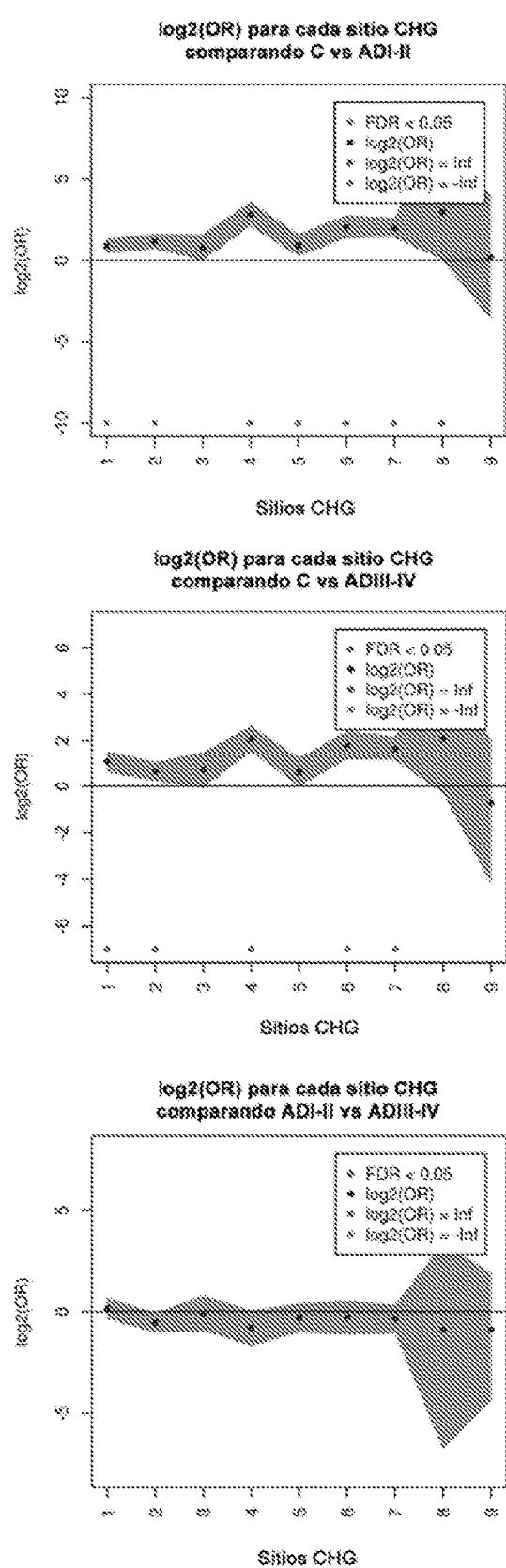
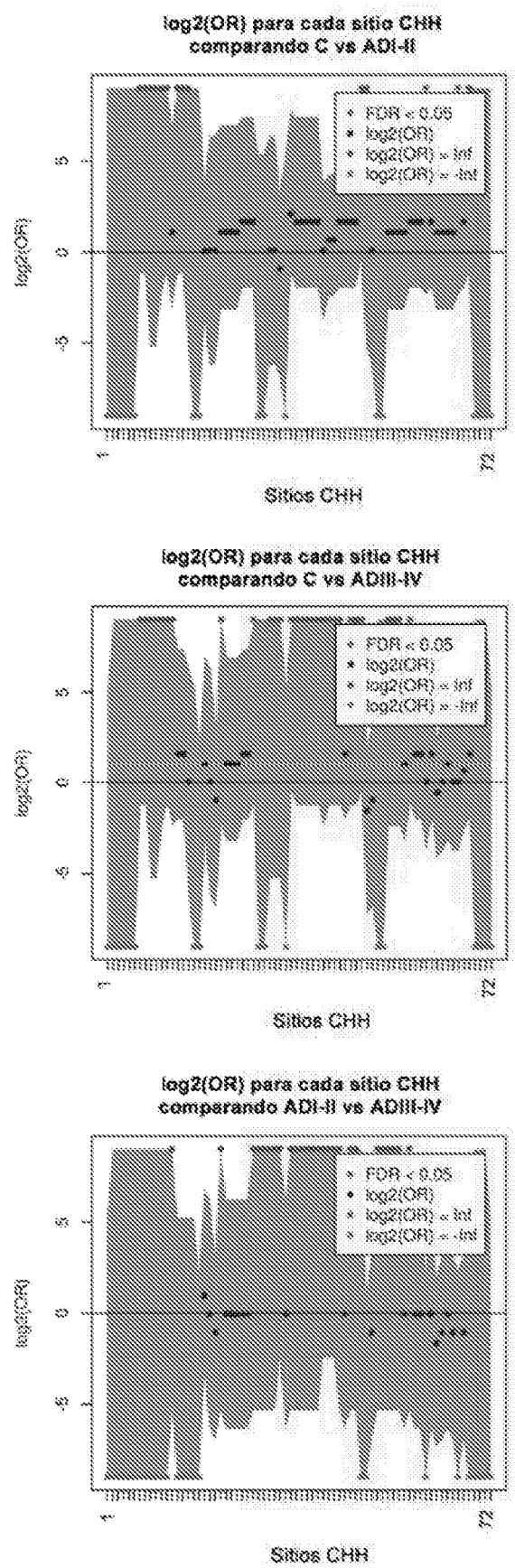
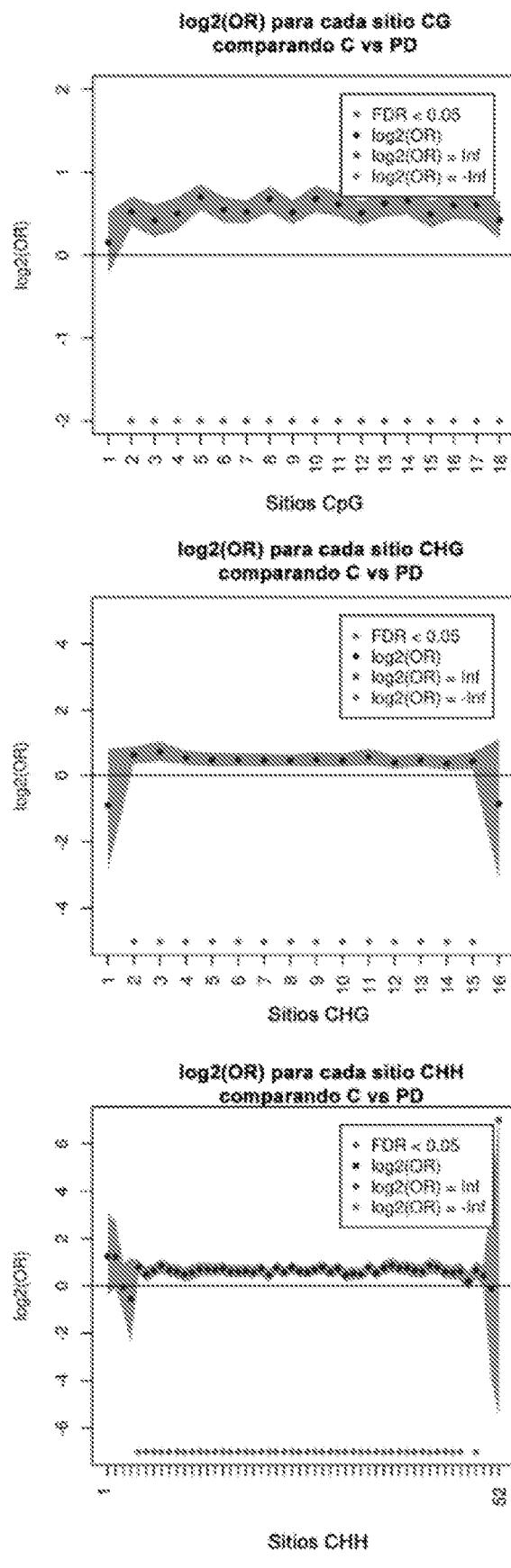
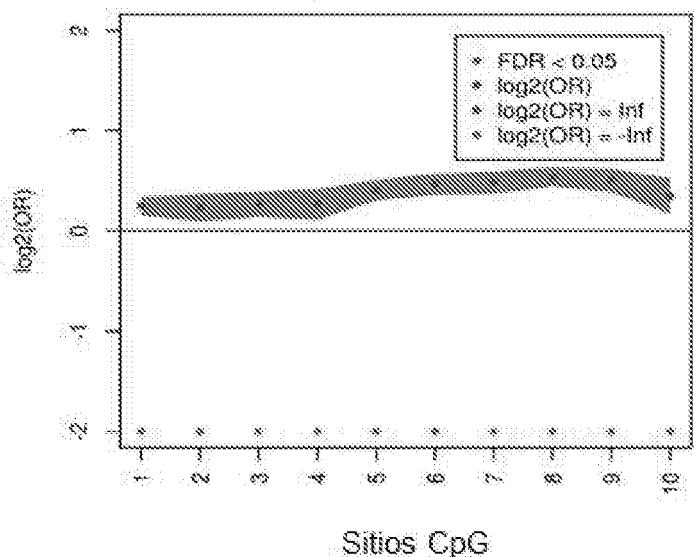
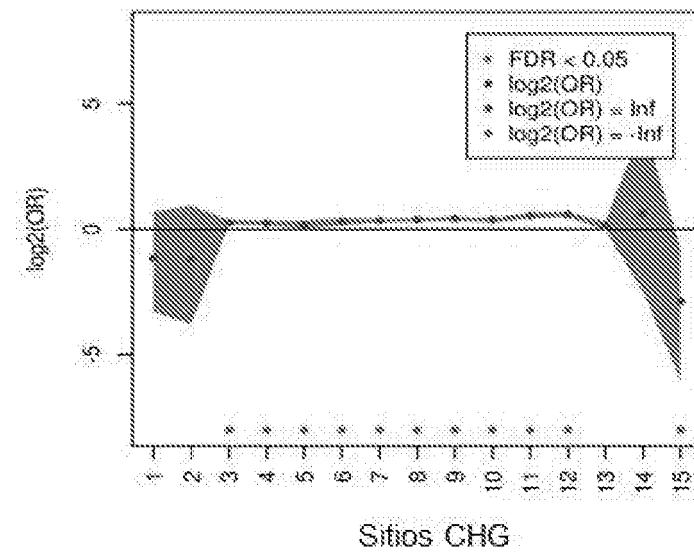


Figura 2 (cont.)

C)

**Figura 2 (cont.)**

**Figura 3**

**A)****Log2(OR) para cada sitio CG  
comparando TG.6 vs WT.6****B)****Log2(OR) para cada sitio CHG  
comparando TG.6 vs WT.6****Figura 4**

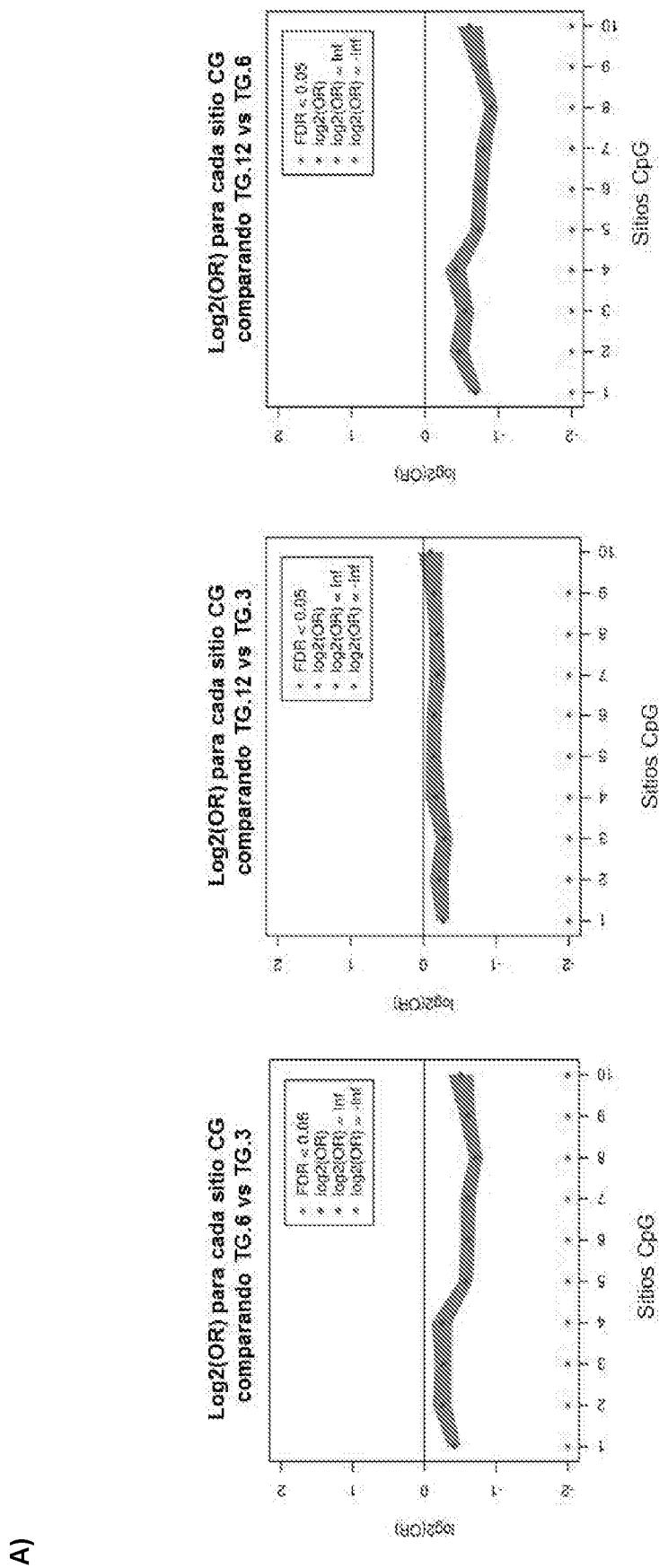


Figura 5

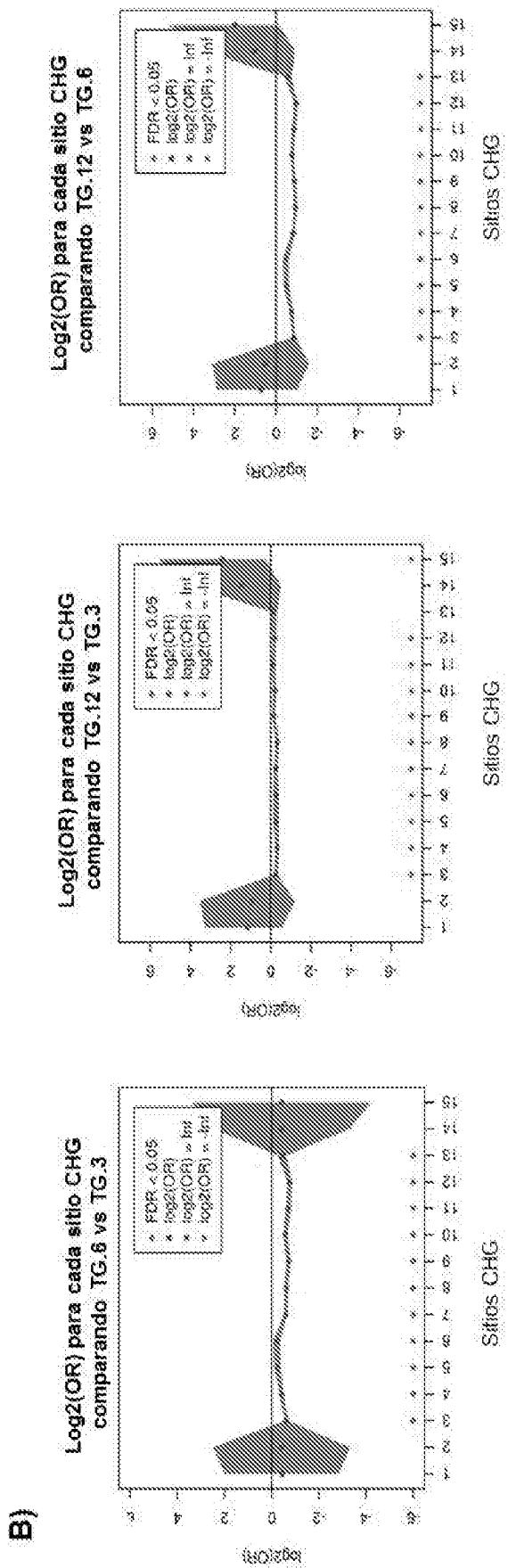


Figura 5 (cont)

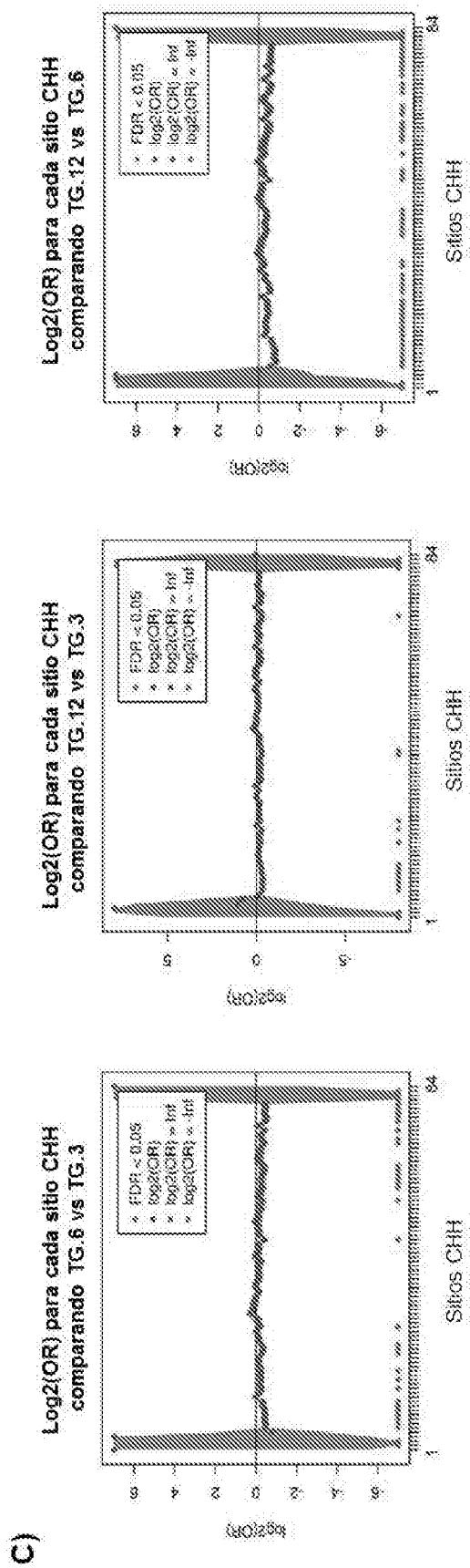


Figura 5 (cont)

2906064\_1.txt  
SEQUENCE LISTING

<110> INSTITUT D'INVESTIGACIÓ BIOMÉDICA DE BELLVITGE (IDIBELL)  
<120> MARCADORES MITOCONDRIALES DE ENFERMEDADES NEURODEGENERATIVAS  
<130> P10635PC00  
<150> ES P201430444  
<151> 2014-03-28  
<160> 34  
<170> PatentIn version 3.5  
<210> 1  
<211> 16569  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<222> (3107)..(3107)  
<223> n is a, c, g, or t

<400> 1  
gatcacaggt ctatcacccct attaaccact cacgggagct ctccatgcat ttggtatTTT 60  
cgtctggggg gtatgcacgc gatagcattg cgagacgctg gagccggagc accctatgtc  
gcagtatctg tctttgattc ctgcctcatc ctattattta tcgcacctac gttcaatatt 120  
acaggcgaac atacttacta aagtgtgtta attaattaat gcttgttagga cataataata 180  
acaattgaat gtctgcacag ccactttcca cacagacatc ataacaaaaaa atttccacca 240  
aaccggccctt ccccgcttc tggccacagc acttaaacac atctctgcca aacccaaaaa 300  
acaagaacc ctaacaccag cctaaccaga tttcaaattt tatctttgg cggtatgcac 360  
ttttaacagt caccggccaa ctaacacatt atttccctt cccactccca tactactaat 420  
ctcatcaata caaccggcc ccatcctacc cagcacacac acaccgctgc taaccata 480  
ccccgaacca accaaacccc aaagacaccc cccacagttt atgttagctt cctcctcaaa 540  
gcaatacact gaaaatgttt agacgggctc acatcacccc ataaacaaat aggtttggc 600  
ctagccttcc tattagctt tagtaagatt acacatgcaa gcatccccgt tccagtgagt 660  
tcaccctcta aatcaccacg atcaaaagga acaagcatca agcacgcagc aatgcagctc 720  
aaaacgctt aactaagcta tactaaccgg agggttggc aatttcgtgc cagccaccgc 780  
acgaaagttt aactaagcta tactaaccgg gaaacagcag tgattaacct ttagcaataa 840  
ggtcacacga ttaacccaaag tcaatagaag ccggcgtaaa gagtgttta gatcacccccc 900  
tcccccaataa agctaaaact cacctgagtt gtaaaaaact ccagttgaca caaaatagac 960  
tacgaaagt gcttaacat atctgaacac acaatagcta agacccaaac tgggattaga 1020  
taccccaacta tgcttagccc taaacctcaa cagttaaatc aacaaaactg ctcgcccagaa 1080  
cactacgagc cacagcttaa aactcaaagg acctggcggt gcttcataatc cctctagagg 1140  
agcctgttct gtaatcgata aaccccgatc aacctcacca cctcttgctc agcctatata 1200  
1260

ccgccccatctt	cagcaaaccc	tgatgaaggc	tacaaagtaa	gcgcaagtac	ccacgtaaag	1320
acgttaggtc	aagggtgtac	ccatgaggtg	gcaagaaatg	ggctacattt	tctacccca	1380
aaaactacga	tagcccttat	gaaacttaag	ggtcgaaggt	ggattttagca	gtaaactaag	1440
agtagagtgc	ttagttgaac	agggccctga	agcgcgtaca	caccgcccgt	caccctcctc	1500
aagtatactt	caaaggacat	ttaactaaaa	cccctacgca	tttatataga	ggagacaagt	1560
cgtaacatgg	taagtgtact	ggaaagtgca	cttggacgaa	ccagagtgt	gcttaacaca	1620
aagcacccaa	cttacactta	ggagatttca	acttaacttgc	accgctctga	gctaaaccta	1680
gccccaaacc	cactccacct	tactaccaga	caaccttagc	caaaccattt	acccaaataa	1740
agtataggcg	atagaaatttgc	aaacctggcg	caatagat	agtaccgcaa	gggaaagatg	1800
aaaaattata	accaagcata	atatagcaag	gactaacccc	tataccttct	gcataatgaa	1860
ttaactagaa	ataactttgc	aaggagagcc	aaagctaaga	cccccgaaac	cagacgagct	1920
acctaagaac	agctaaaaga	gcacacccgt	ctatgttagc	aaatagtggg	aagattata	1980
ggtagaggcg	acaaacctac	cgagcctggt	gatagctggt	tgtccaagat	agaatcttag	2040
ttcaacttta	aatttgccca	cagaaccctc	taaatcccct	tgtaaatttta	actgttagtc	2100
caaagaggaa	cagctcttg	gacactagga	aaaaaccttg	tagagagat	aaaaaattta	2160
acacccatag	taggcctaaa	agcagccacc	aattaagaaa	gcgttcaagc	tcaacaccca	2220
ctacctaaaa	aatcccaaac	atataactga	actcctcaca	cccaatttgg	ccaatctatc	2280
accctataga	agaactaatg	ttagtataag	taacatgaaa	acattctcct	ccgcataagc	2340
ctgcgtcaga	ttaaaacact	gaactgacaa	ttaacagccc	aatatctaca	atcaaccaac	2400
aagtattat	taccctcact	gtcaacccaa	cacaggcatg	ctcataagga	aaggttaaaa	2460
aaagtaaaag	gaactcggca	aatcttaccc	cgcctgttta	ccaaaaacat	cacctctagc	2520
atcaccagta	ttagaggcac	cgcctgccc	gtgacacatg	tttaacggcc	gcggtaccct	2580
aaccgtgcaa	aghtagcata	atcacttgc	ccttaaatag	ggacctgtat	gaatggctcc	2640
acgagggttc	agctgtctct	tacttttac	cagtgaattt	gacctgccc	tgaagaggcg	2700
ggcataacac	agcaagacga	gaagacccta	tggagcttta	atttattaaat	gcaaacagta	2760
cctaacaaac	ccacagggtcc	taaactacca	aacctgcatt	aaaaatttgc	gttggggcga	2820
cctcgagca	gaacccaacc	tccgagcagt	acatgctaag	acttcaccag	tcaaagcgaa	2880
ctactatact	caattgatcc	aataacttgc	ccaacggaaac	aagttaccct	agggataaca	2940
gchgcaatcct	attctagat	ccatatcaac	aatagggttt	acgacctcga	tgttgatca	3000
ggacatcccg	atggtgacgc	cgctattaaa	ggttcggttgc	ttaacacgatt	aaagtccctac	3060
gtgatctgag	ttcagaccgg	agtaatccag	gtcggttct	atctacnttc	aaattcctcc	3120
ctgtacgaaa	ggacaagaga	aataaggcct	acttcacaaa	gcmcctccc	ccgtaaatga	3180
tatcatctca	acttagtatt	atacccacac	ccacccaaga	acagggtttg	ttaagatggc	3240
agagcccggt	aatcgcataa	aactaaaac	tttacagtca	gaggttcaat	tcctttctt	3300
aacaacatac	ccatggccaa	cctcctactc	ctcattgtac	ccattcta	cgcaatggca	3360

2906064\_1.txt

ttcctaata	gc ttaccgaac	gaaaattcta	ggctatata	ac aactacgca	aggccccaa	3420
gtttaggc	cctacggg	ct actacaac	cc ttcgctgac	ccataaaact	cttcaccaaa	3480
gagccc	ctaa aacccgccc	ac atctaccat	cc acccttaca	tc accgcccc	gaccttagct	3540
ctcaccat	cg ctcttact	atgaaccc	cc tccccata	ccaaaccc	ct ggtcaac	3600
aacctagg	cc tcctattt	at tctagccacc	tc tctagcctag	ccg tttactc	aatcctctga	3660
tcagggtg	ag catcaaact	aa actacg	cc ctgatcg	cg cactgcg	ag tagccccaa	3720
acaatct	cat atgaagt	tc ac cccat	attctactat	ca acattact	aataagtggc	3780
tccttaacc	tc tctccacc	ct tatacaca	ca agaa	ca acc	tctgattact	3840
tgaccctt	gg ccataat	at tctatct	cc ac acta	cg ag acca	accccttc	3900
gaccttgcc	g aaggggag	tc cgaacta	tc tagg	ttca acat	cgccgcaggc	3960
cccttcgccc	tattcttcat	agccgaa	atac aacat	ttataataaa	caccctcacc	4020
actacaat	ct tcctagga	ac aacat	at gac	gc actct	ccc ctgaact	4080
tttgcacca	ag accct	act tcta	ac ctcc	ctgatc	taa ac	4140
cgattccg	ct acgacca	act catac	ctc t	ctatgaaa	actcacc	4200
gcattactt	ta tatgat	at gt	ctccat	accaat	cc agcatt	4260
taagaaat	at gtctgata	aa agagtt	actt	tgatagag	taa attag	4320
ccttattt	ct aggactat	gaatc	gaa	catccaa	at tctcg	4380
acctatc	aca cccatc	taa aggt	tc agt	aaataa	gctatcg	4440
aaaatgtt	gg ttatacc	cccgtact	aa ttaatcc	ggcccaac	cc gtcat	4500
ctaccat	tt tgcagg	caca ctc	ctcat	cgacta	gattc	4560
taggc	cttgcaggcaca	ctcatc	acag	cgctaa	acttcc	4620
gttcc	ac a gtc	aa gatatt	cc tc	actgaa	cttacc	4680
taatagct	at cctttca	ac aatata	ctc	ccggaca	ataatcc	4740
atcaata	actc atc	atcattaa	atcata	aaat	acta	4800
ttcactt	ctg agtccc	agag	gttacca	gcaccc	ctcttgc	4860
tcacat	gaca aaaa	acttagcc	cccat	ctca	tcaactaa	4920
taagc	ttcttctc	tcaat	ctt	ccatcat	aggcagt	4980
acccaa	acca	atcttag	cat	actc	ttga	5040
tagcag	ttct	accgtaca	ac	ccatt	tttaact	5100
taact	actac	cgcatt	cct	taact	ccat	5160
ctcg	cacctg	aaacaag	cta	taact	ccat	5220
taggagg	cct	gcttaa	acatgact	accctt	ccacc	5280
caaaaa	acca	atcc	actcc	taat	ccat	5340
acttctac	ct	acgc	cttaa	actcc	ccata	5400

## 2906064\_1.txt

taaaaataaa atgacagttt gaacatacaa aacccacccc attcctcccc acactcatcg	5460
cccttaccac gctactccta cctatctccc cttttatact aataatctta tagaaattta	5520
ggtaaatac agaccaagag cctcaaagc cctcagtaag ttgcaatact taatttctgt	5580
aacagctaag gactgcaaaa ccccaactctg catcaactga acgcaaatca gccactttaa	5640
ttaagctaag cccttactag accaatggga cttaaaccctaa caaacactta gttAACAGCT	5700
aagcacccta atcaactggc ttcaatctac ttctccgccc gcccggaaaa aaggcgggag	5760
aagccccggc aggtttgaag ctgcttcttc gaatttgc aaatcacccctc	5820
ggagctggta aaaagaggcc taacccctgt ctttagattt acagtccat gttcactca	5880
gccatTTac ctcacccca ctgatgttcg ccgaccgtt actattctct acaaaccaca	5940
aagacattgg aacactatac ctattattcg gcgcattgagc tggagtccta ggcacagctc	6000
taagcctcct tattcgagcc gagctggcc agccaggca cttcttaggt aacgaccaca	6060
tctacaacgt tattcgatcaca gcccattgtt ttgtataat cttcttcata gtaataccca	6120
tcataatcgg aggctttggc aactgacttag ttcccctaat aatcggtgcc cccgatatgg	6180
cgtttccccg cataaaacaac ataagcttct gactcttacc tccctctctc ctactcctgc	6240
tcgcattctgc tatagtggag gcccggcag gaacaggtt aacagtctac cttcccttag	6300
cagggacta ctcccacccct ggagcctccg tagaccta acatcttctcc ttacacctag	6360
caggtgtctc ctctatctta gggccatca atttcatcac aacaattatc aatataaaac	6420
cccctgcccatacccaatac caaaacgcccc tcttcgtctg atccgtccta atcacagcag	6480
tcctacttct cctatctctc ccagtcctag ctgctggcat cactatacta ctaacagacc	6540
gcaacctcaa caccacccctt ttcgaccccg ccggaggagg agacccattt ctataccaa	6600
acctattctg atttttcggt caccctgaag tttatattct tattctacca ggcttcggaa	6660
taatctccca tattgttaact tactactccg gaaaaaaaaga accatttggta tacataggta	6720
tggctgagc tatgatcatca attggcttcc tagggtttat cgtgtgagca caccatata	6780
ttacagtagg aatagacgta gacacacgag catatttcac ctccgctacc ataatcatcg	6840
ctatccccac cggcgtcaaa gtatTTAGCT gactcgccac actccacggaa agcaatatga	6900
aatgatctgc tgcagtgttc tgagccctag gattcatctt tcttttccacc gtaggtggcc	6960
tgactggcat tgtatttagca aactcatcac tagacatcgta actacacgac acgtactacg	7020
ttgttagccca cttccactat gtcctatcaa tagagctgt atttgcattc ataggaggct	7080
tcatttcactg atttccctta ttctcaggct acaccctaga ccaaaccctac gccaaatcc	7140
atTTCTACTAT catattcatc ggcgtaaatc taactttctt cccacaacac tttctcgcc	7200
tatccggaaat gccccgacgt tactcgact accccgatgc atacaccaca tgaaacatcc	7260
tatcatctgt aggctcattc atttctctaa cagcagtaat attaataatt ttcatgattt	7320
gagaaggcctt cgcttcgaag cgaaaagtcc taatagtaga agaaccctcc ataaacctgg	7380
agtgactata tggatgcccc ccaccctacc acacattcga agaaccctga tacataaaat	7440
ctagacaaaa aaggaaggaa tcgaacccccc caaagctggt ttcaagccaa ccccatggcc	7500

2906064\_1.txt

tccatgactt tttcaaaaag gtattagaaa aaccattca taactttgtc aaagttaat	7560
tataggctaa atcctatata tcttaatggc acatgcagcg caagtagggtc tacaagacgc	7620
tacttcccct atcatagaag agcttatac ctttcatgtat cacgccccta taatcatttt	7680
ccttatctgc ttcctagtcc tgtatgcctt tttcctaaca ctcacaacaa aactaactaa	7740
tactaacatc tcagacgctc agggaaataga aaccgtctga actatcctgc cggccatcat	7800
ccttagtcctc atcgccctcc catccctacg catccttac ataacagacg aggtcaacga	7860
tccctccctt accatcaaataat caattggcca ccaatggtac tgaacctacg agtacaccga	7920
ctacggcgga ctaatcttca actcctacat acttccccca ttattcctag aaccaggcga	7980
cctgcgactc cttgacgtt acaatcgagt agtactcccg attgaagccc ccattcgtat	8040
aataattaca tcacaagacg tcttgactc atgagctgtc cccacattag gcttaaaaac	8100
agatgcaatt cccggacggtc taaaccaaacc cactttcacc gctacacgac cgggggtata	8160
ctacggtcaa tgctctgaaa tctgtggagc aaaccacagt ttcatgccc tcgtcctaga	8220
attaattccc ctaaaaatct ttgaaatagg gcccgtatcc accctatagc accccctcta	8280
cccccctctag agcccactgt aaagctaact tagcattaac cttttaagtt aaagattaag	8340
agaaccaaca cctctttaca gtgaaatgcc ccaactaaat actaccgtat ggcccaccat	8400
aattacccccc atactcctta cactattcct catcacccaa ctaaaaatataat taaacacaaa	8460
ctaccaccta cctccctcac caaagcccat aaaaataaaaa aattataaca aaccctgaga	8520
accaaaaatga acgaaaaatct gttcgcttca ttcatggccc ccacaatcct aggcctaccc	8580
gccgcagttac tgatcattct atttccccctt ctattgtatcc ccaccccttccaa atatctcatc	8640
aacaaccgac taatcaccac ccaacaatga ctaatcaaacc taacccctaaa acaaattgata	8700
accatacaca acactaaagg acgaacctga tctcttatac tagtacccctt aatcattttt	8760
attgccacaa ctaacccctt cggactcctg cctcactcat ttacaccaac caccaacta	8820
tctataaacc tagccatggc catccccctta tgagcgggca cagtattat aggcttcgc	8880
tctaagatta aaaatgcctt agcccacttc ttaccacaaag gcacacccat acccccttatac	8940
cccatacttag ttattatcga aaccatcagc ctactcatttcc aaccaatagc cctggccgtt	9000
cgcctaaccg ctaacattac tgccaggccac ctactcatgc acctaattgg aagcggccacc	9060
ctagcaatat caaccattaa cctcccttca acacttatca tcttcacaat tctaattcttca	9120
ctgactatcc tagaaatcgc tgctgcctta atccaaagcctt acgttttccatc acttcttagta	9180
agcctctacc tgacgacaa cacataatga cccaccaatc acatgcctat catatagtaa	9240
aacccagccc atgaccccta acaggggccc tctcagccctt cctaatgacc tccggccctag	9300
ccatgtgatt tcacttccac tccataacgc tcctcataactt aggcctacta accaacacac	9360
taaccatata ccaatgatgg cgcatgtaa cacgagaaag cacataccaa ggccaccacaa	9420
caccacctgt ccaaaaaggc cttcgatacg ggataatcctt atttattacc tcagaagttt	9480
ttttcttcgc aggattttc tgagcctttt accactccag cctagccctt accccccaat	9540

taggaggggca	ctggcccca	acaggcatca	ccccgctaaa	tcccctagaa	gtcccactcc	9600
taaacacatc	cgtattactc	gcatcaggag	tatcaatcac	ctgagctcac	catagtctaa	9660
tagaaaacaa	ccgaaaccaa	ataattcaag	cactgcttat	tacaatttta	ctgggtctct	9720
attttaccct	cctacaagcc	tcagagtaact	tcgagctcc	cttcaccatt	tccgacggca	9780
tctacggctc	aacatttttt	gtagccacag	gcttccacgg	acttcacgtc	attattggct	9840
caactttcct	cactatctgc	ttcatccgcc	aactaatatt	tcactttaca	tccaaacatc	9900
actttggctt	cgaagccgcc	gcctgatact	ggcattttgt	agatgtggtt	tgactatttc	9960
tgtatgtctc	catctattga	tgagggtctt	actcttttag	tataaatagt	accgttaact	10020
tccaattaac	tagtttgac	aacattcaaa	aaagagtaat	aaacttcgccc	ttaattttaa	10080
taatcaacac	cctcctagcc	ttactactaa	taattattac	attttgacta	ccacaactca	10140
acggctacat	agaaaaatcc	acccttacg	agtgcggctt	cgaccctata	tccccggccc	10200
gcgtccctt	ctccataaaaa	ttcttcttag	tagctattac	cttcttatta	tttgatctag	10260
aaattgccc	ccttttaccc	ctaccatgag	ccctacaaac	aactaacctg	ccactaatag	10320
ttatgtcatc	cctcttatta	atcatcatcc	tagccctaag	tctggcctat	gagtgactac	10380
aaaaaggatt	agactgaacc	gaattggat	atagttaaa	caaaacgaat	gatttcgact	10440
cattaaatta	tgataatcat	attaccaaa	tgcccctcat	ttacataaat	attatactag	10500
catttaccat	ctcacttcta	ggaatactag	tatatcgctc	acacctcata	tcctccctac	10560
tatgcctaga	aggaataata	ctatcgctgt	tcattatagc	tactctcata	accctcaaca	10620
cccactccct	cttagccaat	attgtgccta	ttgccatact	agtctttgcc	gcctgcgaag	10680
cagcgggtgg	cctagcccta	ctagtctcaa	tctccaaacac	atatggccta	gactacgtac	10740
ataacctaaa	cctactccaa	tgctaaaact	aatcgcccc	acaatttat	tactaccact	10800
gacatgactt	tccaaaaaac	acataatttg	aatcaacaca	accacccaca	gcctaattat	10860
tagcatcatc	cctctactat	tttttaacca	aatcaacaac	aacctattta	gctgtcccc	10920
aacctttcc	tccgaccccc	taacaacccc	cctcctaata	ctaactacct	gactcctacc	10980
cctcacaatc	atggcaagcc	aacgccactt	atccagtgaa	ccactatcac	gaaaaaaact	11040
ctacctctct	atactaattct	ccctacaaat	ctccttaatt	ataacattca	cagccacaga	11100
actaatcata	ttttatatct	tcttcgaaac	cacacttac	cccaccttgg	ctatcatcac	11160
ccgatgaggc	aaccagccag	aacgcctgaa	cgcaggcaca	tacttcctat	tctacaccct	11220
agtaggctcc	cttcccctac	tcatcgact	aatttacact	cacaacaccc	taggctca	11280
aaacattctta	ctactca	tcactgccc	agaactatca	aactcctgag	ccaacaactt	11340
aatatgacta	gcttacacaa	tagttttat	agtaaagata	cctctttacg	gactccactt	11400
atgactccct	aaagcccatg	tcgaagcccc	catcgctggg	tcaatagtac	ttgccgcagt	11460
actcttaaaa	ctaggcggct	atggtataat	acgcctcaca	ctcattctca	accccctgac	11520
aaaacacata	gcctaccct	tccttgtact	atccctatga	ggcataatta	taacaagctc	11580
catctgccta	cgacaaacag	acctaaaatc	gctcattgca	tactcttcaa	tcagccacat	11640

2906064\_1.txt

agccctcgta gtaacagcca ttctcatcca aacccctga agttcacccg ggcgagtc	11700
tctcataatc gcccacgggc ttacatcctc attactattc tgcttagcaa actcaaacta	11760
cgaacgcact cacagtcgca tcataatcct ctctcaagga cttcaaactc tactcccact	11820
aatagcttt tgatgacttc tagcaagcct cgctaaccctc gccttacccc ccactattaa	11880
cctactggga gaactcttg tgcttagtaac cacgttctcc tgatcaaata tcactctcct	11940
acttacagga ctcaacatac tagtcacagc cctatactcc ctctacatat ttaccacaac	12000
acaatggggc tcactcaccc accacattaa caacataaaa ccctcattca cacgagaaaa	12060
caccctcatg ttcatacacc tatccccat tctcctccta tccctcaacc ccgacatcat	12120
taccgggttt tcctttgtta aatatagttt aaccaaaaaca tcagattgtg aatctgacaa	12180
cagaggctta cgaccctta ttaccgaga aagtcacaa gaactgctaa ctcatgcccc	12240
catgtctaac aacatggct tctcaacttt taaaggataa cagctatcca ttggcttag	12300
gccccaaaaa ttttggtgcactccaaata aaagtaataa ccatgcacac tactataacc	12360
accctaaccctgacttccct aattcccccc atccttacca ccctcgtaa ccctaacaaa	12420
aaaaactcat accccattatgtaaaatcc attgtcgcat ccacctttat tatcagtctc	12480
ttccccacaa caatattcat gtgcctagac caagaagttt ttagtcgaa ctgacactga	12540
gccacaaccc aaacaaccca gctccctta agttcaaactt cttccataata	12600
ttcatccctg tagcattgtt cgatccatgg tccatcatag aattctcact gtgatata	12660
aactcagacc caaacattaa tcagttcttc aaatatctac tcatcttccct aattaccata	12720
ctaattcttag ttaccgctaa caacctattc caactgttca tcggctgaga gggcgtagga	12780
attatattcct tcttgctcat cagttgtatgc tacgcccggag cagatgcca cacagcagcc	12840
attcaagcaa tcctatacaa ccgtatcgcc gatatcggtt tcatcctcgc ctttagcatga	12900
tttatcctac actccaactc atgagaccca caacaatag cccttctaaa cgtaatcca	12960
agcctcaccc cactactagg cctccctta gcagcagcag gcaaatcagc ccaatttaggt	13020
ctccacccct gactccctc agccatagaa ggccccaccc cagtcctcgc cttactccac	13080
tcaaggacta tagttgttagc aggaatcttc ttactcatcc gcttccaccc ctttagcagaa	13140
aatagccac taatccaaac tctaacacta tgcttaggcg ctatcaccac tctgtcgca	13200
gcagtctcgcc cccttacaca aaatgacatc aaaaaatcg tagcctctc cacttcaagt	13260
caactaggac tcataatagt tacaatcgcc atcaaccaac cacacctgc attcctgcac	13320
atctgtaccc acgccttctt caaagccata ctatttatgt gctccgggtc catcatccac	13380
aaccttaaca atgaacaaga tattcgaaaa ataggaggac tactcaaaac cataccctc	13440
acttcaacct ccctcaccat tggcagccta gcattagcag gaataccctt cttcacaggt	13500
ttctactcca aagaccacat catgaaacc gcaaacatata catacacaaa cgcctgagcc	13560
ctatctatta ctctcatcgctaccccttgc acaagcgccct atagcactcg aataattctt	13620
ctcaccctaa caggtcaacc tcgcttcccc acccttacta acattaacga aaataacccc	13680

accctactaa	accccattaa	acgcctggca	gccggaagcc	tattcgcagg	atttctcatt	13740
actaacaaca	tttccccgc	atcccccttc	caaacaacaa	tccccctcta	cctaaaactc	13800
acagccctcg	ctgtcacttt	cctaggactt	ctaacagccc	tagacctcaa	ctacctaacc	13860
aacaaactta	aaataaaatc	cccactatgc	acattttatt	tctccaacat	actcgattc	13920
taccctagca	tcacacacccg	cacaatcccc	tatctaggcc	ttcttacgag	ccaaaacctg	13980
cccctactcc	tcctagacct	aacctgacta	gaaaagctat	tacctaaaac	aatttcacag	14040
caccaaattct	ccacctccat	catcacctca	acccaaaaag	gcataattaa	actttacttc	14100
ctctctttct	tcttcccact	catcctaacc	ctactcctaa	tcacataacc	tattcccccg	14160
agcaatctca	attacaatat	atacaccaac	aaacaatgtt	caaccagtaa	ctactactaa	14220
tcaacgcccc	taatcataca	aagccccgc	accaatagga	tcctcccgaa	tcaaccctga	14280
cccctctcct	tcataaatta	ttcagttcc	tacactatta	aagtttacca	caaccaccac	14340
cccatcatac	tcttcaccc	acagcaccaa	tcctacctcc	atcgctaacc	ccactaaaac	14400
actcaccaag	acctcaaccc	ctgaccccca	tgcctcagga	tactcctcaa	tagccatcgc	14460
tgtgtatata	ccaaagacaa	ccatcattcc	ccctaaataa	attaaaaaaa	ctattaaacc	14520
catataacct	ccccaaaaat	tcagaataat	aacacacccg	accacaccgc	taacaatcaa	14580
tactaaaccc	ccataaatag	gagaaggctt	agaagaaaac	cccacaaacc	ccattactaa	14640
acccacactc	aacagaaaca	aagcatacat	cattattctc	gcacggacta	caaccacgac	14700
caatgatatg	aaaaaccatc	gttgtatttc	aactacaaga	acaccaatga	ccccaaatacg	14760
caaaactaac	cccctaataa	aattaattaa	ccactcattc	atcgacctcc	ccacccatc	14820
caacatctcc	gcatgatgaa	acttcggctc	actccttggc	gcctgcctga	tcctccaaat	14880
caccacagga	ctattcctag	ccatgcacta	ctcaccagac	gcctcaaccg	cctttcattc	14940
aatcgcccac	atcactcgag	acgtaaattta	tggctgaatc	atccgctacc	ttcacgccaa	15000
tggcgccctca	atattctta	tctgccttct	cctacacatc	gggcgaggcc	tatattacgg	15060
atcatttctc	tactcagaaa	cctgaaacat	cggcattatc	ctcctgcttg	caactatagc	15120
aacagccttc	ataggctatg	tcctcccggt	aggccaaata	tcattctgag	gggcccacagt	15180
aattacaaac	ttactatccg	ccatcccata	cattgggaca	gacctagttc	aatgaatctg	15240
aggaggctac	tcagtagaca	gtcccaccct	cacacgattc	tttacctttc	acttcattctt	15300
gcccttcatt	attgcagccc	tagcaacact	ccacccctta	ttcttgacag	aaacgggatc	15360
aaacaacccc	ctaggaatca	cctccattc	cgataaaatc	accttccacc	cttactacac	15420
aatcaaagac	gccctcggt	tacttcttct	ccttctctcc	ttaatgacat	taacactatt	15480
ctcaccagac	ctcctaggcg	acccagacaa	ttatacccta	gccaacccct	taaacacccc	15540
tccccacatc	aagcccgaat	gatatttcct	attgcctac	acaattctcc	gatccgtccc	15600
taacaaacta	ggaggcgtcc	ttgccctatt	actatccatc	ctcatcctag	caataatccc	15660
catcctccat	atatccaaac	aacaaagcat	aatattcgc	ccactaagcc	aatcacttta	15720
ttgactccta	gccgcagacc	tcctcattct	aacctgaatc	ggaggacaac	cagtaagcta	15780

2906064\_1.txt

ccctttacc atcattggac aagtagcatc cgtactatac ttcacaacaa tcctaattct	15840
aataccaaact atctccctaa ttgaaaacaa aataactcaa tgggcctgtc cttgttagtat	15900
aaactaatac accagtctg taaaccggag atgaaaacct ttttccaagg acaaattcaga	15960
gaaaaagtct ttaactccac cattagcacc caaagctaag attctaattt aaactattct	16020
ctgttcttc atgggaaagc agatttgggt accacccaag tattgactca cccatcaaca	16080
accgctatgt atttcgtaca ttactgccag ccaccatgaa tattgtacgg taccataaat	16140
acttgaccac ctgttagtaca taaaaaccca atccacatca aaacccctc cccatgctta	16200
caagcaagta cagcaatcaa ccctcaacta tcacacatca actgcaactc caaagccacc	16260
cctcaccac taggatacca acaaacctac ccaccctaa cagtacatag tacataaagc	16320
catttaccgt acatagcaca ttacagtcaa atcccttctc gtccccatgg atgacccccc	16380
tcagataggg gtcccttgcac caccatcctc cgtgaaatca atatccgca caagagtgct	16440
acttcctcg ctccgggccc ataacacttg gggtagcta aagtgaactg tatccgacat	16500
ctggttccta cttcagggtc ataaaagccta aatagcccac acgttcccct taaataagac	16560
atcacgatg	16569

<210> 2	
<211> 25	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> secuencia específica de molde para el amplicón D-loop	
<400> 2	
taggggaaaa ttgattattt ttttt	25

<210> 3	
<211> 30	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> secuencia específica de molde para el amplicón D-loop	
<400> 3	
acaaaacattc aattattattt attatatcct	30

<210> 4	
<211> 30	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> secuencia específica de molde para el amplicón ND1	
<400> 4	
atggtaatt ttttattttt tattgtat	30

<210> 5	
<211> 30	
<212> DNA	

2906064\_1.txt

<213> Artificial Sequence	
<220>	
<223> Secuencia específica de molde para el amplicón ND1	
<400> 5	
taatttaaat ttaatactca ccctaataa	30
<210> 6	
<211> 10	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Secuencia MID	
<400> 6	
acgagtgcgt	10
<210> 7	
<211> 10	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Secuencia MID	
<400> 7	
acgctcgaca	10
<210> 8	
<211> 10	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Secuencia MID	
<400> 8	
agacgcactc	10
<210> 9	
<211> 10	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Secuencia MID	
<400> 9	
agcactgtag	10
<210> 10	
<211> 10	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Secuencia MID	
<400> 10	
atcagacacg	10

2906064\_1.txt

<210> 11  
<211> 10  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Secuencia MID

<400> 11  
atatcgcgag 10

<210> 12  
<211> 10  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Secuencia MID

<400> 12  
cgtgtctcta 10

<210> 13  
<211> 10  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Secuencia MID

<400> 13  
ctcgcgtgta 10

<210> 14  
<211> 10  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Secuencia MID

<400> 14  
tagtatcagc 10

<210> 15  
<211> 10  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Secuencia MID

<400> 15  
tctctatgcg 10

<210> 16  
<211> 10  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Secuencia MID

<400> 16

tgatacgtct

<210> 17  
<211> 10  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Secuencia MID

<400> 17  
tactgagcta

<210> 18  
<211> 10  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Secuencia MID

<400> 18  
catagtagtg

<210> 19  
<211> 10  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Secuencia MID

<400> 19  
cgagagatac

<210> 20  
<211> 10  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Secuencia MID

<400> 20  
atacgacgta

<210> 21  
<211> 10  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Secuencia MID

<400> 21  
tcacgtacta

<210> 22  
<211> 10  
<212> DNA  
<213> Artificial Sequence

<220>

2906064\_1.txt

<223> Secuencia MID

<400> 22  
cgtctagtagc

10

<210> 23  
<211> 10  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Secuencia MID

<400> 23  
tctacgttagc

10

<210> 24  
<211> 10  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Secuencia MID

<400> 24  
tgtactactc

10

<210> 25  
<211> 10  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Secuencia MID

<400> 25  
acgactacag

10

<210> 26  
<211> 10  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> secuencia MID

<400> 26  
cgtagactag

10

<210> 27  
<211> 10  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Secuencia MID

<400> 27  
tacgagtagt

10

<210> 28  
<211> 10  
<212> DNA

## 2906064\_1.txt

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Secuencia MID

&lt;400&gt; 28

tactctcg

10

&lt;210&gt; 29

&lt;211&gt; 10

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Secuencia MID

&lt;400&gt; 29

tagagacgag

10

&lt;210&gt; 30

&lt;211&gt; 502

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 30

ataccaacaa acctacccac ccttaacagt acatagtaca taaagccatt taccgtacat

60

agcacattac agtcaaatcc cttctcgcc ccatggatga cccccctcag ataggggtcc

120

cttgaccacc atcctccgtg aaatcaatat cccgcacaag agtgctactc tcctcgctcc

180

gggcccataa cacttgggg tagctaaagt gaactgtatc cgacatctgg ttcctacttc

240

agggtcataa agcctaaata gcccacacgt tccccttaaa taagacatca cgatggatca

300

caggtctatc accctattaa ccactcacgg gagctctcca tgcatttggt atttcgtct

360

gggggtatg cacgcgatag cattgcgaga cgctggagcc ggagcaccct atgtcgact

420

atctgtcttt gattcctgcc tcattctatt atttatcgca cctacgttca atattacagg

480

cgaacatact tactaaagt tg 502

&lt;210&gt; 31

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; cebador directo para detectar el polimorfismo T16519C

&lt;400&gt; 31

ataccaacaa acctacccac cc

22

&lt;210&gt; 32

&lt;211&gt; 24

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; cebdor reverso para detectar el polimorfismo T16519C

&lt;400&gt; 32

ggcgaacata cttactaaag tgtg

24

2906064\_1.txt

<210> 33  
<211> 21  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Cebador A

<400> 33  
cgtatcgccct ccctcgcgcc a

21

<210> 34  
<211> 21  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Cebador B

<400> 34  
ctatgcgcct tgccagcccg c

21