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

#### (19) World Intellectual Property Organization

International Bureau





(10) International Publication Number WO 2012/076846 A1

(43) International Publication Date 14 June 2012 (14.06.2012)

(51) International Patent Classification: C12Q 1/68 (2006.01) **C07K 14/47** (2006.01) C07K 16/18 (2006.01) A61K 39/00 (2006.01) A01K 67/027 (2006.01)

(21) International Application Number:

PCT/GB2011/001696

(22) International Filing Date:

7 December 2011 (07.12.2011)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

1020982.3 10 December 2010 (10.12.2010)

GB

- (71) Applicant (for all designated States except US): THE UNIVERSITY OF NOTTINGHAM [GB/GB]; University Park, Nottingham NG7 2RD (GB).
- (72) Inventor; and
- Inventor/Applicant (for US only): VINKEMEIER, Uwe [DE/GB]; The University of Nottingham, University Park, Nottingham NG7 2RD (GB).
- Agent: O'NEILL, Michelle; Harrison IP Limited, Box Tree House, Northminster Business Park, Northfield Lane, York YO26 6QU (GB).

- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, 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, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, 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) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (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, ML, MR, NE, SN, TD, TG).

#### Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))

(54) Title: POLYMORPHISM

Figure 7 STAT1 - wild type STAT1 - E705Q Coiled Coll DNA Binding Linker SH2 TAD 576

(57) Abstract: The invention relates to a nucleic acid molecule comprising a nucleic acid sequence encoding a Signal Transducer Activator of Transcription 1 (STAT1) protein wherein the sequence comprises a polymorphism within a polymorphic region represented by nucleotides 2104 to 2115.





1

## **POLYMORPHISM**

#### Field of the Invention

The present invention relates to isolated nucleic acid sequences, vectors, host cells, screens and non-human transgenic animals. Also provided is a method for diagnosing, or determining predisposition to disease. In addition, the invention provides methods for preventing or treating disease, and kits for use in such methods.

#### Background to the Invention

The Signal Transducer and Activator of Transcription (STAT) proteins are best known for their role in cytokine signaling (Levy and Darnell (2002) Nat Rev. Mol. Cell Biol. 3, 651-662)). STAT1 is the main transcription factor involved in interferon (IFN)-dependent induction of antiviral and tumor-suppressing genes (Levy and Darnell, 2002 ibid). It is a nucleocytoplasmic shuttling protein that forms high-affinity homodimers involving N-domain interactions (Vinkemeier, (2004) J. Cell Biol. 167, 197-201; Wenta et al., (2008) Proc. Natl. Acad. Sci. USA 105, 9238-9243). Structural evidence indicates that such dimers are recruited to the IFNv receptor for subsequent activation by tyrosine phosphorylation (Mao et al., (2005) Mol. Cell 17, 761-771). Upon phosphorylation on C-terminal Tyr<sup>701</sup> an equally stable additional STAT1 homodimer conformation emerges where the phosphortyrosine of one monomer reaches into the SH2 domain of the second monomer (Chen et al., (1998) Cell 93, 827-839). Phosphortyrosine-SH2 domainmediated dimers, also called parallel dimers, possess DNA-binding activity and evoke the transcriptional responses to IFNy. Moreover, activation blocks STAT nuclear export, which is discernable in the light microscope as transient accumulation (Meyer et al., (2003) Genes Dev. 17, 1992-2005). In the nucleus STAT1 is highly mobile and distributes homogeneously outside of the nucleoli (Lillemeier et al., (2001) EMBO J. 20, 2508-2517).

The biological actions of the STAT proteins are dependent on adequate activation and reliable inactivation. A number of molecules have been implicated in the negative regulation of STAT1, among them ligation of a ~100 residue

protein termed SUMO (Small Ubiquitin-related Modifier). SUMO modifications have been shown to alter the interactions of proteins or to direct changes in their subcellular distribution (Geiss-Friedlander and Melchior, (2007) Nat. Rev. Mol. Cell Biol. 8, 947-956). In contrast to the other STATs, STAT1 harbors a functional sumoylation consensus sequence (ΨΚxΕ; where Ψ is a large hydrophobic residue and x any residue comprising residues 702-705 with the SUMO acceptor lysine<sup>703</sup> situated just downstream of the phosphorylation site (Rogers et al., (2003) J. Biol. Chem. 278, 30091-30097). Yet, with highly efficient SUMO-deconjugation, only 1-2% of STAT1 are modified at steady state (M.D., U.V., unpublished observation). This and the only modest increase in Tyr phosphorylation and transcriptional activity caused by lack of SUMO modification lead some to conclude that it played a negligible role for STATs (Song et al., (2006) Blood 108, 3237-3244). However, this sentiment is disputed (Ungureanu et al., (2005) Blood 106, 224-226; Vanhatupa et al., (2008) Biochem. J. 409, 179-185) since the mechanism of STAT1 inhibition is not understood, and because very small modified fractions usually suffice to maximally affect SUMO targets.

The present inventors have identified a genetic polymorphism within a consensus sequence of a STAT1 protein where a SUMO protein binds which is shown to, surprisingly, enhance phosphorylation of the proximal tyrosine and this is associated with an increased level of STAT1 activity in response to IFNy and can be used as a diagnostic tool. In addition, the present inventors have found that an increased level of STAT1 activity in response to IFNy in an animal is associated with phenotypic and behavioural changes in the animal which could not have been predicted.

#### Statements of Invention

According to a first aspect of the invention there is provided an isolated nucleic acid molecule comprising a nucleic acid sequence encoding a Signal Transducer Activator of Transcription 1 (STAT1) protein wherein the sequence comprises a polymorphism within a polymorphic region represented by nucleotides 2104 to 2115.

The nucleic acid molecules of the invention may be DNA, e.g. cDNA, RNA, and single or double stranded sequences. Preferably the nucleic acid molecule is cDNA. All the molecules of the present invention are isolated, or alternatively may be recombinant. By isolated is meant a nucleic acid molecule which has been purified, and is substantially free of protein and other nucleic acid. Such molecules may be obtained by PCR amplification, cloning techniques, or synthesis on a synthesiser. By recombinant is meant nucleic acid molecules which have been recombined by the hand of man.

In a preferred aspect of the invention the polymorphism is within a codon 705 of STAT1. Preferably the polymorphism is within a polymorphic region represented by nucleotides 2113 to 2115 of STAT1.

The term "polymorphism" refers to the coexistence of multiple forms of a sequence. Thus, a polymorphic site is the location at which sequence divergence occurs. The different forms of the sequence which exist as a result of the presence of a polymorphism are referred to as "alleles". The region comprising a polymorphic site may be referred to as a polymorphic region.

Examples of the ways in which polymorphisms are manifested include restriction fragment length polymorphisms (Bostein et al Am J Hum Genet 32 314-331 (1980)), variable number of tandem repeats, hypervariable regions, minisatellites, di-or multi- nucleotide repeats, insertion elements and nucleotide or amino acid deletions, additions or substitutions. A polymorphic site may be as small as one base pair, which may alter a codon thus resulting in a change in the encoded amino acid sequence.

The polymorphism may be a knock-out or knock-in mutation. Preferably the polymorphism is a knock-in mutation.

In a preferred aspect of the invention the polymorphism is a single nucleotide polymorphism. Single nucleotide polymorphisms arise due to the substitution, deletion or insertion of a nucleotide residue at a polymorphic site. Such variations

are referred to as SNPs. SNPs may occur in protein coding regions, in which case different polymorphic forms of the sequence may give rise to variant protein sequences. Other SNPs may occur in non-coding regions. In either case, SNPs may result in defective proteins or regulation of genes, thus resulting in disease. Other SNPs may have no direct phenotypic effects, but may show linkage to disease states, thus serving as markers for disease. SNPs typically occur more frequently throughout the genome than other forms of polymorphism discussed above, and there is therefore a greater probability of finding a SNP associated with a particular disease state.

In one embodiment of the invention the SNP is at nucleotide 2113 of the STAT1 nucleic acid sequence. Preferably the nucleotide at position 2113 of the STAT1 cDNA sequence is replaced by a C.

Preferably the STAT 1 protein is a mammalian STAT1 protein for example primate, canine, feline, ovine or rodent e.g. rat, mouse or rabbit. Preferably the STAT1 protein is a human STAT1 protein. Alternatively the STAT1 protein may be a rodent STAT1 protein for example a mouse STAT1 protein.

In one embodiment of the invention, the STAT1 protein is a human STAT1 protein. Thus the invention provides an isolated nucleic acid molecule comprising a nucleic acid sequence selected from:

- i) a nucleic acid sequence as shown in Figure 1 wherein the sequence comprises a polymorphism at one or more of nucleotides 2104 to 2115 of Figure 1; and
- ii) a nucleic acid sequence which hybridises under stringent conditions to a sequence shown in (i).

In an alternative embodiment of the invention the STAT1 protein is a mouse STAT1 protein. Thus the invention provides an isolated nucleic acid molecule comprising a nucleic acid sequence selected from:

5

i) a nucleic acid sequence as shown in Figure 2 wherein the sequence comprises a polymorphism at one or more of nucleotides 2104 to 2115 of Figure 2; and

ii) a nucleic acid sequence which hybridises under stringent conditions to a sequence shown in (i).

Hybridization of a nucleic acid molecule occurs when two complementary nucleic acid molecules undergo an amount of hydrogen bonding to each other. The stringency of hybridization can vary according to the environmental conditions surrounding the nucleic acids, the nature of the hybridization method, and the composition and length of the nucleic acid molecules used. Calculations regarding hybridization conditions required for attaining particular degrees of stringency are discussed in Sambrook et al., Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001); and Tijssen, Laboratory Techniques in Biochemistry and Molecular Biology—Hybridization with Nucleic Acid Probes Part I, Chapter 2 (Elsevier, New York, 1993). The  $T_{\rm m}$  is the temperature at which 50% of a given strand of a nucleic acid molecule is hybridized to its complementary strand. The following is an exemplary set of hybridization conditions and is not limiting:

# Very High Stringency (allows sequences that share at least 90% identity to hybridize)

Hybridization:

5xSSC at 65°C for 16 hours

Wash twice:

2xSSC at room temperature (RT) for 15 minutes each

Wash twice:

0.5xSSC at 65°C for 20 minutes each

#### High Stringency (allows sequences that share at least 80% identity to hybridize)

Hybridization:

5x-6x SSC at 65°C-70°C for 16-20 hours

Wash twice:

2xSSC at RT for 5-20 minutes each

Wash twice:

1xSSC at 55°C-70°C for 30 minutes each

#### Low Stringency (allows sequences that share at least 50% identity to hybridize)

Hybridization:

6xSSC at RT to 55°C for 16-20 hours

6

Wash at least twice: 2x-3xSSC at RT to 55°C for 20-30 minutes each.

The nucleic acid sequence of the invention may comprise the sequence set out in Figure 1 or a sequence which is at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, for example 98%, or 99%, identical to the nucleic acid sequence set out in Figure 1 at the nucleic acid residue level.

"Identity", as known in the art, is the relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, identity also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. Identity can be readily calculated (Computational Molecular Biology, Lesk, A.M. ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D.W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A.M., AND Griffin, H.G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje. G., Academic Press, 1987; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991). While there exist a number of methods to measure identity between two polynucleotide or two polypeptide sequences, the term is well-known to skilled artisans (Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., SIAM J. Applied Math., 48: 1073 (1988). Methods commonly employed to determine identity between sequences include, but are not limited to those disclosed in Carillo, H., and Lipman, D., SIAM J. Applied Math., 48: 1073 (1988). Preferred methods to determine identity are designed to give the largest match between the sequences tested. Methods to determine identity are codified in computer programs. Preferred computer program methods to determine identity between two sequences include, but are not limited to BLASTP, BLASTN, and FASTA (Atschul, S.F. et al., J. Molec. Biol. 215: 403 (1990)).

In a further aspect, the invention provides a STAT 1 polypeptide wherein at least one of residues 702 to 705 inclusive is modified by addition, deletion or substitution. The STAT1 polypeptide may be isolated, recombinant or in substantially pure form.

As used herein, the term "polypeptide" means, in general terms, a plurality of amino acid residues joined together by peptide bonds. It is used interchangeably and means the same as peptide and protein.

Preferably at least one of residues 703 and/or 705 is modified. The residues may be modified by substitution with a different residue, for example reside 703 may be modified by substitution with an arginine residue.

In a preferred aspect of the invention, residue 705 is modified. Residue 705 may be replaced by a residue selected from glutamine, asparagine and alanine. Preferably residue 705 is replaced by glutamine or asparagine. Preferably still residue 705 is replaced by glutamine.

In one embodiment of the invention, the STAT1 polypeptide is a human STAT1 polypeptide. Thus the invention provides a STAT1 polypeptide comprising an amino acid sequence as shown in Figure 3 wherein at least one of residues 702 to 705 inclusive is modified by addition, deletion or substitution.

In an alternative embodiment of the invention, the STAT1 polypeptide is a mouse STAT1 polypeptide. Thus the invention provides a STAT1 polypeptide comprising an amino acid sequence as shown in Figure 4 wherein at least one of residues 702 to 705 inclusive is modified by addition, deletion or substitution.

A further aspect of the invention provides a vector comprising an isolated nucleic molecule according to the invention.

The vector may be a plasmid, cosmid, phage or virus based vector. The vector may include a transcription control sequence (promoter sequence) which

mediates cell specific expression, for example, a cell specific, inducible or constitutive promoter sequence. The vector may be an expression vector adapted for prokaryotic or eukaryotic gene expression, for example, the vector may include one or more selectable markers and/or autonomous replication sequences which facilitate the maintenance of the vector in either a eukaryotic cell or prokaryotic host (Sambrook et al (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory, Cold Spring Harbour, NY and references therein; Marston, F (1987) DNA Cloning Techniques: A Practical Approach Vol III IRL Press, Oxford UK; DNA Cloning: F M Ausubel et al, Current Protocols in Molecular Biology, John Wiley & Sons, Inc.(1994). Vectors which are maintained autonomously are referred to as episomal vectors.

The vector of the second aspect of the invention may include a transcription termination or polyadenylation sequences. This may also include an internal ribosome entry sites (IRES) or Shine-Dalgamo (bacterial ribosome binding) site. The vector may include a nucleic acid sequence that is arranged in a bicistronic or multi-cistronic expression cassette.

In a further aspect, there is provided a host cell comprising an isolated nucleic acid molecule according to invention. The nucleic acid molecule may be provided in a vector. The host cell may comprise an expression vector, or naked DNA encoding the nucleic acid molecules of the invention. A wide variety of suitable host cells are available, both eukaryotic and prokaryotic. Examples include bacteria such as E. coli, yeast, filamentous fungì, insect cells, mammalian cells, preferably immortalised, such as mouse, CHO, HeLa, myeloma, neuronal cell lines or Jurkat cell lines, human and monkey cell lines and derivatives thereof.

The host cells are preferably capable of expression of the nucleic acid sequence to produce a gene product (i. e. RNA or protein). Such host cells are useful in drug screening systems to identify agents for use in the diagnosis or treatment of subjects having, or being susceptible to diseases as defined herein.

The method by which said nucleic acid molecules are introduced into a host cell will usually depend upon the nature of both the vector/DNA and the target cell, and will include those known to a person skilled in the art. Suitable known methods include but are not limited to fusion, conjugation, liposomes, immunoliposomes, lipofectin, transfection, transduction, electroporation or injection, as described in Sambrook et al.

A further aspect of the invention provides a transgenic non-human animal, or progeny thereof, comprising a vector or isolated nucleic acid molecule according to the invention.

The transgenic non-human animal provides a model for use in the screening and development of candidate drugs for example in the treatment of neurological diseases/disorders, in particular depression, and inflammatory diseases. The transgenic non-human animal may be a mammal for example primate, canine, feline, ovine or rodent e.g. rat, mouse or rabbit.

Preferably the transgenic non-human animal is a mouse. The transgenic mouse may have an γIFN blood level higher than the corresponding wild type mouse.

In a preferred transgenic non-human animal, the isolated nucleic molecule comprises a nucleic acid sequence selected from:

- i) a nucleic acid sequence as shown in Figure 2 wherein the sequence comprises a polymorphism at one or more of nucleotides 2104 to 2115 of Figure 2; and
- ii) a nucleic acid sequence which hybridises under stringent conditions to a sequence shown in (i).

The invention also provides a knock-in non-human animal, or progeny thereof, comprising a vector or isolated nucleic acid molecule according to the invention.

Expression of an isolated nucleic acid molecule of the invention in a transgenic non-human animal is usually achieved by operably linking the nucleic acid

molecule to a promoter and/or enhancer sequence, preferably to produce a vector of the invention, and introducing this into an embryonic stem cell of a host animal by microinjection techniques (Hogan et al., A laboratory Manual, Cold Spring Harbour and Capecchi, Science (1989) 244: 1288-1292). The transgene construct should then undergo homologous recombination with the endogenous gene of the host. Those embryonic stem cells comprising the desired nucleic acid molecule may be selected, usually by monitoring expression of a marker gene, and used to generate a non-human transgenic animal.

Thus, the invention provides a method of preparing a transgenic non-human animal according to the invention comprising the steps of:

- incorporating an isolated nucleic acid molecule, or vector, according to the invention into non-human embryonic stem cells;
- ii) transferring the embryonic stem cells to a recipient female nonhuman animal; and
- iii) growing the embryonic stem cells into a mature transgenic nonhuman animal.

In a further aspect of the invention there is provided a kit for the diagnosis of disease or predisposition to disease, comprising a means for determining the presence or absence of a polymorphism at one or more of nucleotides 2104 to 2115 of a nucleic acid sequence encoding STAT1 wherein the polymorphism is diagnostic of disease or predisposition to disease.

Preferably the nucleic acid sequence encoding STAT1 is as represented in Figure 1.

Preferably the kit will comprise the components necessary to determine the presence or absence of a risk allele, in accordance with the methods described below. Such components include PCR primers and/or probes, PCR enzymes, restriction enzymes, and DNA or RNA purification means. Preferably, the kit will contain at least one pair of primers, or probes. The primers are preferably allele specific primers. Other components include labelling means, buffers for the

reactions. In addition, a control nucleic acid sample may be included, which comprises a wild type or variant nucleic acid sequence as defined above, or a PCR product of the same. The kit will usually also comprise instructions for carrying out the diagnostic method, and a key detailing the correlation between the results and the likelihood of disease. The kit may also comprise an agent for the prevention or treatment of disease.

In a preferred aspect of the invention, the kit comprises one or more containers comprising a pair of primers, wherein one of the primers within the pair is capable of hybridizing directly to, or adjacent to, the nucleic acid sequence represented by nucleotides 2104 to 2115 of STAT1 (e.g. as shown in Figure 1). For example, the kit may comprise one or more of the pairs of primers represented in Figure 5(a) and (b).

In a further aspect, the invention provide a method for the diagnosis of a polymorphism in a subject comprising

- i) obtaining a nucleic acid sample obtained from the subject; and
- ii) determining the presence or absence of polymorphism within a polymorphic region represented by nucleotides 2104 to 2115 of a nucleic acid sequence encoding STAT1.

The method of the invention as defined in step (ii) may include those methods known to persons skilled in the art for identifying differences between nucleic acid sequences (in diseased and non-diseased individuals), for example direct probing, allele specific hybridisation, polymerase chain reaction (PCR) methodology including Pyrosequencing (Ahmadian A, Gharizadeh B, Gustafsson AC, Sterky F, Nyren P, Uhlen M, Lundeberg J. Single-nucleotide polymorphism analysis by pyrosequencing, Anal Biochem. 2000 Apr 10; 280 (1): 103-10; Nordstrom T, Ronaghi M, Forsberg L, de Faire U, Morgenstern R, Nyren P. Direct analysis of single-nucleotide polymorphism on double-stranded DNA by pyrosequencing. Biotechnol Appl Biochem. 2000 Apr; 31 (Pt 2): 107-12) Allele Specific Amplification (ASA) (W093/22456), Allele Specific Hybridisation, single base extension (US patent No. 4,656,127), ARMS-PCR, Tagman (US 4683202;

4683195; and 4965188), oligo ligation assays, single-strand conformational analysis ((SSCP) Orita et al PNAS 86 2766-2770 (1989)), Genetic Bit Analysis (WO 92/15712) and RFLP direct sequencing, mass-spectrometry (MALDI- TOF) and DNA arrays. The appropriate restriction enzyme, will, of course, be dependent upon the polymorphism and restriction site, and will include those known to persons skilled in the art. Analysis of the digested fragments may be performed using any method in the art, for example gel analysis, or southern blots.

Preferably the method according to the invention comprises the steps of:

- i) obtaining a nucleic acid sample from a subject;
- ii) amplifying said isolated nucleic acid (e.g. DNA in particular cDNA) by PCR with primers that flank the polymorphism or polymorphic region to produce amplified nucleic acid regions; and
- iii) sequencing said amplified regions to determine the presence or absence of said polymorphism.

The methods of the invention are preferably carried out on a sample removed from a subject. Any biological sample comprising cells containing nucleic acid is suitable for this purpose. Examples of suitable samples include whole blood, leukocytes, semen, saliva, tears, buccal, skin or hair. For analysis of cDNA, mRNA or protein, the sample must come from a tissue in which the STAT1 is expressed. Blood is a readily accessible sample. Thus, the methods of the invention typically include the steps of obtaining a sample from a subject, and preparing nucleic acid from the sample.

The subject is preferably a mammal, and more preferably a human. Alternatively, the sample may be obtained from the subject prepartum e. g. by amniocentesis, or post mortum brain tissue.

Preferably the method of the invention provides for the diagnosis of a polymorphism in a subject comprising

i) obtaining a nucleic acid sample obtained from the subject; and

ii) determining the presence or absence of polymorphism within a polymorphic region represented by nucleotides 2104 to 2115 of a nucleic acid sequence encoding STAT1 wherein the presence of said polymorphism is indicative of disease or susceptibility to disease.

As used herein the term "disease" includes diseases and disorders including neurological, including neuropsychiatric disorders, inflammatory diseases and infertility. Neurological disorders may include, but are not limited to, including Parkinson's Disease, Attention Deficit Hyperactivity Disorder (ADHD), depression (bipolar disorder) including fatigue, schizophrenia, behavioural disorders such as obsessive compulsive disorders, anxiety and addiction (e.g. drug addiction such as alcohol, nicotine or cocaine addiction). Preferably the disease/disorder is depression. Inflammatory diseases include skin lesions, atopy, asthma, autoimmune diseases and including inflammatory bowel disease, such as ulcerative colitis and Crohn's disease, inflammatory joint disease, such as rheumatoid arthritis and ankylosing spondylitis, and multiple sclerosis.

Alternatively the method of the invention may provide for the diagnosis of a polymorphism in a subject comprising

- i) obtaining a nucleic acid sample obtained from the subject; and
- ii) determining the presence or absence of polymorphism within a polymorphic region represented by nucleotides 2104 to 2115 of a nucleic acid sequence encoding STAT1 wherein the presence of said polymorphism is indicative of a reduced susceptibility, or tolerance, to disease. Where the presence of said polymorphism is indicative, or diagnostic, of a reduced susceptibility, or tolerance, to disease, the term "disease" may include cancer and tumours, viral diseases, auto immune diseases including Hepatitis C and multiple sclerosis, microbial infections for example bacterial infections.

Preferably the method comprises the steps of:

i) obtaining a nucleic acid sample from the subject;

ii) amplifying said isolated nucleic acid (e.g. DNA in particular cDNA) by PCR with primers that flank the polymorphism or polymorphic region to produce amplified nucleic acid regions; and

iii) sequencing said amplified regions to determine the presence or absence of said polymorphism.

The invention further provides a method for diagnosing disease or predisposition to disease, comprising determining the presence or absence of a polymorphism at one or more of nucleotides 2104 to 2115 of a nucleic acid sequence encoding STAT1, wherein the presence of the polymorphism is diagnostic of disease or predisposition to disease.

The present invention is advantageous in that it facilitates the accurate diagnosis of disease, or the determination of predisposition to disease. Thus, by genotyping, an individual may be identified as having or being predisposed to disease. This helps to identify those individuals who are likely to respond positively to particular treatments or preventative measures. Thus, more effective therapies or preventative measures can be administered.

The invention further provides a method of diagnosing disease or predisposition to disease, comprising determining the presence or absence of a polymorphism at one or more of nucleotides 2104 to 2115 of a nucleic acid sequence encoding STAT1, wherein the presence of the polymorphism is diagnostic of a reduced susceptibility, or tolerance, to disease.

In a further aspect of the invention there is provided a method of treating or preventing disease in a subject, the method comprising determining the presence or absence of a polymorphism at one or more of nucleotides 2104 to 2115 of a nucleic acid sequence encoding STAT1; and if the polymorphism is present, administering treatment in order to prevent, delay or reduce the disease.

In a yet further aspect the invention provides a method of identifying an agent for the treatment of disease, comprising (a) administration of an agent to animal tissue comprising an isolated nucleic acid molecule according to the invention; and (b) determining whether the agent modulates the effect(s) of the polymorphism.

The effects of the polymorphism may include downstream effects such as increased sensitivity to IFN and the presence of neurological and inflammatory disorders.

The invention also provides for the use of an isolated nucleic acid molecule according to the invention to screen for agents which are capable of modulating the effect(s) of the polymorphism.

Putative agents will include those known to the person skilled in the art, and include chemical or biological compounds, sense or anti-sense nucleic acid sequences, binding proteins including antibodies, kinases and any other gene or gene product agonist or antagonist. Preferably the agent will be capable of modulating the effects of the polymorphism. Most preferably, the agent is capable of ameliorating the deleterious effects of the polymorphism. Preferably the agent is in inhibitor.

Such agents may be suitable for either prophylactic administration or after disease has been diagnosed. The invention also provides for agents to be administered either as DNA or RNA and thus as a form of gene therapy.

A further aspect of the invention provides an agent as described above. The agent may be an antisense sequence that is capable of hybridizing to a nucleic acid molecule of the invention under stringent conditions and these form part of the invention. Preferred antisense sequences are those which are capable of hybridising to an allele of the polymorphism of an isolated nucleic acid molecule of the invention and, most preferably, is capable of distinguishing between alleles of a polymorphism.

WO 2012/076846

The agent may be an antibody that binds to a polymorphic region of a STAT1 polypeptide wherein the polymorphic region is represented by residues 702 to 705 of a STAT1 polypeptide inclusive. Thus, the invention provides an antibody specifically recognising, and which binds to, a polymorphic region of a STAT1 polypeptide wherein the polymorphic region is represented by residues 702 to 705 of a STAT1 polypeptide.

The agent of the invention may be for use in preventing or treating neurological or inflammatory diseases such as described herein.

The suppression may be therapeutically helpful where a pro-inflammatory environment is actually supportive to counter disease, for example, tumour-infiltrating macrophages, which support tumour growth and invasiveness. In the tumour environment these macrophages are reprogrammed to abandon their pro-inflammatory phenotype. Thus, the reprogramming of macrophages may be inhibited by blocking STAT1 sumoylation, making the macrophages turn on the tumour rather than supporting it. Thus we search for inhibitors that block STAT1 sumoylation and examine their consequences for tumour development.

The invention, therefore, further provides the use of a STAT1 polypeptide in the identification of agents that block the binding of a regulatory protein to residue 703 of the STAT1 polypeptide. The regulatory protein may be a Small Ubiquitin-related Modifier (SUMO) protein which binds to STAT1 and regulated phosphorylation of residue 701 (tyrosine) of STAT1. The agent may be any agent capable of binding to one or more of residues 702 to 705 of STAT1 thereby hindering the binding of the regulatory protein, specifically SUMO, to residue 703 of the STAT1 polypeptide. The agent may be a chemical or biological compound, inhibitor, sense or anti-sense nucleic acid sequence, binding protein including antibodies, kinases and any other gene or gene product agonist or antagonist. The agent may be an antibody, for example an antagonistic antibody, which binds one or more of residues 702 to 705 of a STAT1 polypeptide inclusive. The agent may be for use in preventing or treating

cancer, cancer and tumours, viral diseases, auto immune diseases including Hepatitis C and multiple sclerosis, microbial infections for example bacterial infections.

The methods and uses according to the invention may also be carried out on a STAT 1 polypeptide wherein at least one of residues 702 to 705 inclusive is modified by addition, deletion or substitution. Thus the methods and uses of the invention may include the steps of determining the presence or absence of a polymorphism at one or more of residues 702 to 705 of a STAT1 polypeptide. This step may represent an alternative to determining the presence or absence of a polymorphism at one or more of nucleotides 2104 to 2115 of a nucleic acid sequence encoding STAT1.

Throughout the description and claims of this specification, the words "comprise" and "contain" and variations of the words, for example "comprising" and "comprises", means "including but not limited to", and is not intended to (and does not) exclude other moieties, additives, components, integers or steps.

Throughout the description and claims of this specification, the singular encompasses the plural unless the context otherwise requires. In particular, where the indefinite article is used, the specification is to be understood as contemplating plurality as well as singularity, unless the context requires otherwise.

Features, integers, characteristics, compounds, chemical moieties or groups described in conjunction with a particular aspect, embodiment or example of the invention are to be understood to be applicable to any other aspect, embodiment or example described herein unless incompatible therewith.

The invention will now be defined by way of example only with reference to the following figures in which:

Figure 1 cDNA sequence of human STAT 1;

WO 2012/076846

Figure 2 cDNA sequence of mouse STAT1;

Figure 3 amino acid sequence of human STAT 1 (genbank accession number NP009330);

Figure 4 amino acid sequence of mouse STAT 1;

Figure 5 sequences of the PCR primers used to generate 5' and 3' homology fragments respectively: (a) 5' homologues and (b) 3 homologues;

Figure 6 Schematic drawing depicting the exon/ intron structure in the 3'-region of the mouse STAT1 gene (top). Also shown is a relevant region of the ES cell targeting vector with the neomycine resistance gene and selected restriction sites (bottom). The exon containing codon 705 is indicated (\*). UTR, untranslated region.

Figure 7 Schematic drawing depicting the functional domains of the STAT1 protein. Approximate domain boundaries are numbered according to human STAT1. The positions of tyrosine 701 (black background), and the SUMO consensus with the SUMO conjugation site lysine 703 (boxed), and the glutamine 705 mutation (grey background) are highlighted. Coiled coil, coiled coil domain; DNA binding, DNA binding domain; linker, linker domain; SH2, src-homology 2 domain; TAD, transcription activation domain.

Figure 8 Comparison of STAT1 protein sequences from the indicated species showing the evolutionary conservation of both the location and the sequence of the SUMO targeting site. The tyrosine phosphorylation site is shown with a black background. The sequence (Y+1 to +4) is shown with a grey background if it conforms to the SUMO consensus; residues within this region that divert from the consensus and thus presumably render it non-functional are shown with a white background. Sequences containing a presumably functional SUMO consensus are boxed.

Figure 9 photographs of 9 week old littermates (a) homozygous and (b) heterozygous for SUMO-deficiency of STAT 1.

#### **EXAMPLE**

# Generation of mutant STAT1 knock-in mice

To generate STAT1<sup>SUMO-deficient</sup> knock-in mice carrying the mutation glutamate<sup>705</sup>glutamine, a genomic DNA fragment containing exons 17 to 22 was isolated from a mouse genomic P1 artificial chromosome library (RCPI-21 209N6, Deutsches Ressourcenzentrum für Genomforschung, Germany). The targeting vector was constructed from a 4.7-kb PCR-amplified fragment encompassing exons 17 to 21 (with exon 19 containing codon 705) as 5' homologue, and a 2.1-kb fragment including exon 22 as 3' homologue as follows:

Two genomic fragments were retrieved from a PAC clone by PCR.

The 5' homology fragment was amplified by PCT using the primer pair 5'-aaaactcgagggcacctgcacagggt and 5'-aaaactcgaggctcaccctggacaga.

The 3' homology fragment was amplified by PCR using the primer pair 5'-attttaggtacccgagagtccaccacctcagg and 5'-attttaggtaccggacagagcacatcgtgg.

The 5' homology fragment was subsequently cloned into the Xhol site of vector Bluescript (pBS), followed by site-directed mutagenesis to mutate the codon 705 from Glu (gag) to Gln (cag). The following primer pair was used for mutagenesis. 5'-ggatacatcaagacctCAGttgatttctgtgtctg and the reverse complementary oligo 5'-cagacacagaaatcaaCTGagttcttgatgtatcc. The mutated codon is in capital letters.

Next, the mutated 5' homology fragment was excised from pBS with Xhol and BamHI, and cloned into these sites of pPNT-frt3. The 3' homology fragment was inserted into the KpnI site of pPNT-frt3 to generate the gene targeting vector construct.

Embryonic stem cells (129P2/OlaHsd-derived E14, subclone 14.1) were electroporated with Notl-linearized target vector and grown under double selection with 360 μg of G418 per ml, and 2 μM ganciclovir. Two ES cell clones carrying the desired mutation were identified by Southern blotting due to the presence of an additional 4.7-kb Pstl fragment. One clone was used for injection into C57BL/6 blastocysts to generate germ line chimeras (Max Planck Institute for Molecular Cell Biology and Genetics, Dresden, Germany). Successful germ line transmission was confirmed by coat color. The chimeras were mated with Flip delitas mice (C57BL/6 genetic background) to excise the inserted Neo cassette. Subsequently the Neo-deleted heterozygous offspring was intercrossed under standard pathogen-free conditions to obtain mice homozygous for the Q705 locus.

#### Analysis of knock-in mice

The knock-in approach chosen resulted in animals with C57/BI6 genetic background expressing exclusively mutant STAT1 under control of the endogenous promoter. Glutamate705 in the SUMO consensus sequence (aa 702-705) was mutated to glutamine, which expectedly precluded SUMO addition in cells and in vitro sumoylation assays (data not shown). Mice homozygous for lack of STAT1 sumoylation (E705Q) were born at the expected mendelian rate, and grew to adulthood without overt growth abnormalities. By the age of 6-9 weeks, however, E705Q mice started to develop areas with hairless skin and lesions on their snout regions (Fig. 9a), which did not scar over and progressively grew larger, but only affected the facial region. This phenomenon was nearly 100% penetrant (5/6).

With wild type (WT) or heterozygous (HT) littermates similar observations were not made even when housed in the same cage from birth (Fig. 9b), thus excluding the possibility that lesions were caused by allogrooming or were the result of aggressive encounters with other E705Q mice. In fact E705Q animals did not appear aggressive, but were observed to frequently engage in facial self-grooming. Moreover, in contrast to HT and WT animals, the E705Q mice increasingly displayed poor health and died prematurely at 8-10 months of age.

In addition we did not obtain offspring from pairings of E705Q mice, whereas HT's bred successfully. The cause for this failure is presently unknown but it may not be infertility since we obtained offspring from one pairing of E705Q (male) x HT (female). In summary these observation raise the possibility that E705Q mice have excessive and injurious self-grooming behaviour. Since the animals are raised and kept in isolator cages and are free of common microbial pathogens, the abnormal behaviour and early mortality are unlikely to result from an infectious agent. To our best knowledge this is the first report associating dysregulated JAK-STAT-signalling with behavioural abnormalities in mice.

Despite the relatively modest inhibitory effects of SUMO modification on STAT1, the observed phenotype of mice expressing STAT1 devoid of sumoylation is not only unexpected but also surprisingly severe.

We are quantifying the overgrooming behaviour and other physiologic parameters (e.g. body temperature, locomotor activity) by camera monitoring and radiotelemetric analysis. Skin examinations of the affected facial regions is being performed before and after the skin alterations are obvious to examine possible cutaneous defects, such as inflammation or innervation abnormalities as a cause for the hair loss and lesions. Since our mice display excessive levels of self-grooming, a phenotype reminiscent of compulsive behaviours, we will consider whether their phenotype further resembles obsessive-compulsive disorders. In addition we are testing whether drugs used to treat depression, anxiety or obsessive-compulsory behaviour (fluoxetine, riboxetine) are effective in correcting the abnormal behaviour in STAT1 sumo-minus mice.

# Brain-resident macrophages are activated in mice expressing SUMO-free STAT1

Mice expressing SUMO-free STAT1, among other things, display a neurological phenotype, i.e. defective short term memory. Our published work (Blood 2011) has demonstrated hyperactivity and increased cytotoxicity of peripheral macrophages derived from these mutant mice. Macrophages are important mediators of inflammation, which has been recognized as a major contributor to neurological disorders. However, it has not been reported to date whether the

activation state of macrophages residing in the brain —cells which are called microglia—is altered in mice expressing SUMO-free STAT1, too.

To find out, we examined the expression of the protein Iba1 (ionized calcium binding adaptor molecule 1) in the brain. Iba1 is specifically expressed in macrophages/microglia and is upregulated during the activation of these cells. This protein therefore is a marker for the abundance and the activation of microglia.

We have perfused the brains of 12 weeks old wild type (labelled WT in the images) and mutant mice (labelled hm, that is homozyguous for SUMO-free STAT1), and performed immunohistochemistry on paraffin-embedded brain sections using an Iba1-specific antibody. Four brain regions were investigated; the hippocampus (hip) the cerebellum (cere), and two regions of the cortex. The chromogenic detection used for Iba1 makes microglia appear dark brown in the images provided below. Additionally, the sections have been counterstained with Hoechst dye to reveal the cell nuclei (blue colour). The data is not shown. All brain regions tested show microglia, irrespective of the genotype. The cells show the typical long thin processes emanating from the cell body. However, microglia in the brain tissue from mice expressing mutant STAT1are stained more intensely, and their cell bodies have better defined projections. Thus, expression of Iba1 is elevated in the mutant animals, indicating heightened microglia activation. Whether there is also an increase in the number of microglia is unclear at present, since a detailed cell count has not yet been performed.

We therefore concluded that both peripheral and central-nervous system macrophages show increased activity in the mice expressing SUMO-free STAT1. This is a strong indicator that the cognitive deficiencies seen in these animals are linked to inflammatory events in the brain. This, in turn, strengthens the notion that SUMO modification of STAT1 is a disease-associated modification, and hence a potential drug target for therapy.

### **CLAIMS**

- 1. An isolated nucleic acid molecule comprising a nucleic acid sequence encoding a Signal Transducer Activator of Transcription 1 (STAT1) protein wherein the sequence comprises a polymorphism within a polymorphic region represented by nucleotides 2104 to 2115.
- 2. An isolated nucleic acid molecule as claimed in claim 1 wherein the polymorphism is within a polymorphic region represented by nucleotides 2113 to 2115 of STAT1.
- 3. An isolated nucleic acid molecule as claimed in claim 1 or claim 2 wherein the polymorphism is a knock-in mutation.
- 4. An isolated nucleic acid molecule as claimed in any preceding claim wherein the polymorphism is a single nucleotide polymorphism (SNP).
- 5. An isolated nucleic acid molecule as claimed in claim 4 wherein the SNP is at nucleotide 2113 of the STAT1 nucleic acid sequence.
- 6. An isolated nucleic acid molecule as claimed in claim 5 wherein the nucleotide at position 2113 of the STAT1 nucleic acid sequence is replaced by a C.
- 7. An isolated nucleic acid molecule as claimed in any preceding claim wherein the STAT 1 protein is a mammalian STAT1 protein.
- 8. An isolated nucleic acid molecule as claimed in claim 7 wherein the STAT1 protein is a human STAT1 protein.

- 9. An isolated nucleic acid molecule as claimed in claim 8 wherein the nucleic acid sequence encoding a human STAT1 protein is shown in Figure 1
- 10. An isolated nucleic acid molecule as claimed in claim 7 wherein the STAT1 protein is a rodent STAT1 protein.
- 11. An isolated nucleic acid molecule as claimed in claim 10 wherein the STAT1 protein is a mouse STAT1 protein.
- 12. An isolated nucleic acid molecule as claimed in claim 11 wherein the nucleic acid sequence encoding a mouse STAT1 protein is shown in Figure 2
- 13. An isolated nucleic acid molecule comprising a nucleic acid sequence selected from:
- i) a nucleic acid sequence as shown in Figure 1 wherein the sequence comprises a polymorphism at one or more of nucleotides 2104 to 2115 of Figure 1; and
- ii) a nucleic acid sequence which hybridises under stringent conditions to a sequence shown in (i).
- 14. An isolated nucleic acid molecule comprising a nucleic acid sequence selected from:
- i) a nucleic acid sequence as shown in Figure 2 wherein the sequence comprises a polymorphism at one or more of nucleotides 2104 to 2115 of Figure 2; and
- ii) a nucleic acid sequence which hybridises under stringent conditions to a sequence shown in (i).
- 15. A STAT 1 polypeptide wherein at least one of residues 702 to 705 inclusive is modified by addition, deletion or substitution.
- 16. A STAT1 polypeptide as claimed in claim 15 wherein at least one of residues 703 and/or 705 is modified.

- 17. A STAT1 polypeptide as claimed in claim 16 wherein residue 705 is modified.
- 18. A STAT1 polypeptide as claimed in claim 17 wherein residue 705 is replaced by a residue selected from glutamine, asparagine and alanine.
- 19. A STAT1 polypeptide as claimed in claim 18 wherein residue 705 is replaced by glutamine or asparagine.
- 20. A STAT1 polypeptide as claimed in claim 19 wherein residue 705 is replaced by glutamine.
- 21. A STAT1 polypeptide as claimed in any of claims 15 to 20 wherein the STAT1 polypeptide is a human STAT1 polypeptide.
- 22. A STAT1 polypeptide as claimed in claim 21 wherein the STAT1 polypeptide comprises an amino acid sequence as shown in Figure 3 wherein at least one of residues 702 to 705 inclusive is modified by addition, deletion or substitution.
- 23. A STAT1 polypeptide as claimed in any of claims 15 to 20 wherein the STAT1 polypeptide is a mouse STAT1 polypeptide.
- 24. A STAT1 polypeptide as claimed in claim 23 wherein the STAT1 polypeptide comprises an amino acid sequence as shown in Figure 4 wherein at least one of residues 702 to 705 inclusive is modified by addition, deletion or substitution.
- 25. A vector comprising an isolated nucleic molecule as claimed in any one of claims 1 to 14.

- 26. A host cell comprising an isolated nucleic acid molecule as claimed in any one of claims 1 to 14 or a vector as claimed in claim 25.
- 27. A transgenic non-human animal, or progeny thereof, comprising an isolated nucleic acid molecule as claimed in any one of claims 1 to 14 or a vector as claimed in claim 25.
- 28. A transgenic non-human animal as claimed in claim 27 which is a mammal.
- 29. A transgenic non-human animal as claimed in claim 28 wherein the mammal is a rodent.
- 30. A transgenic non-human animal as claimed in claim 97 wherein the rodent is a mouse.
- 31. A method of preparing a transgenic non-human animal as claimed in any one of claims 27 to 30 comprising the steps of:
  - i) incorporating an isolated nucleic acid molecule as claimed in any one of claims 1 to 14, or a vector as claimed in claim 25, into non-human embryonic stem cells;
  - ii) transferring the embryonic stem cells to a recipient female nonhuman animal; and
  - iii) growing the embryonic stem cells into a transgenic non-human animal.
- 32. A kit for the diagnosis of disease or predisposition to disease, comprising a means for determining the presence or absence of a polymorphism at one or more of nucleotides 2104 to 2115 of a nucleic acid sequence encoding STAT1 wherein the polymorphism is diagnostic of disease or predisposition to disease.
- 33. A kit as claimed in claim 32 wherein the nucleic acid sequence encoding STAT1 is shown in Figure 1.

- 34. A kit as claimed in claim 32 or claim 33 which comprises one or more containers comprising a pair of primers, wherein one of the primers within the pair is capable of hybridizing directly to, or adjacent to, the nucleic acid sequence represented by nucleotides 2104 to 2115 of STAT1.
- 35. A kit as claimed in claim 34 which comprises one or more of the pairs of primers represented in Figure 5(a) and/or (b).
- 36. A method for the diagnosis of a polymorphism in a subject comprising
  - i) obtaining a nucleic acid sample obtained from the subject; and
- ii) determining the presence or absence of polymorphism within a polymorphic region represented by nucleotides 2104 to 2115 of a nucleic acid sequence encoding STAT1.
- 37. A method as claimed in claim 36 wherein the method comprises the steps of:
- i) obtaining a nucleic acid sample from a subject;
- ii) amplifying said isolated nucleic acid by PCR with primers that flank the polymorphism or polymorphic region to produce amplified nucleic acid regions; and
- iii) sequencing said amplified regions to determine the presence or absence of said polymorphism.
- 38. A method as claimed in claim 36 or claim 37 wherein the subject is a mammal.
- 39. A method as claimed in claim 38 wherein the subject is human.
- 40. A method as claimed in any one of claims 36 to 39 wherein the presence of said polymorphism is indicative of disease or susceptibility to disease.

- 41. A method as claimed in claim 40 wherein the disease is selected from the group consisting of neurological diseases/disorders, inflammatory diseases and infertility.
- 42. A method for the diagnosis of a polymorphism in a subject comprising
  - i) obtaining a nucleic acid sample obtained from the subject; and
- ii) determining the presence or absence of polymorphism within a polymorphic region represented by nucleotides 2104 to 2115 of a nucleic acid sequence encoding STAT1 wherein the presence of said polymorphism is indicative of a reduced susceptibility, or tolerance, to disease.
- 43. A method as claimed in claim 42 wherein the disease is selected from the group consisting of cancer, microbial infections and microbial related diseases, and auto immune diseases.
- 44. A method for diagnosing disease or predisposition to disease, comprising determining the presence or absence of a polymorphism at one or more of nucleotides 2104 to 2115 of a nucleic acid sequence encoding STAT1, wherein the presence of the polymorphism is diagnostic of disease or predisposition to disease.
- 45. A method of diagnosing disease or predisposition to disease, comprising determining the presence or absence of a polymorphism at one or more of nucleotides 2104 to 2115 of a nucleic acid sequence encoding STAT1, wherein the presence of the polymorphism is diagnostic of a reduced susceptibility, or tolerance, to disease.
- 46. A method of treating or preventing disease comprising determining the presence or absence of a polymorphism at one or more of nucleotides 2104 to 2115 of a nucleic acid sequence encoding STAT1; and if the polymorphism is present, administering treatment in order to prevent, delay or reduce the disease.

- 47. A method of identifying an agent for the treatment of disease, comprising (a) administration of an agent to animal tissue comprising an isolated nucleic acid molecule as claimed in any one of claims 1 to 14; and (b) determining whether the agent modulates the effect(s) of the polymorphism.
- 48. Use of an isolated nucleic acid molecule as claimed in any one of claims 1 to 14 to screen for agents which are capable of modulating the effect(s) of the polymorphism.
- 49. An agent identified by the method of claim 47.
- 50. An agent as claimed in claim 49 wherein the agent is an antisense sequence that is capable of hybridizing to an isolated nucleic acid molecule as claimed in any one of claims 1 to 14 under stringent conditions.
- 51. An agent as claimed in claim 50 wherein the agent is an antibody that binds to a polymorphic region of a STAT1 polypeptide wherein the polymorphic region is represented by residues 702 to 705 inclusive.
- 52. An antibody specifically recognising a polymorphic region of a STAT1 polypeptide wherein the polymorphic region is represented by residues 702 to 705 of a STAT1 polypeptide.
- 53. An agent as claimed in any one of claims 49 to 51 for use in preventing or treating a neurological or inflammatory disease.
- 54. Use of a STAT1 polypeptide in the identification of agents that block the binding of a regulatory protein to residue 703 of the STAT1 polypeptide.
- 55. Use as claimed in claim 54 wherein the regulatory protein is a Small Ubiquitin-related Modifier (SUMO) protein.

#### Figure 1

1 atg tct cag tgg tac gaa ctt cag cag ctt gac tca aaa 40 ttc ctq gag cag gtt cac cag ctt tat gat gac agt ttt 79 ccc atq qaa atc aga cag tac ctg gca cag tgg tta gaa 118 aag caa gac tgg gag cac gct gcc aat gat gtt tca ttt 157 gcc acc atc cgt ttt cat gac ctc ctg tca cag ctg gat 196 gat caa tat agt cgc ttt tct ttg gag aat aac ttc ttg 235 cta cag cat aac ata agg aaa agc aag cgt aat ctt cag 274 gat aat ttt cag gaa gac cca atc cag atg tct atg atc 313 att tac agc tgt ctg aag gaa gaa agg aaa att ctg gaa 352 aac qcc cag aga ttt aat cag gct cag tcg ggg aat att 391 cag age aca gtg atg tta gae aaa cag aaa gag ett gae 430 agt aaa gtc aga aat gtg aag gac aag gtt atg tgt ata 469 gag cat gaa atc aag agc ctg gaa gat tta caa gat gaa 508 tat gac ttc aaa tgc aaa acc ttg cag aac aga gaa cac 547 gag acc aat ggt gtg gca aag agt gat cag aaa caa gaa 586 cag ctg tta ctc aag aag atg tat tta atg ctt gac aat 625 aag aga aag gaa gta gtt cac aaa ata ata gag ttg ctg 664 aat gtc act gaa ctt acc cag aat gcc ctg att aat gat 703 gaa cta gtg gag tgg aag cgg aga cag cag agc gcc tgt 742 att ggg ggg ccg ccc aat gct tgc ttg gat cag ctg cag 781 aac too ttc act ata gtt gcg gag agt ctg cag caa gtt 820 cgg cag cat aaa aag ttg gag gaa ttg gaa cag aaa 859 tac acc tac gaa cat gac cct atc aca aaa aac aaa caa

898 gtg tta tgg gac cgc acc ttc agt ctt ttc cag cag ctc 937 att cag agc tcg ttt gtg gtg gaa aga cag ccc tgc atg 976 cca acg cac cct cag agg ccg ctg gtc ttg aag aca ggg 1015 gtc cag ttc act gtg aag ttg aga ctg ttg gtg aaa ttg 1054 caa gag ctg aat tat aat ttg aaa gtc aaa gtc tta ttt 1093 gat aaa gat gtg aat gag aga aat aca gta aaa gga ttt 1132 agg aag ttc aac att ttg ggc acg cac aca aaa gtg atg 1171 aac atg gag gag tcc acc aat ggc agt ctg gcg gct gaa 1210 ttt cgg cac ctg caa ttg aaa gaa cag aaa aat gct ggc 1249 acc aga acg aat gag ggt cct ctc atc gtt act gaa gag 1288 ctt cac tcc ctt agt ttt gaa acc caa ttg tgc cag cct 1327 ggt ttg gta att gac ctc gag acg acc tct ctg ccc gtt 1366 gtg gtg atc tcc aac gtc agc cag ctc ccg agc ggt tgg 1405 gcc tcc atc ctt tgg tac aac atg ctg gtg gcg gaa ccc 1444 agg aat ctg tcc ttc ttc ctg act cca cca tgt gca cga 1483 tgg gct cag ctt tca gaa gtg ctg agt tgg cag ttt tct 1522 tct gtc acc aaa aga ggt ctc aat gtg gac cag ctg aac 1561 atg ttg gga gag aag ctt ctt ggt cct aac gcc agc ccc 1600 gat ggt ctc att ccg tgg acg agg ttt tgt aag gaa aat 1639 ata aat gat aaa aat ttt ccc ttc tgg ctt tgg att gaa 1678 agc atc cta gaa ctc att aaa aaa cac ctg ctc cct ctc 1717 tgg aat gat ggg tgc atc atg ggc ttc atc agc aag gag 1756 cga gag cgt gcc ctg ttg aag gac cag cag ccg qqq acc 1795 ttc ctg ctg cgg ttc agt gag agc tcc cgg gaa ggg qcc

1834 atc aca ttc aca tgg gtg gag cgg tcc cag aac gga ggc
1873 gaa cct gac ttc cat gcg gtt gaa ccc tac acg aag aaa
1912 gaa ctt tct gct gtt act ttc cct gac atc att cgc aat
1951 tac aaa gtc atg gct gct gag aat att cct gag aat ccc
1990 ctg aag tat ctg tat cca aat att gac aaa gac cat gcc
2029 ttt gga aag tat tac tcc agg cca aag gaa gca cca gag
2068 cca atg gaa ctt gat ggc cct aaa gga act gga tat atc
2107 aag act gag ttg att tct gtg tct gaa gtt cac cct tct
2146 aga ctt cag acc aca gac aac ctg ctc ccc atg tct cct
2185 gag gag ttt gac gag gtg tct cgg ata gtg ggc tct gta
2224 gaa ttc gac agt atg atg aac aca gta tag

#### Figure 2

1 atg tca cag tgg ttc gag ctt cag cag ctg gac tcc aag 40 ttc ctg gag cag gtc cac cag ctg tac gat gac agt ttc 79 ccc atq gaa atc aga cag tac ctg gcc cag tgg ctg gaa 118 aag caa gac tgg gag cac gct gcc tat gat gtc tcg ttt 157 gcg acc atc cgc ttc cat gac ctc ctc tca cag ctg gac 196 gac cag tac agc cgc ttt tct ctg gag aat aat ttc ttg 235 ttg cag cac aac ata cgg aaa agc aag cgt aat ctc cag 274 gat aac ttc caa gaa gat ccc gta cag atg tcc atg atc 313 atc tac aac tgt ctg aag gaa gaa agg aag att ttg gaa 352 aat qcc caa aga ttt aat cag gcc cag gag gga aat att 391 cag aac act gtg atg tta gat aaa cag aag gag ctg gac 430 agt aaa gtc aga aat gtg aag gat caa gtc atg tgc ata 469 gag cag gaa atc aag acc cta gaa gaa tta caa gat gaa 508 tat gac ttt aaa tgc aaa acc tct cag aac aga gaa ggt 547 gaa gcc aat ggt gtg gcg aag agc gac caa aaa cag gaa 586 cag ctg ctc cac aag atg ttt tta atg ctt gac aat 625 aag aga aag gag ata att cac aaa atc aga gag ttg ctg 664 aat tcc atc gag ctc act cag aac act ctg att aat gac 703 gag ctc gtg gag tgg aag cga agg cag cag agc gcc tgc 742 atc qgg gga ccg ccc aac gcc tgc ctg gat cag ctg caa 781 agc tgg ttc acc att gtt gca gag acc ctg cag cag atc 820 cqt cag cag ctt aaa aag ctg gag gag ttg gaa cag aaa 859 ttc acc tat gag ccc gac cct att aca aaa aac aag cag

898 gtg ttg tca gat cga acc ttc ctc ctc ttc cag cag ctc 937 att cag agc tcc ttc gtg gta gaa cga cag ccg tgc atg 976 ccc act cac ccg cag agg ccc ctg gtc ttg aag act ggg 1015 gta cag ttc act gtc aag ctg aga ctg ttg gtg aaa ttg 1054 caa gag ctg aac tat aac ttg aaa gtg aaa gtc tca ttt 1093 gac aaa gat gtg aac gag aaa aac aca gtt aaa gga ttt 1132 cgg aag ttc aac atc ttg ggt acg cac aca aaa gtg atg 1171 aac atg gaa gaa tcc acc aac gga agt ctg gca gct gag 1210 ttc cga cac ctg caa ctg aag gaa cag aaa aac gct ggg 1249 aac aga act aat gag ggg cct ctc att gtc acc gaa gaa 1288 ctt cac tct ctt agc ttt gaa acc cag ttg tgc cag cca 1327 ggc ttg gtg att gac ctg gag acc acc tct ctt cct gtc 1366 gtg gtg atc tcc aac gtc agc cag ctc ccc agt ggc tgg 1405 gcg tct atc ctg tgg tac aac atg ctg gtg aca gag ccc 1444 agg aat ctc tcc ttc ttc ctg aac ccc ccg tgc gcg tgg 1483 tgg tcc cag ctc tca gag gtg ttg agt tgg cag ttt tca 1522 tca qtc acc aag aga ggt ctg aac gca gac cag ctg agc 1561 atg ctg gga gag aag ctg ctg ggc cct aat gct ggc cct 1600 gat ggt ctt att cca tgg aca agg ttt tgt aag gaa aat 1639 att aat gat aaa aat ttc tcc ttc tgg cct tgg att gac 1678 acc atc cta gag ctc att aag aag cac ctg ctg tgc ctc 1717 tgg aat gat ggg tgc att atg ggc ttc atc agc aag gag 1756 cga gaa cgc gct ctg ctc aag gac cag cag cca ggg acg 1795 ttc ctg ctt aga ttc agt gag agc tcc cgg gaa ggg gcc

1834 atc aca ttc aca tgg gtg gaa cgg tcc cag aac gga ggt
1873 gaa cct gac ttc cat gcc gtg gag ccc tac acg aaa aaa
1912 gaa ctt tca gct gtt act ttc cca gat att att cgc aac
1951 tac aaa gtc atg gct gcc gag aac ata cca gag aat ccc
1990 ctg aag tat ctg tac ccc aat att gac aaa gac cac gcc
2029 ttt ggg aag tat tat tcc aga cca aag gaa gca cca gaa
2068 ccg atg gag ctt gac gac cct aag cga act gga tac atc
2107 aag act gag ttg att tct gtg tct gaa gtc cac cct tct
2146 aga ctt cag acc aca gac aac ctg ctt ccc atg tct cca
2185 gag gag ttt gat gag atg tcc cgg ata gtg ggc ccc gaa
2224 ttt gac agt atg atg atg aca gta taa

#### Figure 3

MSQWFELQQLDSKFLEQVHQLYDDSFPMEIRQYLAQWLEKQDWEHAAYDVSFATIRF HDLLSQLDDQYSRFSLENNFLLQHNIRKSKRNLQDNFQEDPVQMSMIIYNCLKEERKILE NAQRFNQAQEGNIQNTVMLDKQKELDSKVRNVKDQVMCIEQEIKTLEELQDEYDFKCK TSQNREGEANGVAKSDQKQEQLLLHKMFLMLDNKRKEIIHKIRELLNSIELTQNTLINDEL VEWKRRQQSACIGGPPNACLDQLQSWFTIVAETLQQIRQQLKKLEELEQKFTYEPDPIT KNKQVLSDRTFLLFQQLIQSSFVVERQPCMPTHPQRPLVLKTGVQFTVKLRLLVKLQEL NYNLKVKVSFDKDVNEKNTVKGFRKFNILGTHTKVMNMEESTNGSLAAEFRHLQLKEQ KNAGNRTNEGPLIVTEELHSLSFETQLCQPGLVIDLETTSLPVVVISNVSQLPSGWASIL WYNMLVTEPRNLSFFLNPPCAWWSQLSEVLSWQFSSVTKRGLNADQLSMLGEKLLGP NAGPDGLIPWTRFCKENINDKNFSFWPWIDTILELIKKHLLCLWNDGCIMGFISKERERA LLKDQQPGTFLLRFSESSREGAITFTWVERSQNGGEPDFHAVEPYTKKELSAVTFPDIIR NYKVMAAENIPENPLKYLYPNIDKDHAFGKYYSRPKEAPEPMELDDPKRTGY<u>IKTE</u>LISV SEVHPSRLQTTDNLLPMSPEEFDEMSRIVGPEFDSMMSTV\*

#### Figure 4

MSQWYELQQLDSKFLEQVHQLYDDSFPMEIRQYLAQWLEKQDWEHAANDVSFATIRF HDLLSQLDDQYSRFSLENNFLLQHNIRKSKRNLQDNFQEDPIQMSMIIYSCLKEERKILE NAQRFNQAQSGNIQSTVMLDKQKELDSKVRNVKDKVMCIEHEIKSLEDLQDEYDFKCK TLQNREHETNGVAKSDQKQEQLLLKKMYLMLDNKRKEVVHKIIELLNVTELTQNALINDE LVEWKRRQQSACIGGPPNACLDQLQNWFTIVAESLQQVRQQLKKLEELEQKYTYEHD PITKNKQVLWDRTFSLFQQLIQSSFVVERQPCMPTHPQRPLVLKTGVQFTVKLRLLVKL QELNYNLKVKVLFDKDVNERNTVKGFRKFNILGTHTKVMNMEESTNGSLAAEFRHLQLK EQKNAGTRTNEGPLIVTEELHSLSFETQLCQPGLVIDLETTSLPVVVISNVSQLPSGWAS ILWYNMLVAEPRNLSFFLTPPCARWAQLSEVLSWQFSSVTKRGLNVDQLNMLGEKLLG PNASPDGLIPWTRFCKENINDKNFPFWLWIESILELIKKHLLPLWNDGCIMGFISKERER ALLKDQQPGTFLLRFSESSREGAITFTWVERSQNGGEPDFHAVEPYTKKELSAVTFPDII RNYKVMAAENIPENPLKYLYPNIDKDHAFGKYYSRPKEAPEPMELDGPKGTGY<u>IKTE</u>LIS VSEVHPSRLQTTDNLLPMSPEEFDEVSRIVGSVEFDSMMNTV\*

# Figure 5

(a)

5'-aaaactcgagggcacctgcacagggt-3'

5'-aaaactcgaggctcaccctggacaga-3'

(b)

5'-attttaggtacccgagagtccaccacctcagg-3'

5'-attttaggtaccggacagagcacatcgtgg-3'

Figure 6

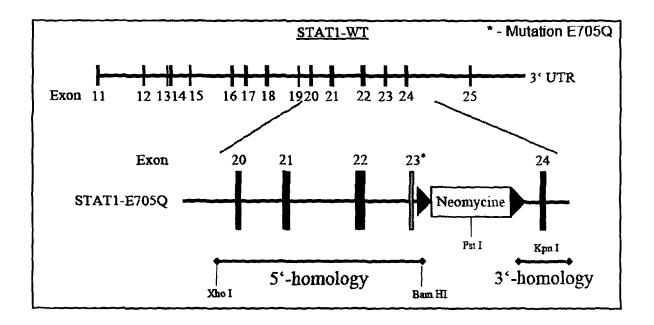


Figure 7

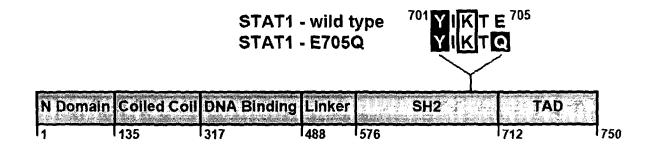


Figure 8

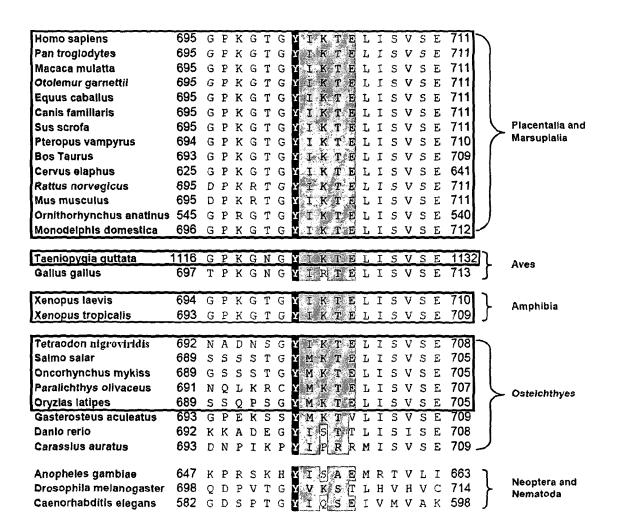
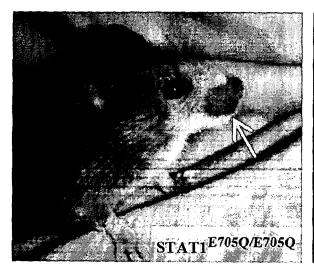


Figure 9





International application No
PCT/GB2011/001696

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K14/47 C07K16/18 A01K67/027 C12Q1/68 A61K39/00
ADD.

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

#### B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C07K A01K C12Q A61K

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

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, BIOSIS, Sequence Search, EMBASE, FSTA, WPI Data

	Ī	
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	D. UNGUREANU: "PIAS proteins promote SUMO-1 conjugation to STAT1", BL00D, vol. 102, no. 9, 1 November 2003 (2003-11-01), pages 3311-3313, XP55024432, ISSN: 0006-4971, DOI: 10.1182/blood-2002-12-3816 page 3311, right-hand column, paragraph 1st full figures 1a, 1c	1-55

X See patent family annex.
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention  "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone  "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled
in the art. "&" document member of the same patent family
Date of mailing of the international search report
24/04/2012
Authorized officer Ripaud, Leslie
•

International application No
PCT/GB2011/001696

C(Continua	tion). DOCUMENTS CONSIDERED TO BE RELEVANT	
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	D. UNGUREANU: "SUMO-1 conjugation selectively modulates STAT1-mediated gene responses", BLOOD, vol. 106, no. 1, 1 January 2005 (2005-01-01), pages 224-226, XP55024433, ISSN: 0006-4971, DOI: 10.1182/blood-2004-11-4514 cited in the application page 224, right-hand column, paragraph 1st-2nd full page 225, left-hand column, paragraph 2nd full figures 1A, 1B	1-55
X	R. S. ROGERS: "SUMO Modification of STAT1 and Its Role in PIAS-mediated Inhibition of Gene Activation", JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 278, no. 32, 4 June 2003 (2003-06-04), pages 30091-30097, XP55024434, ISSN: 0021-9258, DOI: 10.1074/jbc.M301344200 cited in the application page 30095, left-hand column, paragraph bridging - right-hand column	1-55
X	L. SONG ET AL: "Stat1 and SUMO modification", BL00D, vol. 108, no. 10, 15 November 2006 (2006-11-15), pages 3237-3244, XP55024435, ISSN: 0006-4971, DOI: 10.1182/blood-2006-04-020271 cited in the application page 3238, left-hand column, paragraph 1st full page 3238, right-hand column, paragraph 2nd full	1-55
X	WO 2010/015376 A1 (HANNOVER MED HOCHSCHULE [DE]; NIEDENTHAL RAINER [DE]; JAKOBS ASTRID [D) 11 February 2010 (2010-02-11) page 14, lines 15-28 page 18, line 24 page 19, line 30 - page 20, line 9	1-55
X	WO 2010/076322 A1 (SIEMENS HEALTHCARE DIAGNOSTICS [US]; KRONENWETT RALF [DE]; VON TOERNE) 8 July 2010 (2010-07-08) sequences 110, 171 page 48, lines 19-27 page 61, lines 10-13	13,14, 32-34, 49,50

International application No
PCT/GB2011/001696

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Anonymous: "STAT1 (phosphoY701) polyclonal antibody", Abnova catalog #: PAB12752, 15 April 2010 (2010-04-15), pages 1-4, XP55024350, Retrieved from the Internet: URL:http://www.abnova.com/PDFServer/output s/PAB12752.pdf [retrieved on 2012-04-12] page 1	49,51,52
X,P	A. BEGITT ET AL: "SUMO conjugation of STAT1 protects cells from hyperresponsiveness to IFN", BLOOD, vol. 118, no. 4, 28 July 2011 (2011-07-28), pages 1002-1007, XP55024436, ISSN: 0006-4971, DOI: 10.1182/blood-2011-04-347930 the whole document	1-55
Α	US 2010/233718 A1 (AUBERT JEROME [FR] ET AL) 16 September 2010 (2010-09-16) abstract table S2	1-55
А	WO 2005/038052 A2 (ALTANA PHARMA AG [DE]; WITTENBERGER TIMO [DE]; HOFMANN HANS-PETER [DE]) 28 April 2005 (2005-04-28) page 2, line 30 - page 3, line 12	1-55
Α	LOUISA VARINOU ET AL: "Phosphorylation of the Statl Transactivation Domain Is Required for Full-Fledged IFN-[gamma]-Dependent Innate Immunity", IMMUNITY, vol. 19, no. 6, 1 December 2003 (2003-12-01), pages 793-802, XP55024437, ISSN: 1074-7613, DOI: 10.1016/S1074-7613(03)00322-4 page 794, paragraph bridging - page 795	1-55
Α	EP 1 637 540 A1 (FORSCHUNGSVERBUND BERLIN EV [DE]) 22 March 2006 (2006-03-22) paragraphs [0083], [0084], [0090]/	1-55

International application No
PCT/GB2011/001696

0/0 ::	PRODUCTION OF THE PRODUCTION O	PCT/GB2011/001696
C(Continua	tion). DOCUMENTS CONSIDERED TO BE RELEVANT	<u> </u>
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KHATIB H ET AL: "Effects of signal transducer and activator of transcription (STAT) genes STAT1 and STAT3 genotypic combinations on fertilization and embryonic survival rates in Holstein cattle", JOURNAL OF DAIRY SCIENCE, AMERICAN DAIRY SCIENCE ASSOCIATION, US, vol. 92, no. 12, 1 December 2009 (2009-12-01), pages 6186-6191, XP026954859, ISSN: 0022-0302 [retrieved on 2009-12-01] page 6186	1-55
Α	L Adámková ET AL: "Transcription Protein STAT1: Biology and Relation to Cancer", Folia Biologica, 1 January 2007 (2007-01-01), pages 1-6, XP55024442, Retrieved from the Internet: URL:http://fb.cuni.cz/Data/files/folia_biologica/volume_53_2007_1/FB2007A0001.pdf [retrieved on 2012-04-13] the whole document	1-55
A	PFITZNER ET AL: "THE ROLE OF STATS in inflammation and inflammatory diseases", CURRENT PHARMACEUTICAL DESIGN, BENTHAM SCIENCE PUBLISHERS, NL, vol. 10, 1 January 2004 (2004-01-01), pages 2839-2850, XP008106241, ISSN: 1381-6128, DOI: 10.2174/1381612043383638 the whole document	1-55

Information on patent family members

International application No
PCT/GB2011/001696

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2010015376 A1	11-02-2010	EP 2154533 A1 EP 2329264 A1 US 2011124011 A1 WO 2010015376 A1	17-02-2010 08-06-2011 26-05-2011 11-02-2010
WO 2010076322 A1	08-07-2010	NONE	
US 2010233718 A1	16-09-2010	CA 2703333 A1 EP 2212442 A1 US 2010233718 A1 WO 2009053493 A1	30-04-2009 04-08-2010 16-09-2010 30-04-2009
WO 2005038052 A2	28-04-2005	NONE	
EP 1637540 A1	22-03-2006	NONE	