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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 December 2013 (27.12.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/190321 A1

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

(21) International Application Number:
PCT/GB2013/051635

(22) International Filing Date:
21 June 2013 (21.06.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
1211158.9 22 June 2012 (22.06.2012) GB

(71) Applicant: **NOTTINGHAM TRENT UNIVERSITY** [GB/GB]; Burton Street, Nottingham NG1 4BU (GB).

(72) Inventor: **GRAHAM, Roy Ball**; School of Science and Technology, Nottingham Trent University, Clifton Lane, Nottingham NG11 8NS (GB).

(74) Agent: **BARKER BRETTELL LLP**; 100 Hagley Road, Edgbaston, Birmingham, West Midlands B16 8QQ (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, 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, IS, JP, KE, KG, 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, PA, 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, 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, KM, 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))*

[Continued on next page]

(54) Title: BIOMARKERS FOR DETERMINING THE M. TUBERCULOSIS INFECTION STATUS

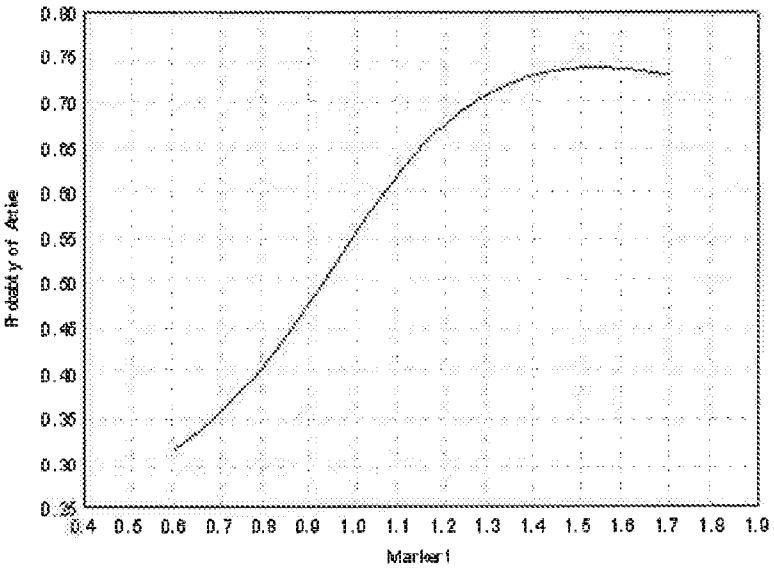

Figure 9B

(57) Abstract: The present invention relates to novel biomarkers for determining the Mycobacterium tuberculosis infection status of a subject, and to uses of novel panels of biomarkers. A method of determining the M. tuberculosis infection status of a subject comprising providing a sample of material obtained from the subject determining the expression level of one or more genes selected from the group consisting of: NXNL1, PSMA7, C6orf61, EMP1, CLIC1, LACTB and DUSP3 in the sample.

WO 2013/190321 A1

— *with sequence listing part of description (Rule 5.2(a))*

BIOMARKERS FOR DETERMINING THE M. TUBERCULOSIS INFECTION STATUS

The present invention relates to novel biomarkers for determining the *Mycobacterium tuberculosis* infection status of a subject, and to uses of novel panels of biomarkers.

Tuberculosis (TB) is an infection caused by *Mycobacterium tuberculosis (M. tuberculosis)* and is a major cause or morbidity and mortality worldwide. *M. tuberculosis* is an airborne bacterial infection that primarily affects the lungs. It is estimated that approximately 2.2 billion people or a third of the world's population are infected with *M. tuberculosis*.

The majority of infected people remain asymptomatic. Infected people who remain asymptomatic are said to have latent *M. tuberculosis* infection (latent TB). A person infected with M. tuberculosis has about a 10% lifetime risk of developing active *M. tuberculosis* infection (active TB) where symptoms of *M. tuberculosis* infection are shown.

M. tuberculosis is a substantial management and cost burden for healthcare systems and could be reduced with improvements in diagnosis and informed patient management. Treatments for *M. tuberculosis* infection vary depending on the type of *M. tuberculosis* infection status that a person has. A person who has latent *M. tuberculosis* infection and does not have active *M. tuberculosis* infection may be given preventative therapy.

Current tests for *M. tuberculosis* infection can not distinguish between latent and active *M. tuberculosis* infection and cannot identify which individuals having latent *M. tuberculosis* infection will go on to develop active *M. tuberculosis* infection.

It would be advantageous to be able to test for the infection status of an individual subject to distinguish between uninfected individuals and individuals having latent *M. tuberculosis* infection and active *M. tuberculosis* infection. Such a test could also be used to follow the infection status of individuals to

determine whether TB is being activated or during treatment to determine when the *M. tuberculosis* infection has been cleared.

In a first aspect the present invention provides a method of determining the *M. tuberculosis* infection status of a subject comprising:

(a) providing a sample of material obtained from the subject; and

(b) determining the expression level of one or more genes selected from the group consisting of: NXNL1, PSMA7, C6orf61, EMP1, CLIC1, LACTB and DUSP3 in the sample.

Optionally, the method could comprise a step (c) determining the expression level of one or more genes selected from table 3 or table 4.

The method may further comprise: (d) employing the expression level determined in (b) and optionally (c) to distinguish between subjects not infected with *M. tuberculosis*, subjects with latent *M. tuberculosis* infection and subjects with active *M. tuberculosis* infection.

Subjects not infected with *M. tuberculosis* may be defined as subjects that test negative for *M. tuberculosis* using a tuberculin-skin test (TST) and/or test negative for *M. tuberculosis* using an antigen-specific IFN-gamma release assay (IGRA).

Subjects with latent *M. tuberculosis* infection may be defined as subjects that test positive for *M. tuberculosis* using a tuberculin-skin test (TST) and/or test positive for *M. tuberculosis* using an antigen-specific IFN-gamma release assay (IGRA) but do not have symptoms of tuberculosis such as a persistent cough.

Subjects with active *M. tuberculosis* infection may be confirmed by culture for *M. tuberculosis* from a blood, serum or sputum sample and also test positive for *M. tuberculosis* using a tuberculin-skin test (TST) and/or test positive for *M. tuberculosis* using an antigen-specific IFN-gamma release assay (IGRA) and also have symptoms of tuberculosis.

The sample of material may be a sample of blood, sputum, saliva, wound exudate, urine, faeces, peritoneal fluid or any respiratory secretion. A sample of blood may be whole blood, blood plasma or blood serum.

Preferably the sample is a sample of whole blood. Blood samples have the advantage that they are readily obtainable and tend to be more homogenous in nature than other sample types. Samples of whole blood contain RNA which can be extracted to generate a transcriptional profile.

The expression level of one or more genes selected from the group consisting of: NXNL1, PSMA7, C6orf61, EMP1, CLIC1, LACTB and DUSP3 or selected from table 3 or table 4 in a sample may be determined by any suitable method. Such methods include methods that quantify the nucleotide products of these genes such as: quantitative PCR using suitable oligonucleotide primers designed to adhere within the sequence of an mRNA encoded by the gene of interest; analysis of expression arrays; next generation sequencing, comparative genomic hybridisation arrays (CGH arrays); multiplexed PCR. The expression level of one or more genes selected from the group consisting of: NXNL1, PSMA7, C6orf61, EMP1, CLIC1, LACTB and DUSP3 or selected from table 3 or table 4 in a sample may also be determined using methods that quantify the protein products of the selected genes, for example: ELISA immunohistochemistry; protein aptamer arrays or protein immunological arrays.

The expression level of the gene may be determined by measuring the rate of polymerisation of the RNA using standard techniques.

The method of the invention may not include the step of obtaining the sample.

The method of the invention may include a further step of comparing the determined value of the expression level of one or more genes selected from the group consisting of: NXNL1, PSMA7, C6orf61, EMP1, CLIC1, LACTB and DUSP3 or selected from table 3 or table 4 in a sample with a reference value.

The reference value may be the value for the expression level of the same gene in a sample of the same sample type, from an individual who is known to have or not to have infection with *M. tuberculosis*. Alternatively, or additionally, the reference value may be the expression level of the same gene in the same sample type in a sample taken previously from the same subject, for example, prior to or during the course of a particular treatment. The reference sample may be a sample of the same type, for example, both samples may be blood samples. In this way the expression level of one or more genes selected from the group consisting of: NXNL1, PSMA7, C6orf61, EMP1, CLIC1, LACTB and DUSP3 or selected from table 3 or table 4 in a sample may be used to monitor the progression of an infection in a subject, and/or to monitor the efficacy of a particular treatment in a subject.

Alternatively or in addition the reference value may be the expression level for the selected gene in an individual that is infected with *M. tuberculosis* or the reference value may be the expression level for the selected gene in an individual that is not infected with *M. tuberculosis*.

The method of the invention may be carried out *in vitro*.

The subject may be a mammal, and is preferably a human, but may alternatively be a monkey, ape, cat, dog, cow, horse, deer, badger, rabbit or rodent.

The method may be used in one or more of the following; diagnosing whether or not a subject has *M. tuberculosis* infection; advising on the prognosis for a subject with a *M. tuberculosis* infection; and monitoring the effectiveness or response of a subject to a particular treatment for infection by *M. tuberculosis*.

In the method altered expression level of one or more genes selected from: NXNL1, PSMA7, C6orf61 and EMP1 and optionally one or more genes selected from table 3 can differentiate between subjects not infected with *M. tuberculosis* and subjects with latent *M. tuberculosis* infection. NXNL1, PSMA7, C6orf61 and EMP1 and optionally one or more genes selected from table 3 may be used as a biomarker panel to distinguish between subjects not infected with *M. tuberculosis*

and subjects with latent *M. tuberculosis* infection with a high degree of accuracy. This biomarker panel is advantageous because it provides a high degree of accuracy with only a small number of biomarkers. The evaluation error of this biomarker panel is 0.93%. Evaluation error is calculated as the square root of the averaged difference between model predictions and the actual TB status squared, expressed as a percentage.

In subjects with latent *M. tuberculosis* infection the expression levels of NXNL1, PSMA7 and C6orf61 are lower then in subjects not infected with *M. tuberculosis*. In subjects with latent *M. tuberculosis* infection the expression levels of EMP1 are higher than in subjects not infected with *M. tuberculosis*. Therefore, if these four markers are used together as a biomarker panel a pattern of decreased expression of NXNL1, PSMA7 and C6orf61 and also increased expression of EMP1 (compared to the level expected in a control subject not infected with *M. tuberculosis*) indicates that the subject may have latent *M. tuberculosis* infection.

In the method altered expression level of one, two, three or all of the genes NXNL1, PSMA7, C6orf61 and EMP1 may be evaluated together to differentiate between subjects not infected with *M. tuberculosis* and subjects with latent *M. tuberculosis* infection.

In order to improve the accuracy of the biomarker panel for distinguishing subjects that are not infected with *M. tuberculosis* and subjects with latent *M. tuberculosis* infection one, two or three biomarkers listed in table 3 may be tested in addition to NXNL1, PSMA7, C6orf61 and EMP1.

One, two or three of the following biomarkers may be tested in addition to NXNL1, PSMA7, C6orf61 and EMP1 wherein increased expression of LOC389541 and/or increased expression of MID1IP1 and/or increased expression of KLRC3 and/or increased expression of KLF9 and/or decreased expression GPR117 and/or increased expression of FBXO32 and/or decreased expression of TAZ and/or increased expression of C5ORF29 and/or decreased expression of HSDL1 and/or increased expression of CHUK and/or increased expression of LOC652062 and/or decreased expression of HIP1 and/or increased

expression of C6ORF60 and/or increased expression of MTMR11 indicates that a subject has latent *M. tuberculosis* infection.

The full names and nucleotide sequences of each of the biomarkers useful in distinguishing between subjects that are not infected with *M. tuberculosis* and subjects with latent *M. tuberculosis* infection is shown in figure 3. The right hand column of figure 3 shows whether the biomarker expression is increased (up) or decreased (down) in subjects that have latent *M. tuberculosis* infection.

In the method altered expression level of one or more genes selected from: CLIC1, LACTB and DUSP3 can differentiate between subjects with latent *M. tuberculosis* infection and subjects with active *M. tuberculosis* infection.

Altered expression level of one, two or all of the genes CLIC1, LACTB and DUSP3 may be evaluated together to differentiate between subjects with latent *M. tuberculosis* infection and subjects with active *M. tuberculosis* infection.

In the method altered expression level of one or more genes selected from: CLIC1, LACTB and DUSP3 and optionally one or more genes selected from table 4 can differentiate between subjects with active *M. tuberculosis* infection and subjects with latent *M. tuberculosis* infection. CLIC1, LACTB and DUSP3 and optionally one or more genes selected from table 4 may be used as a biomarker panel to distinguish between subjects with active *M. tuberculosis* infection and subjects with latent *M. tuberculosis* infection with a high degree of accuracy. This biomarker panel is advantageous because it provides a high degree of accuracy with only a small number of biomarkers. The evaluation error of this biomarker panel is 0.26%.

In subjects with active *M. tuberculosis* infection the expression levels of CLIC1, LACTB and DUSP3 are higher than in subjects with latent *M. tuberculosis* infection or subjects not infected with *M. tuberculosis*. Therefore, if these three markers are used together as a biomarker panel a pattern of increased expression of CLIC1, LACTB and DUSP3 (compared to the level expected in a subject with

latent *M. tuberculosis* infection or subjects not infected with *M. tuberculosis*) indicates that the subject may have active *M. tuberculosis* infection.

The sensitivity of the biomarker panel may be increased by additionally testing one or more biomarkers listed in table 4. If the level of any of the biomarkers in table 4 is increased this indicates that the subject may have active *M. tuberculosis* infection.

The full names and nucleotide sequences of each of the biomarkers useful for testing whether a subject has active *M. tuberculosis* infection or latent *M. tuberculosis* infection is shown in figure 4. The right hand column of figure 4 shows whether the biomarker expression is increased (up) or decreased (down) in subjects that have active *M. tuberculosis* infection.

A subject may be tested for expression levels of one, two, three, four, five, six, or all seven biomarkers selected from the group consisting of: NXNL1, PSMA7, C6orf61, EMP1, CLIC1, LACTB and DUSP3. This may allow the *M. tuberculosis* infection status of the subject to be determined as comparing the expression levels of a number of or all of these biomarkers may allow a subject to be classified as a subject not infected with *M. tuberculosis*, a subject with active *M. tuberculosis* infection or a subject with latent *M. tuberculosis* infection.

The altered expression level is an expression level that is higher or lower than the expression level expected in a subject not infected with *M. tuberculosis*. Higher or lower expression level means an expression level that is statistically significantly higher or lower than the control expression level that it is compared to. Statistical significance may be measured using standard statistical methods. Higher or lower expression level may be a statistically significantly higher or lower expression level when the significance is corrected using the number of samples.

The expression level of the selected gene may be determined using the level of mRNA encoded by that gene present in the sample.

The expression level of the selected gene may be determined using the level of mRNA encoded by that gene present in the sample.

The expression level of one or more further genes may also be determined.

According to another aspect the invention provides a kit for use in determining the *M. tuberculosis* infection status in a subject comprising at least one agent for determining the expression level of one or more genes selected from the group consisting of: NXNL1, PSMA7, C6orf61, EMP1, CLIC1, LACTB and DUSP3 in a sample from a subject and instructions for determining the *M. tuberculosis* infection status of the subject. The instructions may include instructions to perform an assay for expression level of one or more of the selected genes. The instructions may provide reference values for comparison with values for the expression level of the selected gene.

The agent may an oligonucleotide, for example an oligonucleotide that adheres to a mRNA encoded by the selected gene. The kit may provide a pair of oligonucleotides suitable for amplifying the selected gene or an mRNA thereof.

The expression level of one or more genes selected from the group consisting of: NXNL1, PSMA7, C6orf61, EMP1, CLIC1, LACTB and DUSP3 may be used as a means to determine the *M. tuberculosis* infection status in a subject.

In another aspect the present invention provides a gene expression product from a gene selected from the group consisting of NXNL1, PSMA7, C6orf61, EMP1, CLIC1, LACTB and DUSP3 or a gene selected from table 3 or table 4 for use as biomarker for infection by *M. tuberculosis*.

In another aspect the present invention provides an oligonucleotide capable of detecting the presence or expression level of a gene expression product from a gene selected from the group consisting of NXNL1, PSMA7, C6orf61, EMP1, CLIC1, LACTB and DUSP3 or a gene selected from table 3 or table 4 in a sample from a subject.

The nucleotide sequences shown for each gene in Figure 1 and Figure 2, Figure 3 and Figure 4 are sections of the nucleotide sequence for each gene that are particularly useful as representative sequences for diagnosis. Therefore it is particularly advantageous to design a pair of primers for identification of each gene that would bind within the sequence given for that gene in Figure 1, Figure 2, Figure 3 or Figure 4. The primers may be designed to amplify part of all of the given nucleotide sequence.

In another aspect the present invention provides a method, kit, use, gene or oligonucleotide as described herein with reference to the examples.

The use of an altered expression level of one or more genes selected from: CLIC1, LACTB and DUSP3 to differentiate between subjects with active *M. tuberculosis* infection and subjects with latent *M. tuberculosis* infection or the use of an altered expression level of one or more genes selected from: NXNL1, PSMA7, C6orf61 and EMP1 to differentiate between subjects not infected with *M. tuberculosis* and subjects with latent *M. tuberculosis* infection has many applications positively impacting on clinical care. It permits both the regular screening of susceptible populations and also the testing of all individuals who are suspected of *M. tuberculosis* infection, and thus enables more timely eradication treatment for initial infection, possibly preventing chronic infection. It also permits a quantitative assessment of the efficacy of antibiotic therapy resulting in interventions that are customised to the response of the individual patient. It also provides a tool for widespread use in epidemiological studies, an important consideration *M. tuberculosis* infection becomes more prevalent.

NXNL1, PSMA7, C6orf61, EMP1, CLIC1, LACTB and DUSP3 and a gene selected from table 3 or table 4 may be targets for development of new therapeutics for treating *M. tuberculosis* infection.

The skilled man will appreciate that preferred features of any one embodiment and/or aspect of the invention may be applied to all other embodiments or aspects of the invention.

The present invention will be further described in more detail, by way of example only, with reference to the following figures in which:

Figure 1 – shows a table of information about the panel of biomarkers that are useful in differentiating between subjects that are not infected with *M. tuberculosis* and subjects with latent *M. tuberculosis* infection, these genes are: Homo sapiens nucleoredoxin-like 1 (NXNL1); Homo sapiens proteasome (prosome, macropain) subunit, alpha type, 7 (PSMA7); Homo sapiens chromosome 6 open reading frame 61 (C6orf61); and Homo sapiens epithelial membrane protein 1 (EMP1). The RNA sequences shown are mRNA sequences.

Figure 2 shows a table of information about the panel of biomarkers that are useful in differentiating between subjects that have active *M. tuberculosis* infection and subjects with latent *M. tuberculosis* infection, these genes are: homo sapiens chloride intracellular channel 1 (CLIC1); Homo sapiens lactamase, beta (LACTB, nuclear gene encoding mitochondrial protein, transcript variant 1; and Homo sapiens dual specificity phosphatise 3 (vaccinia virus phosphatise VH1-related)(DUSP3). The RNA sequences shown are mRNA sequences.

Figure 3 – shows a table of information about the panel of additional biomarkers that are useful in differentiating between subjects that are not infected with *M. tuberculosis* and subjects with latent *M. tuberculosis* infection, column 6 shows a sequence that is useful in identifying the gene and column 8 shows whether the gene is up or down regulated in subjects with latent *M. tuberculosis* infection.

Figure 4 – shows a table of information about the panel of biomarkers that are useful in differentiating between subjects that have active *M. tuberculosis* infection and subjects with latent *M. tuberculosis* infection, column 6 shows a sequence that is useful in identifying the gene and column 8 shows whether the gene is up or down regulated in subjects with active *M. tuberculosis* infection.

Figure 5 – shows a ROC curve for the core set of biomarkers to distinguish between subjects that are not infected with *M. tuberculosis* and subjects with latent *M. tuberculosis* infection (NXNL1, PSMA7, C6orf61 and EMP1), ROC curves show the performance of the classifier incorporating both sensitivity and specificity. The higher the area under the ROC curve the better the performance of the classifier.

Figure 6 – shows a ROC curve for the core set of biomarkers to distinguish between subjects that have active *M. tuberculosis* infection and subjects with latent *M. tuberculosis* infection (CLIC1, LACTB and DUSP3), ROC curves show the performance of the classifier incorporating both sensitivity and specificity. The higher the area under the ROC curve the better the performance of the classifier.

Figure 7 – shows a stepwise summary for the core set of biomarkers to distinguish between subjects that are not infected with *M. tuberculosis* and subjects with latent *M. tuberculosis* infection (NXNL1, PSMA7, C6orf61 and EMP1). Markers are added in a stepwise fashion to build an optimised panel for classification.

Figure 8 – shows a stepwise summary for the core set of biomarkers to distinguish between subjects that have active *M. tuberculosis* infection and subjects with latent *M. tuberculosis* infection (CLIC1, LACTB and DUSP3) Markers are added in a stepwise fashion to build an optimised panel for classification.

Figure 9 – shows response data and curve for the core set of biomarkers to distinguish between subjects that have active *M. tuberculosis* infection and subjects with latent *M. tuberculosis* infection (CLIC1, LACTB and DUSP3) relating the expression level of the marker to the probability of class membership.

Raw data used to produce gene panes was taken from the GENE EXPRESSION OMNIBUS database Code: E-GEOD-22098.

The data was analysed using a data mining algorithm and method both described in patent application number PCT/GB2009/051412 published as: WO2010046697 and claiming priority to GB 0819221.3. This method provided two panels of biomarkers.

1) A panel of biomarkers that are useful in differentiating between subjects that are not infected with *M. tuberculosis* and subjects with latent *M. tuberculosis* infection, these genes are: Homo sapiens nucleoredoxin-like 1 (NXNL1); Homo sapiens proteasome (prosome, macropain) subunit, alpha type, 7 (PSMA7); Homo sapiens chromosome 6 open reading frame 61 (C6orf61); and Homo sapiens epithelial membrane protein 1 (EMP1).

2. A panel of biomarkers that are useful in differentiating between subjects that have active *M. tuberculosis* infection and subjects with latent *M. tuberculosis* infection, these genes are: homo sapiens chloride intracellular channel 1 (CLIC1); Homo sapiens lactamase, beta (LACTB, nuclear gene encoding mitochondrial protein, transcript variant 1; and Homo sapiens dual specificity phosphatise 3 (vaccinia virus phosphatise VH1-related)(DUSP3).

Tables 1 shows the log of mean expression value for each gene in subjects 10 with latent M. tuberculosis infection and Control subjects not infected with M. tuberculosis and the p-value showing the statistical significance of the difference.

Table 1 - Control v Latent

	NXNL1	PSMA7	C6orf61	EMP1
Latent TB	0.93	0.9	0.97	1.50
Control uninfected	1.15	1.29	1.49	1.05
P-value	0.0016	0.0044	$0.8x10^{-5}$	0.045

Tables 2 shows the log of mean expression value for each gene in subjects with latent M. tuberculosis infection and subjects with active M. tuberculosis infection and the p-value showing the statistical significance of the difference.

Table 2 –Latent v Active

	CLIC1	LACTB	DUSP3
Latent TB	0.92	0.90	0.9
Active TB	1.40	1.5	2.2
P-value	$0.4x10^{-5}$	$0.14x10^{-5}$	$0.1x10^{-6}$

Table 3 - Control versus Latent probes of significance

ILMN_Gene	Definition
LOC389541	Homo sapiens similar to CG14977-PA (LOC389541), mRNA
MID1IP1	Homo sapiens MID1 interacting protein 1 (gastrulation specific G12 homolog (zebrafish)) (MID1IP1), mRNA
KLRC3	Homo sapiens killer cell lectin-like receptor subfamily C, member 3 (KLRC3), transcript variant 1, mRNA
KLF9	Homo sapiens Kruppel-like factor 9 (KLF9), mRNA
GPR177	Homo sapiens G protein-coupled receptor 177 (GPR177), transcript variant 1, mRNA
FBXO32	Homo sapiens F-box protein 32 (FBXO32), transcript variant 2, mRNA.
TAZ	Homo sapiens tafazzin (cardiomyopathy, dilated 3A (X-linked); endocardial fibroelastosis 2; Barth syndrome) (TAZ), transcript variant 3, mRNA
C5ORF29	Homo sapiens chromosome 5 open

	reading frame 29 (C5orf29), mRNA
HSDL1	Homo sapiens hydroxysteroid dehydrogenase like 1 (HSDL1), mRNA.
CHUK	Homo sapiens conserved helix-loop-helix ubiquitous kinase (CHUK), mRNA
LOC652062	PREDICTED: Homo sapiens similar to Mitochondrial carnitine/acylcarnitine carrier protein (Carnitine/acylcarnitine translocase) (CAC) (LOC652062), mRNA
HIP1	Homo sapiens huntingtin interacting protein 1 (HIP1), mRNA
C6ORF60	Homo sapiens chromosome 6 open reading frame 60 (C6orf60), transcript variant 1, mRNA
MTMR11	Homo sapiens myotubularin related protein 11 (MTMR11), transcript variant 1, mRNA.

Table 4 - Active versus Latent probes of significance

ILMN_Gene	Definition
LACTB	Homo sapiens lactamase, beta (LACTB), nuclear gene encoding mitochondrial protein, transcript variant 1, mRNA.
SRBD1	Homo sapiens S1 RNA binding domain 1 (SRBD1), mRNA.
DUSP3	Homo sapiens dual specificity phosphatase 3 (vaccinia virus phosphatase VH1-related) (DUSP3), mRNA.
ATG3	Homo sapiens ATG3 autophagy related 3 homolog (S. cerevisiae) (ATG3),

	mRNA.
JAK2	Homo sapiens Janus kinase 2 (a protein tyrosine kinase) (JAK2), mRNA.
PSMB8	Homo sapiens proteasome (prosome, macropain) subunit, beta type, 8 (large multifunctional peptidase 7) (PSMB8), transcript variant 2, mRNA.
PSME1	Homo sapiens proteasome (prosome, macropain) activator subunit 1 (PA28 alpha) (PSME1), transcript variant 1, mRNA.
ACOT9	Homo sapiens acyl-CoA thioesterase 9 (ACOT9), transcript variant 2, mRNA.
IFI30	Homo sapiens interferon, gamma-inducible protein 30 (IFI30), mRNA.
SORT1	Homo sapiens sortilin 1 (SORT1), mRNA.
GSR	Homo sapiens glutathione reductase (GSR), mRNA.
TAP1	Homo sapiens transporter 1, ATP-binding cassette, sub-family B (MDR/TAP) (TAP1), mRNA.
GRN	Homo sapiens granulin (GRN), mRNA.
IRF1	Homo sapiens interferon regulatory factor 1 (IRF1), mRNA.

Sequences used in the application

Table 5 –gene names and SEQ ID NO

ILMN Gene (names shown in figures 3 and 4)	SEQ ID NO	Sequence used to identify the gene
NXNL1	1	GAGACCCTGACTCTACGAAAATTAAAAGTTAGCCCGGTGTGGTGGCGCGC
PSMA7	2	CGCTTGCATGCTCACCTCTGGCAGCAGGGC

		AGTCACGGCTCCGCCATGGA
C6ORF61	3	GGCAGGATGATTGCTTAAGCCCAGGAGTTC GGGGTTACAGTGAGCTGTGG
EMP1	4	GAGGGCAAGCCACCAAATTACCTAGGCTGA GGTTAGAGAGATTGGCCAGC
CLIC1	5	GAGCTCGCCTATGAGCAAGTGGCAAAGGCC CTCAAATAAGCCCCTCCTGG
LACTB	6	ATACTGGAGGGGCAGTGGGTGCCAGTAGTG TCCTGCTGGTCCTTCCTGAA
DUSP3	7	CCATGGTGATGGATGGTTTGGAAAGGGAAT GTTGGTGCCTTTTGTGCCAC
LOC389541	8	TGTGGTGAAGAGGCAGAACCGAGGTCGGG AGCCCATTGATGTCTGAGCCT
MID1IP1	9	CCCCAGTGTGTATAAGCTGGCATTTCGCCA GCTTGTACGTAGCTTGCCAC
KLRC3	10	GGGAAGAGAGTTTGCAGGCCTGTGCTTCAA AGAACTCTTCTAGTCTGCTT
KLF9	11	GCCCTTCACCATTGTGGAATGATGCCCTGG CTTTAAGGTTTAGCTCCACA
GPR177	12	TGAGATCTACAAGTTGACCCGCAAGGAGGC CCAGGAGTAGGAGGCTGCAG
FBXO32	13	GAAGGGTCCCCCTGCTGACTGGAGAGCTGG GAATATGGCATTTGGACACT
TAZ	14	GACAGATTTGTTCATAGACCCTCTCAAGTG CCCTCTCCGAGCTGGTAGGC
C5ORF29	15	GCCTCTTCTCTCAAGCCTGCTTCAGATCATA AGTTCTTCCACACATCTCC
HSDL1	16	GTGCCGAGCTGTCCATAGCTGCAGTGAAAG GTGAAGAGCAAGACCTTCTC
C6ORF60	17	GCTTCTCCAGATCCCCAGCGCCAGGAGTGG TTTGCCCGGTACTTCACATT
CHUK	18	CCTTTATTTTGCTGCTTGATGATGAGAGGG AGGGCTGCTGCCACAGACTG
LOC652062	19	ATGGAGCTGTCTTTCAGATCTTTCCTGGGAT

		TACCCCTGCCTACCCCCAG
HIP1	20	TGCAGCCGTCCATAGCAGTACCCCTAAAAT CCCACCAGAATACGGGTCCC
MTMR11	21	CACCAGAAGTAGCAGAGAAGCAGGGGGCC AGAGCTACAACAGTATTCTTC
SRBD1	22	ACTTCTACTTGCCAACATCTGCCTTGCTGGA CTTGTATGGGATTGTCTCC
ATG3	23	AGTGACCATTGAAAATCACCCTCATCTGCC ACCACCTCCCATGTGTTCAG
JAK2	24	TTGTCATCCTTTGAGCTGCTGACTGCCAATA ACATTCTTCGATCTCTGGG
PSMB8	25	GTCGCTCGGACCCAGGACACTACAGTTTCT CTATGCGATCTCCAGAGCTC
PSME1	26	ACCGGGACATCCGGCTGATGGTCATGGAGA TCCGCAATGCTTATGCTGTG
ACOT9	27	GGACATTAAGTTCAGTGGCCATGTTAGCTG GGTCGGGAAGACATCCATGG
IFI30	28	TGGAAGATCAGACCCAGCTCCTTACCCTTG TCTGCCAGTTGTACCAGGGC
LACTB	29	ATACTGGAGGGGCAGTGGGTGCCAGTAGTG TCCTGCTGGTCCTTCCTGAA
DUSP3	30	CCATGGTGATGGATGGTTTGGAAAGGGAAT GTTGGTGCCTTTTGTGCCAC
SORT1	31	GGTCCCCATGTGCCTGTTGTTCAGCCCTCTC TCTTGTTCCCTTTCTGAGC
GSR	32	GAACCAGGAGACACGTGTGGCGGGCAGTG GGACCCATAGATCTTCTGAAA
TAP1	33	GTAACGGAGTTTAGAGCCAGGGCTGATGCT TTGGTGTGGCCAGCACTCTG
GRN	34	AAGGCTCGATCCTGCGAGAAGGAAGTGGTC TCTGCCCAGCCTGCCACCTT
IRF1	35	CCTCAACAGGCCCAGGGAGGGAAGTGTGA GCGCCTTGGTATGACTTAAAA

CLAIMS

1. A method of determining the *M. tuberculosis* infection status of a subject comprising:

(a) providing a sample of material obtained from the subject; and

(b) determining the expression level of one or more genes selected from the group consisting of: NXNL1, PSMA7, C6orf61, EMP1, CLIC1, LACTB and DUSP3 in the sample.

2. The method of claim 1 further comprising the step of:

(c) determining the expression level of one or more genes listed in table 3 or table 4.

3. The method of claim 1 or claim 2 further comprising:

(d) employing the expression level determined in (b) and optionally (c) to distinguish between subjects not infected with *M. tuberculosis*, subjects with latent *M. tuberculosis* infection and subjects with active *M. tuberculosis* infection.

4. The method of any one of the preceding claims for use in one or more of the following; diagnosing whether or not a subject has *M. tuberculosis* infection; advising on the prognosis for a subject with a *M. tuberculosis* infection; and monitoring the effectiveness or response of a subject to a particular treatment for infection by *M. tuberculosis*.

5. The method according to any one of the preceding claims wherein altered expression level of one or more genes selected from: NXNL1, PSMA7, C6orf61 and EMP1 can differentiate between subjects not infected with *M. tuberculosis* and subjects with latent *M. tuberculosis* infection.

6. The method according to any one of the preceding claims wherein decreased expression level of the genes NXNL1, PSMA7 and C6orf61 and

increased expression level of the gene EMP1 indicates that a subject has latent *M. tuberculosis* infection.

7. The method according to any one of the preceding claims wherein altered expression level of one or more genes selected from NXNL1, PSMA7, C6orf61, EMP1 and at least one gene selected from table 3 can differentiate between subjects not infected with *M. tuberculosis* and subjects with latent *M. tuberculosis* infection.

8. The method according to any one of the preceding claims wherein altered expression level of one or more genes selected from NXNL1, PSMA7, C6orf61, EMP1 and one, two or three genes selected from table 3 can differentiate between subjects not infected with *M. tuberculosis* and subjects with latent *M. tuberculosis* infection.

9. The method according to any one of the preceding claims wherein decreased expression level of the genes NXNL1, PSMA7 and C6orf61 and increased expression level of the gene EMP1 and altered expression level of one, two or three genes selected from table 3 indicates that a subject has latent *M. tuberculosis* infection.

10. The method according to claim 9 wherein increased expression of LOC389541 and/or increased expression of MID1IP1 and/or increased expression of KLRC3 and/or increased expression of KLF9 and/or decreased expression GPR117 and/or increased expression of FBXO32 and/or decreased expression of TAZ and/or increased expression of C5ORF29 and/or decreased expression of HSDL1 and/or increased expression of CHUK and/or increased expression of LOC652062 and/or decreased expression of HIP1 and/or increased expression of C6ORF60 and/or increased expression of MTMR11 indicates that a subject has latent *M. tuberculosis* infection.

11. The method according to any one of claims 1 to 4 wherein altered expression level of one or more genes selected from: CLIC1, LACTB and DUSP3

can differentiate between subjects with latent *M. tuberculosis* infection and subjects with active *M. tuberculosis* infection.

12. The method according to any one of claims 1 to 4 wherein altered expression level of all of the genes CLIC1, LACTB and DUSP3 can differentiate between subjects with latent *M. tuberculosis* infection and subjects with active *M. tuberculosis* infection.

13. The method according to any one of claims 1 to 4 wherein increased expression level of all of the genes CLIC1, LACTB and DUSP3 indicates a subject with active *M. tuberculosis* infection.

14. The method according to any one of claims 1 to 4 wherein increased expression level of all of the genes CLIC1, LACTB and DUSP3 and increased expression level of one, two or three genes selected from table 4 indicates a subject with active *M. tuberculosis* infection.

15. The method according to any one of the preceding claims wherein the altered expression level is an expression level that is higher or lower than the expression level expected in a subject not infected with *M. tuberculosis.*

16. A method according to any one of the preceding claims wherein the expression level of the gene is determined using the level of mRNA encoded by the selected gene present in the sample.

17. A method according to any one of the preceding claims wherein the expression level of the gene is determined using quantitative PCR.

18. A method according to any one of the preceding claims wherein the expression level of one or more further genes is also determined.

19. A kit for use in determining the *M. tuberculosis* infection status in a subject comprising at least one agent for determining the expression level of one

or more genes selected from the group consisting of: NXNL1, PSMA7, C6orf61, EMP1, CLIC1, LACTB and DUSP3 in a sample from a subject and instructions for determining the *M. tuberculosis* infection status of the subject.

20. The kit of claim 19 wherein the agent is an oligonucleotide.

21. Use of the determination of the expression level of one or more genes selected from the group consisting of: NXNL1, PSMA7, C6orf61, EMP1, CLIC1, LACTB and DUSP3 as a means to determine the *M. tuberculosis* infection status in a subject.

22. A gene expression product from a gene selected from the group consisting of NXNL1, PSMA7, C6orf61, EMP1, CLIC1, LACTB and DUSP3 for use as biomarker for infection by *M. tuberculosis*.

23. An oligonucleotide capable of detecting the presence or expression level of a gene expression product from a gene selected from the group consisting of NXNL1, PSMA7, C6orf61, EMP1, CLIC1, LACTB and DUSP3 in a sample from a subject.

24. A method, kit, use, gene or oligonucleotide as described herein with reference to the example.

Gene	ID	Species	Source	Search Key	Transcript	ILMN Gene	Source Reference ID	RefSeq ID	Unigene ID	Entrez Gene ID	GI
NXNL1	ILMN_1742917	Homo sapiens	RefSeq	ILMN_12581	ILMN_12581	NXNL1	NM_138454.1	NM_138454.1		115861	19923986
PSMA7	ILMN_1701962	Homo sapiens	RefSeq	ILMN_13260	ILMN_13260	PSMA7	NM_152255.1	NM_152255.1		5688	23110947
C6orf61	ILMN_1680867	Homo sapiens	RefSeq	ILMN_43474	ILMN_43474	C6ORF61	XM_927492.1	XM_927492.1		54844	88997736
EMP1	ILMN_1801616	Homo sapiens	RefSeq	ILMN_25097	ILMN_163266	EMP1	NM_001423.1	NM_001423.1		2012	4503558

Figure 1 a

Gene	Accession	Symbol	Protein Product	Array Address Id	Probe Type	Probe Start	SEQUENCE	Chromosome	Probe Chr Orientation	Probe Coordinates	Cytoband
NXNL1	NM_138454.1	NXNL1	NP_612463.1	2480180	S	836	GAGACCCTGACTCTACGAAAATTAAAAGTTAGCCCGGTGTGGTGGCGCGC	19	-	17427256-17427305	19p13.11d
PSMA7	NM_152255.1	PSMA7	NP_689468.1	5050279	I	477	CGCTTGCATGCTCACCTCTGGCAGCAGGGCAGTCACGGCTCCGCCATGGA	20	-	60148168-60148217	20q13.33c
C6orf61	XM_927492.1	C6orf61	XP_932585.1	3290255	S	2997	GGCAGGATGATTGCTTAAGCCCAGGAGTTCGGGGTTACAGTGAGCTGTGG	6	-	119241824-119241873	
EMP1	NM_001423.1	EMP1	NP_001414.1	3940435	S	2380	GAGGGCAAGCCACCAAATTACCTAGGCTGAGGTTAGAGAGATTGGCCAGC	12	+	13260606-13260655	12p13.1b

Figure 1b

Gene	Definition	Ontology Component	Ontology Process	Ontology Function	Synonyms	Obsolete Probe Id	GB ACC
NXNL1	Homo sapiens nucleoredoxin-like 1 (NXNL1), mRNA.		electron transport [goid 6118] [evidence IEA]	electron carrier activity [goid 9055] [evidence IEA]; protein disulfide oxidoreductase activity [goid 15035] [evidence IEA]	TXNL6; RDCVF	RDCVF	NM_138454.1
PSMA7	Homo sapiens proteasome (prosome, macropain) subunit, alpha type, 7 (PSMA7), transcript variant 2, mRNA.	proteasome core complex (sensu Eukaryota) [goid 5839] [evidence IEA]; cytosol [goid 5829] [evidence IEA]; protein complex [goid 43234] [evidence IEA]	ubiquitin-dependent protein catabolism [goid 6511] [evidence IEA]	threonine endopeptidase activity [goid 4298] [evidence IEA]	MGC3755; HSPC; C6; XAPC7; RC6-1	MGC3755; HSPC; C6; XAPC7; RC6-1	NM_152255.1
C6orf61	PREDICTED: Homo sapiens chromosome 6 open reading frame 61 (C6orf61), mRNA.						XM_927492.1
EMP1	Homo sapiens epithelial membrane protein 1 (EMP1), mRNA.	membrane fraction [goid 5624] [pmid 8996089] [evidence TAS]; membrane [goid 16020] [evidence IEA]; integral to membrane [goid 16021] [pmid 8884260] [evidence TAS]	cell proliferation [goid 8283] [pmid 8996089] [evidence TAS]; development [goid 7275] [pmid 8996089] [evidence TAS]; cell growth [goid 16049] [evidence IEA]; epidermis development [goid 8544] [pmid 7499420] [evidence TAS]; cell death [goid 8219] [evidence NR]		CL-20; EMP-1; TMP	EMP-1; CL-20; TMP	NM_001423.1

Figure 1c

ID	Species	Source	Search Key	Transcript	ILMN Gene	Source Reference ID	RefSeq_ID	Unigene ID	Entrez Gene ID	GI
ILMN_1756982	Homo sapiens	RefSeq	ILMN_29900	ILMN_29900	CLIC1	NM_001288.4	NM_001288.4		1192	48375182
ILMN_1703335	Homo sapiens	RefSeq	ILMN_24565	ILMN_24565	LACTB	NM_032857.2	NM_032857.2		114294	26051230
ILMN_1797522	Homo sapiens	RefSeq	ILMN_14523	ILMN_180655	DUSP3	NM_004090.2	NM_004090.2		1845	37655179

Figure 2a

ID	Symbol	Protein Product	Array Address Id	Probe Type	Probe Start	SEQUENCE
ILMN_1756982	CLIC1	NP_001279.2	3450072	S	952	GAGCTCGCCTATGAGCAAGTGGCAAAGGCCCTCAAATAAGCCCCTCCTGG
ILMN_1703335	LACTB	NP_116246.2	1570669	I	1493	ATACTGGAGGGGCAGTGGGTGCCAGTAGTGTCCTGCTGGTCCTTCCTGAA
ILMN_1797522	DUSP3	NP_004081.1	6560156	S	3634	CCATGGTGATGGATGGTTTGGAAAGGGAATGTTGGTGCCTTTTGTGCCAC

Figure 2b

ID	Chromosome	Probe Chr Orientation	Probe Coordinates	Cytoband	Definition
ILMN_1756982	6	-	31806586-31806597: 31806598-31806635	6p21.33a	Homo sapiens chloride intracellular channel 1 (CLIC1), mRNA.
ILMN_1703335	15	+	40256913-40256962	15q22.2b	Homo sapiens lactamase, beta (LACTB), nuclear gene encoding mitochondrial protein, transcript variant 1, mRNA.
ILMN_1797522	17	-	37608003-37608052	17q21.31b	Homo sapiens dual specificity phosphatase 3 (vaccinia virus phosphatase VH1-related) (DUSP3), mRNA.

Figure 2c

ID	Ontology Component	Ontology Function	Synonyms	Obsolete Probe Id	GB_ACC
ILMN_1756982	membrane fraction [goid 5624] [pmid 10793131] [evidence IDA]; nucleus [goid 5634] [pmid 16130169] [evidence TAS]; membrane [goid 16020] [evidence IEA]; brush border [goid 5903] [pmid 10793131] [evidence TAS]; soluble fraction [goid 5625] [pmid 10793131] [evidence IDA]; cytoplasm [goid 5737] [pmid 10793131] [evidence IDA]	voltage-gated chloride channel activity [goid 5247] [evidence IEA]; protein binding [goid 5515] [pmid 14667819] [evidence IPI]	NCC27; G6	NCC27; G6	NM_001288.4
ILMN_1703335	membrane [goid 16020] [evidence IEA]; integral to membrane [goid 16021] [evidence IEA]	beta-lactamase activity [goid 8800] [evidence IEA]; hydrolase activity [goid 16787] [evidence IEA]	FLJ14902; G24; MRPL56	FLJ14902; G24; MRPL56	NM_032857.2
ILMN_1797522		hydrolase activity [goid 16787] [evidence IEA]; protein tyrosine phosphatase activity [goid 4725] [pmid 1281549] [evidence TAS]; protein tyrosine/serine/threonine phosphatase activity [goid 8138] [evidence IEA]	VHR	VHR	NM_004090.2

Figure 2d

ID	ILMN_Gene	Source_Reference_ID	RefSeq_ID	Protein_Product	SEQUENCE	Definition	Expression in latent cf control
ILMN_1742917	NXNL1	NM_138454.1	NM_138454.1	NP_612463.1	GAGACCCTGACTCTACGAAAATTAAAAGTTAGCCCGGTGTGGTGGCGCGC	Homo sapiens nucleoredoxin-like 1 (NXNL1), mRNA.	down
ILMN_1779735	LOC389541	NM_001008395.2	NM_001008395.2	NP_001008396.1	TGTGGTGAAGAGGCAGAACCGAGGTCGGGAGCCCATTGATGTCTGAGCCT	Homo sapiens similar to CG14977-PA (LOC389541), mRNA.	up
ILMN_2165478	MID1IP1	NM_021242.3	NM_021242.3	NP_067065.1	CCCCAGTGTGTATAAGCTGGCATTTCGCCAGCTTGTACGTAGCTTGCCAC	Homo sapiens MID1 interacting protein 1 (gastrulation specific G12 homolog (zebrafish)) (MID1IP1), mRNA.	up
ILMN_2386790	KLRC3	NM_002261.2	NM_002261.2	NP_002252.2	GGGAAGAGAGTTTGCAGGCCTGTGCTTCAAAGAACTCTTCTAGTCTGCTT	Homo sapiens killer cell lectin-like receptor subfamily C, member 3 (KLRC3), transcript variant 1, mRNA.	up
ILMN_1778523	KLF9	NM_001206.2	NM_001206.2	NP_001197.1	GCCCTTCACCATTGTGGAATGATGCCCTGGCTTTAAGGTTTAGCTCCACA	Homo sapiens Kruppel-like factor 9 (KLF9), mRNA.	up
ILMN_1753813	GPR177	NM_024911.4	NM_024911.4	NP_079187.3	TGAGATCTACAAGTTGACCCGCAAGGAGGCCCAGGAGTAGGAGGCTGCAG	Homo sapiens G protein-coupled receptor 177 (GPR177), transcript variant 1, mRNA.	down
ILMN_1703955	FBXO32	NM_148177.1	NM_148177.1	NP_680482.1	GAAGGGTCCCCCTGCTGACTGGAGAGCTGGGAATATGGCATTTGGACACT	Homo sapiens F-box protein 32 (FBXO32), transcript variant 2, mRNA.	up
ILMN_2358733	TAZ	NM_181312.1	NM_181312.1	NP_851829.1	GACAGATTTGTTCATAGACCCTCTCAAGTGCCCTCTCCGAGCTGGTAGGC	Homo sapiens tafazzin (cardiomyopathy, dilated 3A (X-linked); endocardial fibroelastosis 2; Barth syndrome) (TAZ), transcript variant 3, mRNA.	down
ILMN_1675191	C5ORF29	NM_152687.2	NM_152687.2	NP_689900.1	GCCTCTTCTCTCAAGCCTGCTTCAGATCATAAGTTCTTCCACACATCTCC	Homo sapiens chromosome 5 open reading frame 29 (C5orf29), mRNA.	up
ILMN_2112755	HSD11	NM_031463.3	NM_031463.3	NP_113651.3	GTGCCGAGCTGTCCATAGCTGCAGTGAAAGGTGAAGAGCAAGACCTTCTC	Homo sapiens hydroxysteroid dehydrogenase like 1 (HSD11), mRNA.	down
ILMN_1701962	PSMA7	NM_152255.1	NM_152255.1	NP_689468.1	CGCTTGCATGCTCACCTCTGGCAGCAGGGCAGTCACGGCTCCGCCATGGA	Homo sapiens proteasome (prosome, macropain) subunit, alpha type, 7 (PSMA7), transcript variant 2, mRNA.	down
ILMN_1680867	C6ORF61	XM_927492.1	XM_927492.1	XP_932585.1	GGCAGGATGATTGCTTAAGCCCAGGAGTTCGGGGTTACAGTGAGCTGTGG	PREDICTED: Homo sapiens chromosome 6 open reading frame 61 (C6orf61), mRNA.	up
ILMN_1801616	EMP1	NM_001423.1	NM_001423.1	NP_001414.1	GAGGGCAAGCCACCAAATTACCTAGGCTGAGGTTAGAGAGATTGGCCAGC	Homo sapiens epithelial membrane protein 1 (EMP1), mRNA.	down
ILMN_1677041	CHUK	NM_001278.3	NM_001278.3	NP_001269.3	CCTTTATTTTGCTGCTTGATGATGAGAGGGAGGGCTGCTGCCACAGACTG	Homo sapiens conserved helix-loop-helix ubiquitous kinase (CHUK), mRNA.	up
ILMN_1808991	LOC652062	XM_941382.1	XM_941382.1	XP_946475.1	ATGGAGCTGTCTTTCAGATCTTTCCTGGGATTACCCCTGCCTACCCCCAG	PREDICTED: Homo sapiens similar to Mitochondrial carnitine/acylcarnitine carrier protein (Carnitine/acylcarnitine translocase) (CAC) (LOC652062), mRNA.	up
ILMN_1701403	HIP1	NM_005338.4	NM_005338.4	NP_005329.3	TGCAGCCGTCCATAGCAGTACCCCTAAAATCCCACCAGAATACGGGTCCC	Homo sapiens huntingtin interacting protein 1 (HIP1), mRNA.	down
ILMN_1696699	C6ORF60	NM_024581.4	NM_024581.4	NP_078857.5	GCTTCTCCAGATCCCCAGCGCCAGGAGTGGTTTGCCCGGTACTYCACATT	Homo sapiens chromosome 6 open reading frame 60 (C6orf60), transcript variant 1, mRNA.	up
ILMN_1698246	MTMR11	NM_006697.1	NM_006697.1	NP_006688.1	CACCAGAAGTAGCAGAGAAGCAGGGGGCCAGAGCTACAACAGTATTCTTC	Homo sapiens myotubularin related protein 11 (MTMR11), transcript variant 1, mRNA.	up

Figure 3

ID	HMM_Gene	Source_Reference ID	RefSeq_ID	Protein_Product	SEQUENCE	Definition	Regulation wrt Active
HMM_1756982	CLIC1	NM_001288.4	NM_001279.2	NP_001270.2	GAGCTCGGCCTATGAGCAAGTGGCAAAGGCCCTGAAATAAGCCCCTCCTGG	Homo sapiens chloride intracellular channel 1 (CLIC1), mRNA.	up
HMM_1703335	LACTB	NM_032857.2	NM_032857.2	NP_116246.2	ATACTGGAGGGGGCAGTGGGTGCCAGTAGTGTCCTGCTGGTCCTTCCTGAA	Homo sapiens lactamase, beta (LACTB), nuclear gene encoding mitochondrial protein, transcript variant 1, mRNA.	up
HMM_1756827	SRBD1	NM_018079.3	NM_060549.3	NP_060549.3	ACTTCTACTTGCCAACATCTGCCTTGCTGGACTTGTATGGGATTGTCTCC	Homo sapiens S1 RNA binding domain 1 (SRBD1), mRNA.	up
HMM_1797522	DUSP3	NM_004090.2	NM_004081.1	NP_004081.1	CCATGGTGATGGATGGTTTGGAAAGGGAATGTTGGTGCCTTTGTGCCAC	Homo sapiens dual specificity phosphatase 3(vaccinia virus phosphatase VH1-related) (DUSP3), mRNA.	up
HMM_1768566	ATG3	NM_022488.3	NM_022488.3	NP_071931.2	AGTGACCATTGAAAATCACCCTCATCTGCCACCACCTCCCATGTGTTCAG	Homo sapiens ATG3 autophagy related 3 homolog (S. cerevisiae) (ATG3), mRNA.	up
HMM_1683178	JAK2	NM_004972.2	NM_004963.1	NP_004963.1	TTGTCATCCTTTGAGCCTGCTGACTGCCAATAACATTCTTCGATCTCTGGG	Homo sapiens Janus kinase 2 (a protein tyrosine kinase) (JAK2), mRNA.	up
HMM_1767096	PSMB8	NM_148919.3	NM_148919.3	NP_683720.2	GTCGCTCGGGACCCAGGGACACTACAGTTCTCTATGCGATCTCCAGAGGTC	Homo sapiens proteasome (prosome, macropain) subunit, beta type, 8(large multifunctional peptidase 7) (PSMB8), transcript variant 2, mRNA.	up
HMM_1805017	PSME1	NM_006263.2	NM_006263.2	NP_006254.1	ACCGGGGACATCCGGGCTGATGGTCATGGAGATCCGCAATGCTTATGCTGTG	Homo sapiens proteasome (prosome, macropain) activator subunit 1 (PA28 alpha) (PSME1), transcript variant 1, mRNA.	up
HMM_1638995	ACOT9	NM_001033583.2	NM_001033583.2	NP_001028755.2	GGACATTAAGTTCAGTGGCCATGTTAGCTGGGTGGGGAAGAATCATCCATGG	Homo sapiens acyl-CoA thioesterase 9 (ACOT9), transcript variant 2, mRNA	up
HMM_1807277	IFI30	NM_006332.3	NM_006323.2	NP_006323.2	TGGGAAGATGAGACCCAGCTCCTTACCCTTGTCTGCCAGTTGTACCAGGGC	Homo sapiens interferon, gamma-inducible protein 30 (IFI30), mRNA.	up
HMM_1703315	LACTB	NM_032857.2	NM_032857.2	NP_116246.2	ATACTGGAGGGGGCAGTGGGTGCCAGTAGTGTCCTGCTGGTCCTTCCTGAA	Homo sapiens lactamase, beta (LACTB), nuclear gene encoding mitochondrial protein, transcript variant 1, mRNA.	up
HMM_1797522	DUSP3	NM_004090.2	NM_004081.1	NP_004081.1	CCATGGTGATGGATGGTTTGGAAAGGGAATGTTGGTGCCTTTGTGCCAC	Homo sapiens dual specificity phosphatase 3 (vaccinia virus phosphatase VH1-related) (DUSP3), mRNA.	up
HMM_1707077	SORT1	NM_002959.4	NM_002959.3	NP_002950.3	GGTCCCCATGTGCCTGTGTTCAGCCCTCTCTGTTCCTCTTCTGAGC	Homo sapiens sortilin 1 (SORT1), mRNA.	up
HMM_1775182	GSR	NM_000637.2	NM_000628.2	NP_000628.2	GAACCAGGAGACACGTGTGGCGGCAGTGGGACCCATAGATCTTCTGAAA	Homo sapiens glutathione reductase (GSR), mRNA.	up
HMM_1751679	TAP1	NM_000593.5	NM_000593.5	NP_000584.2	GTAACGGAGTTTAGAAGCCAGGGCTGATGCTTTGGTGTGGCCAGCACTCTG	Homo sapiens transporter 1, ATP-binding cassette, sub-family B (MDR/TAP) (TAP1), mRNA.	up
HMM_1724258	GRN	NM_002087.2	NM_002087.2	NP_002078.1	AAGGCTCGATCCTGCGAAGAAGGAAGTGGTCTCTGCCCAGGCTTGCCACCTT	Homo sapiens granulin (GRN), mRNA.	up
HMM_1703375	IRF1	NM_002198.1	NM_002189.1	NP_002180.1	CCTCAACAGGCCCAGGGAAGGGAAGTGTGAGCGCCTTTGGTATGACTTAAAA	Homo sapiens interferon regulatory factor 1 (IRF1), mRNA.	up

Figure 4

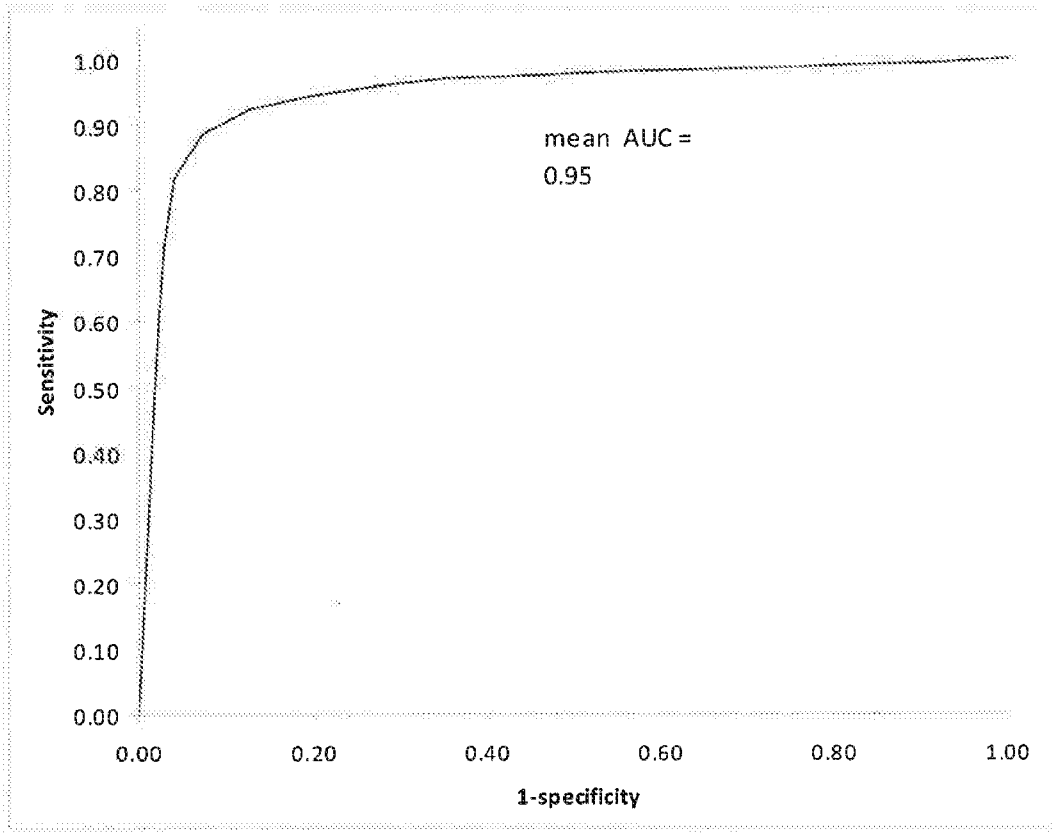
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Sensitivity
mean AUC =
0.95
0.00
0.20
0.40
0.60
0.80
1.00
1-specificity

Figure 5

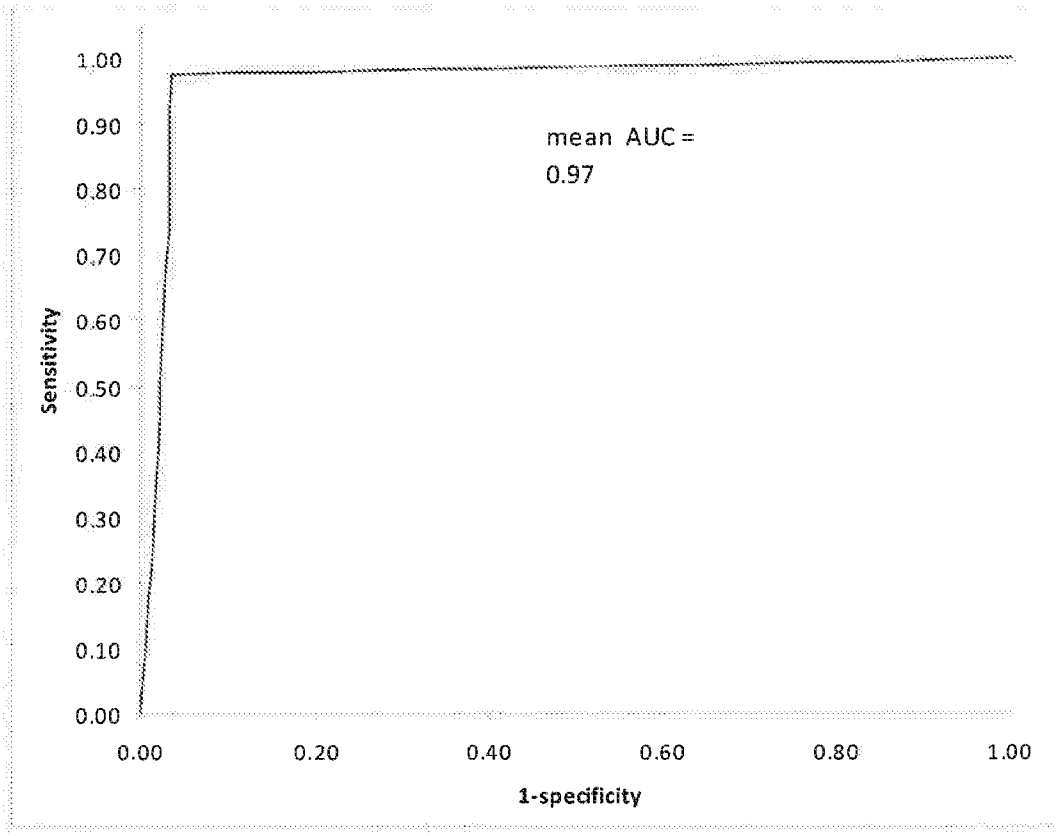
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Sensitivity
mean AUC = 0.97
0.00
0.20
0.40
0.60
0.80
1.00
1-specificity

Figure 6

	Mean 1	Mean 0	t-value	df	p	Valid N 1	Valid N 0	Std.Dev. 1	Std.Dev. 0	F-ratio Variances	p Variances
ILMN_1680867	0.932032	1.150736	3.50969	27	0.001593	17	12	0.142096	0.194108	1.86606	0.249162
ILMN_1701962	0.898495	1.294530	3.10941	27	0.004387	17	12	0.346869	0.324178	1.14489	0.838257
ILMN_1742917	0.965240	1.488988	5.51248	27	0.000008	17	12	0.152273	0.349487	5.26766	0.002996
ILMN_1801616	1.497231	1.050381	2.09965	27	0.045242	17	12	0.729113	0.093825	60.38778	0.000000

Figure 7A

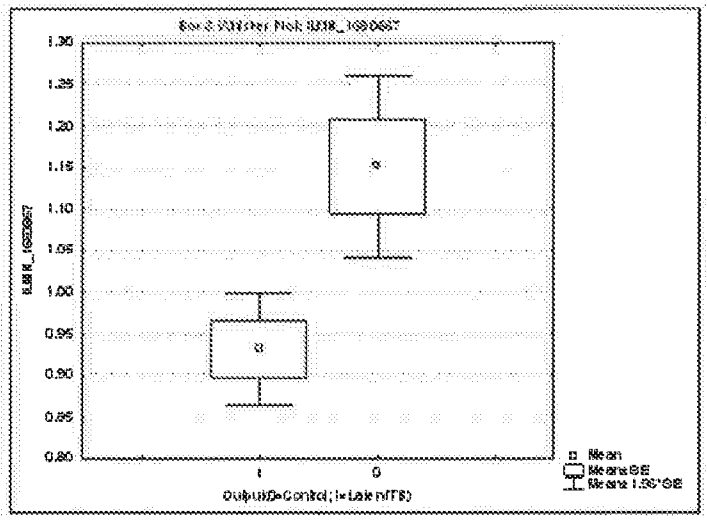

Figure 7B

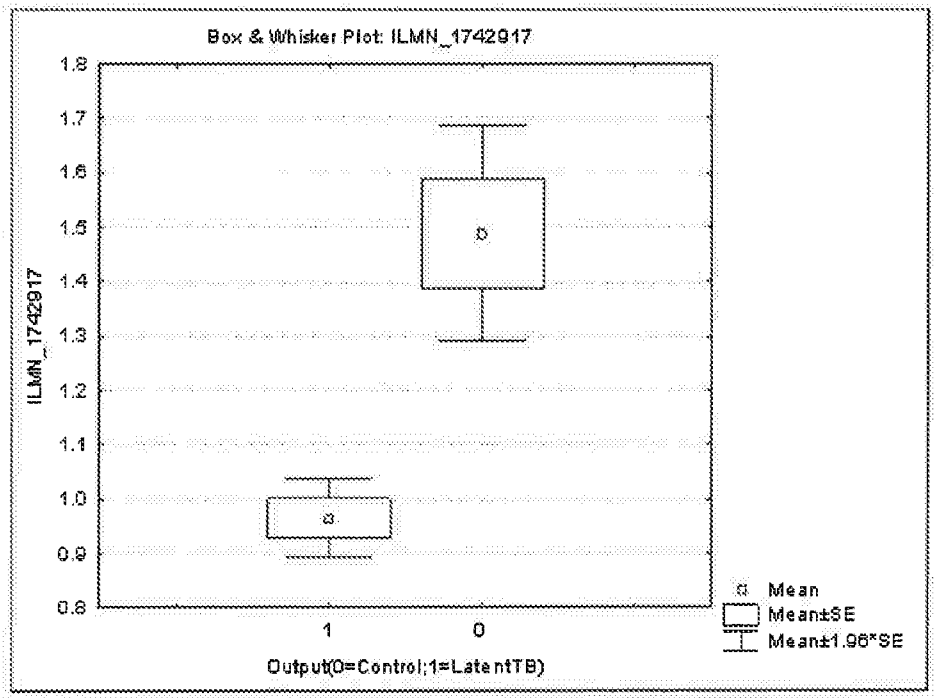

Figure 7C

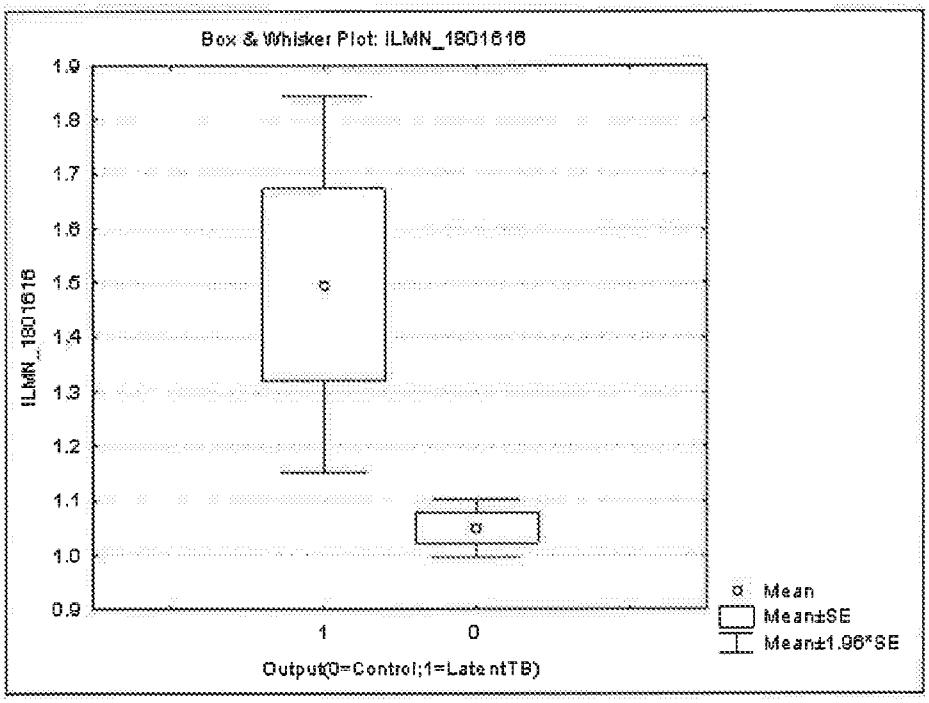

Figure 7D

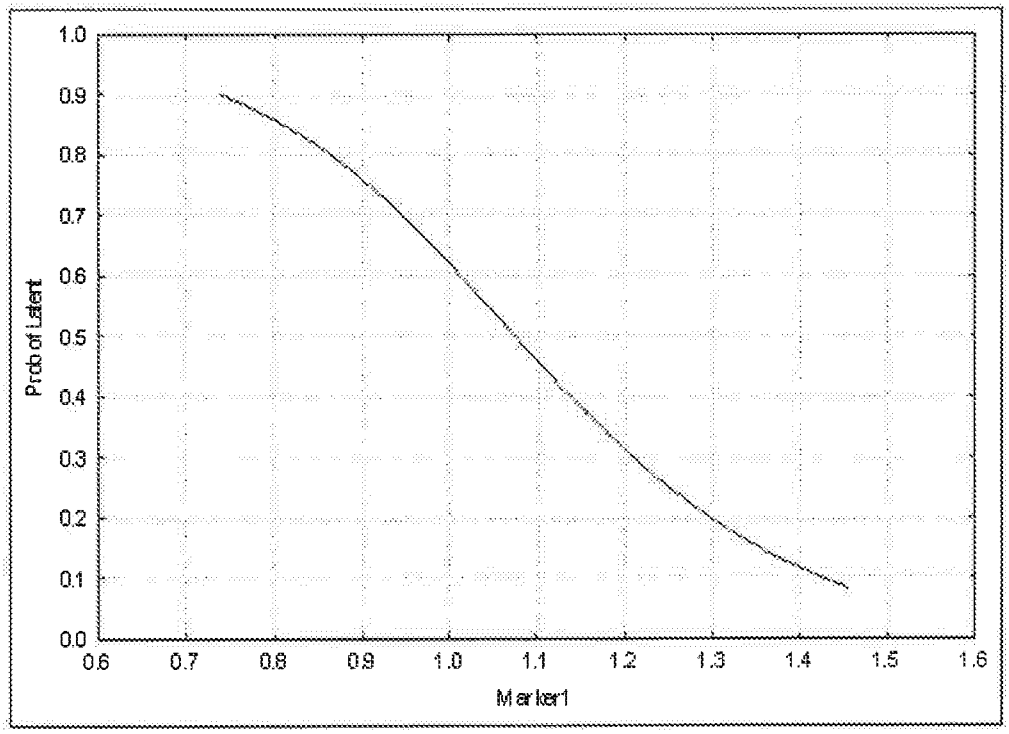

Figure 7E

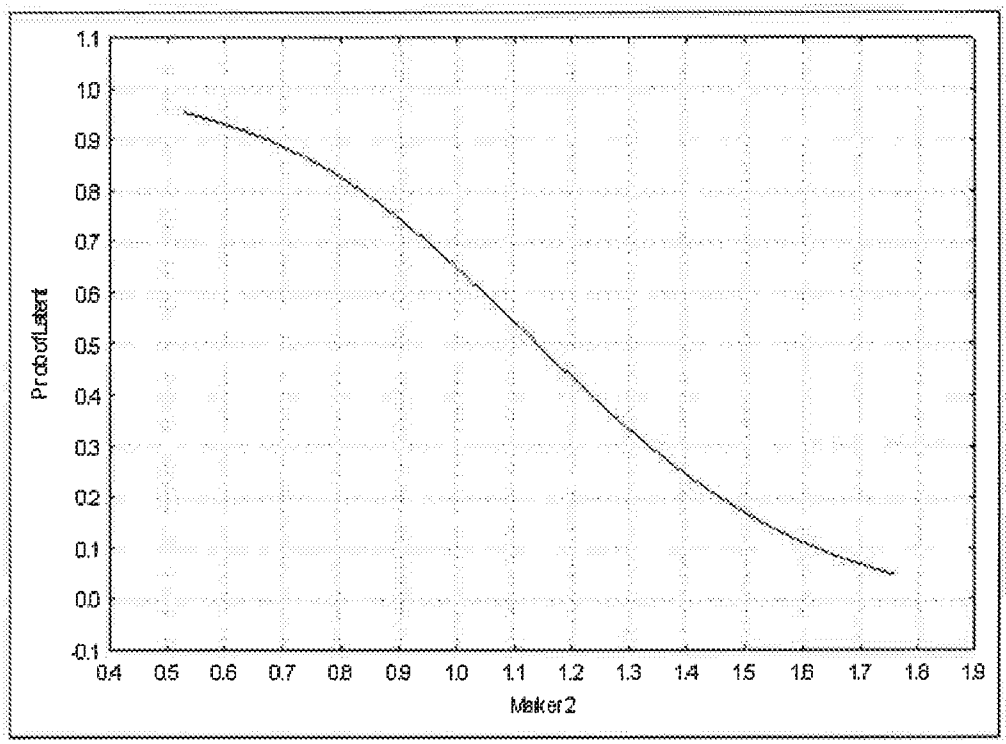

Figure 7F

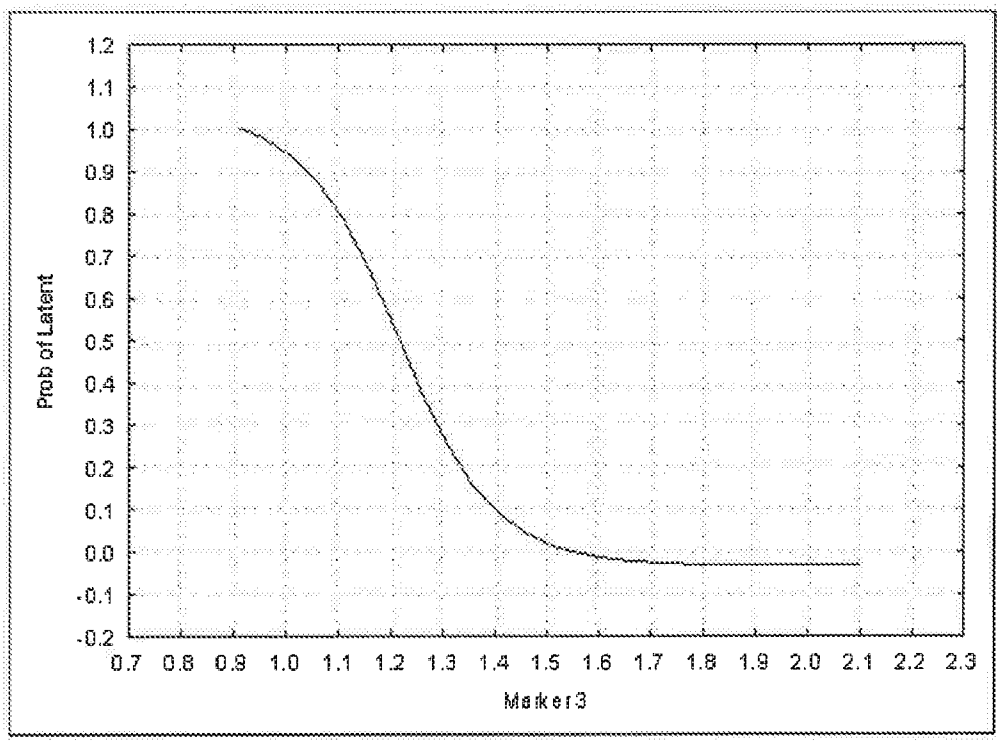

Figure 7G

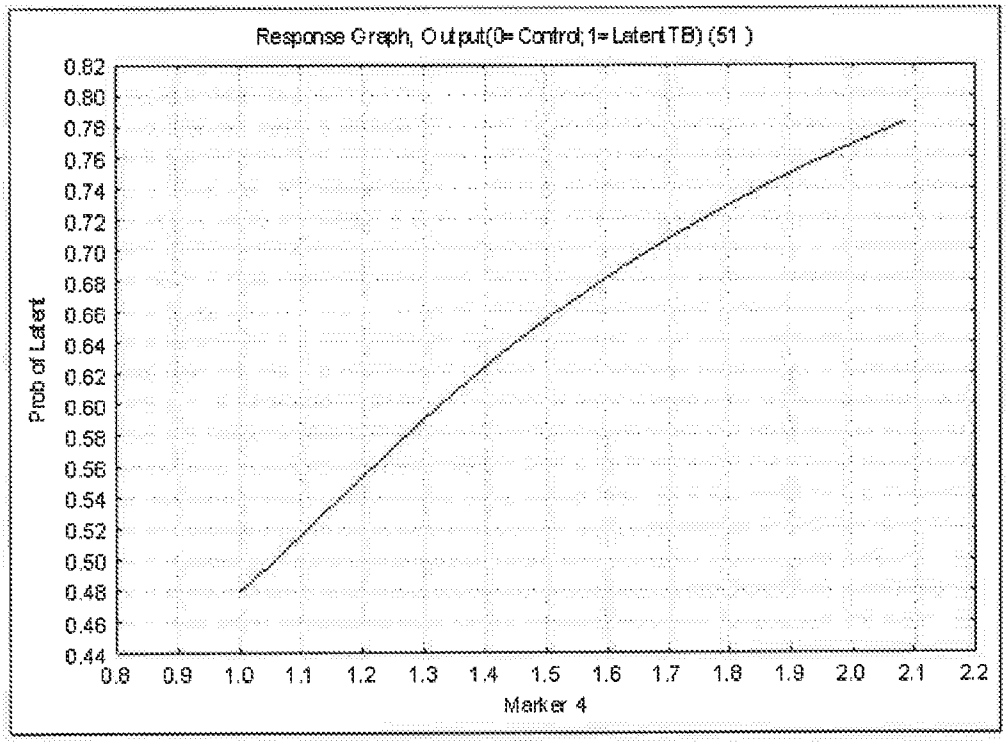
Response Graph, Output(0=Control;1=LatentTB) (51)
Prob of Latent
0.82
0.80
0.78
0.76
0.74
0.72
0.70
0.68
0.66
0.64
0.62
0.60
0.58
0.56
0.54
0.52
0.50
0.48
0.46
0.44
0.8
0.9
1.0
1.1
1.2
1.3
1.4
1.5
1.6
1.7
1.8
1.9
2.0
2.1
2.2
Marker 4

Figure 7H

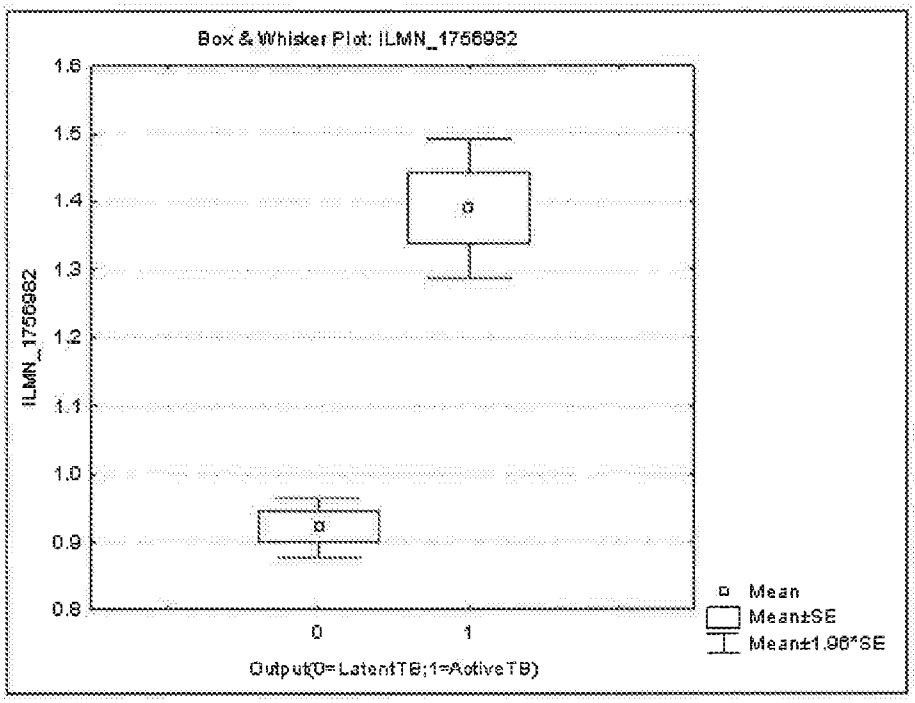

Figure 8A

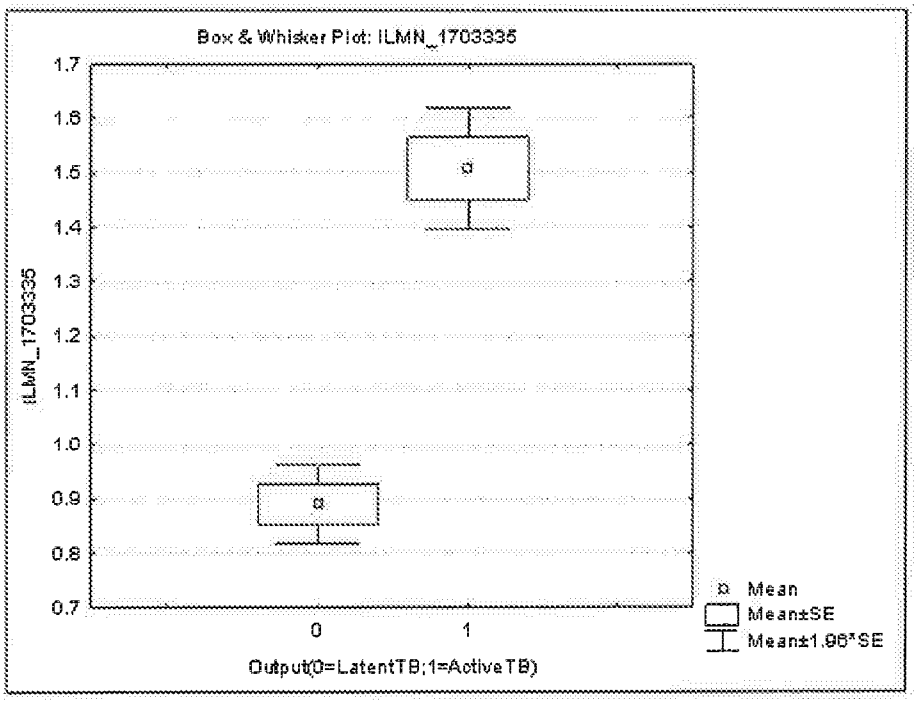

Figure 8B

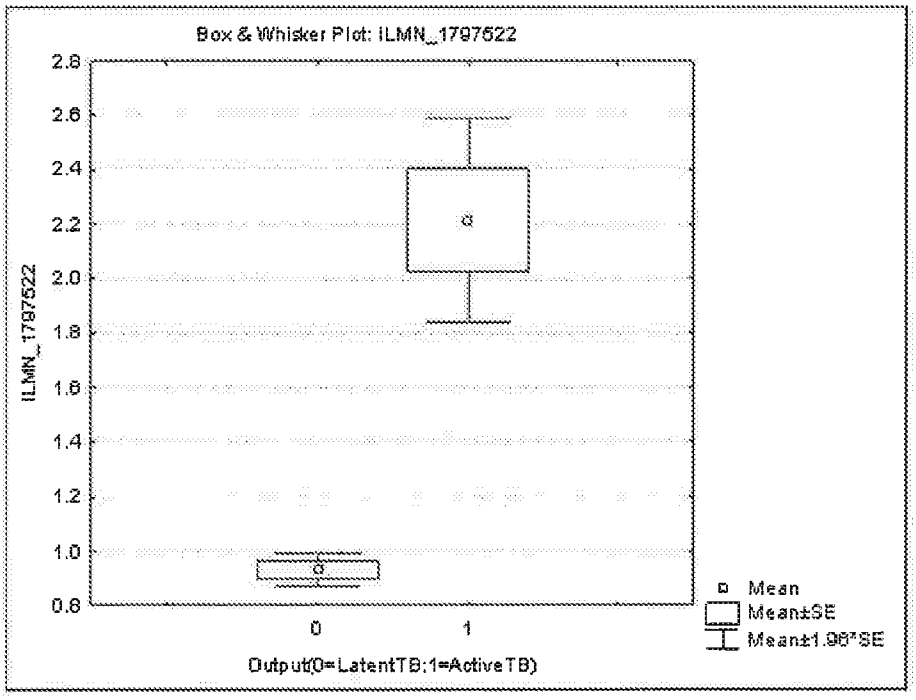
Box & Whisker Plot: ILMN_1797522
ILMN_1797522
2.8
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0
1
Output(0=LatentTB;1=ActiveTB)
Mean
Mean±SE
Mean±1.96*SE

Figure 8C

	Var1	Var5.51
1	0.598071	0.314299
2	0.644157	0.331317
3	0.690243	0.350748
4	0.736328	0.372930
5	0.782414	0.398122
6	0.828500	0.426400
7	0.874586	0.457555
8	0.920672	0.491017
9	0.966758	0.525812
10	1.012843	0.560586
11	1.058929	0.593820
12	1.105015	0.624180
13	1.151101	0.650807
14	1.197187	0.673374
15	1.243272	0.691957
16	1.289358	0.706861
17	1.335444	0.718469
18	1.381530	0.727151
19	1.427616	0.733211
20	1.473701	0.736892
21	1.519787	0.738407
22	1.565873	0.738007
23	1.611959	0.736044
24	1.658045	0.732999
25	1.704130	0.729424

	Var 2	Var 5.51
1	0.576960	0.271947
2	0.648845	0.306070
3	0.720731	0.343585
4	0.792617	0.384713
5	0.864502	0.429413
6	0.936388	0.477120
7	1.008273	0.526495
8	1.080159	0.575446
9	1.152044	0.621650
10	1.223930	0.663335
11	1.295816	0.699723
12	1.367701	0.730894
13	1.439587	0.757395
14	1.511472	0.779906
15	1.583358	0.799069
16	1.655243	0.815427
17	1.727129	0.829425
18	1.799015	0.841423
19	1.870900	0.851715
20	1.942786	0.860544
21	2.014671	0.868115
22	2.086557	0.874602
23	2.158442	0.880151
24	2.230328	0.884889
25	2.302214	0.888925

	Var 3	Var 5.51
1	0.613502	0.122972
2	0.731129	0.162712
3	0.848756	0.213758
4	0.966383	0.279598
5	1.084010	0.362529
6	1.201637	0.460264
7	1.319265	0.562780
8	1.436892	0.656315
9	1.554519	0.732486
10	1.672146	0.790756
11	1.789773	0.834559
12	1.907401	0.867858
13	2.025028	0.893757
14	2.142655	0.914380
15	2.260282	0.931139
16	2.377909	0.944984
17	2.495536	0.956574
18	2.613164	0.966376
19	2.730791	0.974736
20	2.848418	0.981914
21	2.966045	0.988109
22	3.083672	0.993479
23	3.201299	0.998151
24	3.318927	1.002226
25	3.436554	1.005788

Figure 9A

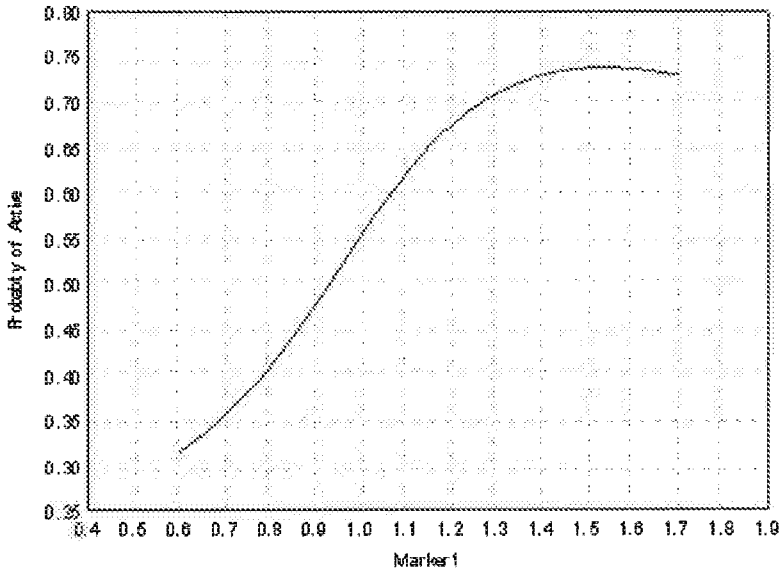
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
Probability of Active
0.4 0.5 0.6 0.7 0.8 0.9 1.0 1.1 1.2 1.3 1.4 1.5 1.6 1.7 1.8 1.9
Marker1

Figure 9B

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2013/051635

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12Q1/68
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)
C12Q

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 practicable, search terms used)
EPO-Internal, BIOSIS, Sequence Search, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2004/001070 A1 (GLAXO GROUP LTD [GB]; DUNCAN KENNETH [GB]; LUKEY PAULINE TERESA [GB];) 31 December 2003 (2003-12-31) the whole document -----	1-24
X	JOEL N. H. STERN ET AL: "Molecular signatures distinguishing active from latent tuberculosis in peripheral blood mononuclear cells, after in vitro antigenic stimulation with purified protein derivative of tuberculin (PPD) or Candida: a preliminary report", IMMUNOLOGIC RESEARCH, vol. 45, no. 1, 22 July 2008 (2008-07-22), pages 1-12, XP055029887, ISSN: 0257-277X, DOI: 10.1007/s12026-008-8024-2 see whole doc. esp. Table 1, ----- -/--	1-24

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"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 the actual completion of the international search	Date of mailing of the international search report
13 September 2013	04/12/2013

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Mueller, Frank

1

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/GB2013/051635

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-24(partially)

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-24(partially)

method of determining the M. tuberculosis infection status by determining the expression level of NXNL1 or combinations therewith

---

2-7. claims: 1-24(partially)

method of determining the M. tuberculosis infection status by determining the expression level of a further gene or gene combination therewith, for each gene or combination individually

---

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2013/051635

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2009/158521 A2 (BAYLOR RES INST [US]; NAT INST FOR MEDICAL RES [GB]; IMP COLLEGE HEALT) 30 December 2009 (2009-12-30) the whole document	1-24
X	WO 2011/066008 A2 (BAYLOR RES INST [US]; BANCHEREAU JACQUES F [US]; CHAUSSABEL DAMIEN [US) 3 June 2011 (2011-06-03) the whole document	1-24
X	WO 2006/125973 A2 (ST GEORGE S ENTPR LTD [GB]; FERNANDEZ-REYES DELMIRO [GB]; KRISHNA SANJ) 30 November 2006 (2006-11-30) the whole document	1-24

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/GB2013/051635

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2004001070 A1	31-12-2003	AU 2003241055 A1	06-01-2004
		WO 2004001070 A1	31-12-2003
WO 2009158521 A2	30-12-2009	AU 2009262112 A1	30-12-2009
		CA 2729000 A1	30-12-2009
		CN 102150043 A	10-08-2011
		EA 201170088 A1	31-10-2011
		EP 2300823 A2	30-03-2011
		JP 2011526152 A	06-10-2011
		KR 20110036590 A	07-04-2011
		NZ 590341 A	27-07-2012
		PE 03862011 A1	03-07-2011
		SG 182951 A1	30-08-2012
		TW 201022492 A	16-06-2010
		US 2011196614 A1	11-08-2011
		WO 2009158521 A2	30-12-2009
WO 2011066008 A2	03-06-2011	AR 080570 A1	18-04-2012
		AU 2010325179 A1	05-07-2012
		CA 2782211 A1	03-06-2011
		CN 102844444 A	26-12-2012
		EA 201270650 A1	28-06-2013
		EP 2519652 A2	07-11-2012
		JP 2013511981 A	11-04-2013
		KR 20120107979 A	04-10-2012
		PE 16902012 A1	16-12-2012
		US 2011129817 A1	02-06-2011
		WO 2011066008 A2	03-06-2011
WO 2006125973 A2	30-11-2006	EP 1896848 A2	12-03-2008
		JP 2008545960 A	18-12-2008
		US 2009104602 A1	23-04-2009
		WO 2006125973 A2	30-11-2006