(12) UK Patent Application (19) GB (11) 2460717 (13) A

(43) Date of A Publication 16.12.2009

(21) Application No: 0810709.6

(22) Date of Filing: 11.06.2008

(71) Applicant(s):
Sense Proteomic Limited
(Incorporated in the United Kingdom)
Unit 4, The Switchback, Gardner Road, Maidenhead,
Berkshire, SL6 7RJ, United Kingdom

(72) Inventor(s):
Martin Johannes Ebner
Colin Henry Wheeler
Rachel Alison Fallon
Sarah Paula Joyce
Jens-Oliver Koopman
Michael Bernard McAndrew
Nicholas Ian Workman

(74) Agent and/or Address for Service:
Carpmaels & Ransford
43-45 Bloomsbury Square, LONDON, WC1A 2RA,
United Kingdom

(51) INT CL:
G01N 33/564 (2006.01)   *A61K 39/00* (2006.01)
A61P 37/00 (2006.01)

(56) Documents Cited:
WO 2009/068481 A
Nature Genet; Vol 40, pp 211-216 (2008). Kozyrev et al.
"Functional variants in the B-cell gene BANK1 are
associated with systemic lupus erythematosus"
Arth Rheumat; Vol 58, pp S221- S222 (2008). Soejima
et al. "Association of SLE with STAT4, BANK1, BLK
and PXK: a replication study in a Japanese
population"
Arth Rheumat; Vol 58, pp S220 (2008). Guthridge et al.
"Confirmation of the genetic association of BANK1
with systemic lupus erythematosus"
EMBO J; Vol 21, pp 83-92 (2002). Yokoyama et al.
"BANK regulates BCR-induced calcium mobilization
by promoting tyrosine phosphorylation of IP3
receptor"
Genes Immunity; Vol 10, pp 373-379 (2009). Moser et
al. "Recent insights into the genetic basis of systemic
lupus erythematosus"

(58) Field of Search:
INT CL **A61K, G01N**
Other: **WPI, EPODOC, MEDLINE, BIOSIS**

(54) Abstract Title: **Autoantibodies for the detection of a predisposition to Lupus**

(57) Methods for diagnosing a patient as having or being predisposed to developing Lupus are described. The methods
comprise detecting in a sample taken from a patient one or more auto- antibodies that bind to auto-antigens,
wherein the presence of one or more auto-antibodies, or an increase in the concentration of one or more antibodies,
indicates that the patient suffers from or is predisposed to develop Lupus. A total of 31 separate autoantigens are
claimed. Methods of monitoring the progression of Lupus or the efficacy of a therapeutic agent to Lupus by
detecting the presence of these autoantigens are also claimed, along with diagonistic kits, and the uses of the
autoantigens in a vaccine for the treatment of Lupus.

GB 2460717 A

Figure 1. Cartoon illustrating the assay detecting auto-antibodies in human serum.

1. apply patient serum

2. incubate and wash

3. detect with -human antibody
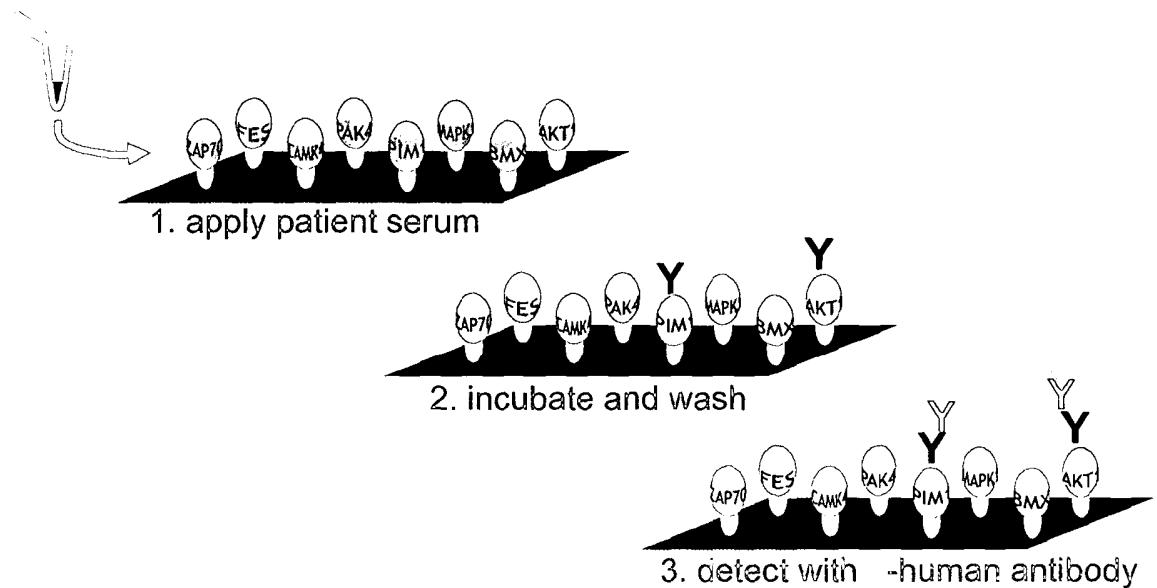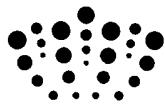
The following terms are registered trademarks and should be read as such wherever they occur in this document:

Triton

Biomarkers for Lupus

The present invention relates to novel biomarkers for lupus and methods and kits for their use.

Systemic lupus erythematosus (SLE) or lupus is a chronic autoimmune disease that can affect the joints and almost every major organ in the body, including heart, kidneys, skin, lungs, blood vessels, liver, and the nervous system. As in other autoimmune diseases, the body's immune system attacks the body's own tissues and organs, leading to inflammation. A person's risk to develop lupus appears to be determined mainly by genetic factors, but environmental factors, such as infection or stress may trigger the onset of the disease. The course of lupus varies, and is often characterised by alternating periods of flares, i.e. increased disease activity, and periods of remission.

It has been estimated that in the US, about 1 million people suffer from lupus, world-wide, conservative estimates indicate about 4 million lupus patients. Lupus occurs about 9 times more frequently in women than in men. Although lupus can occur at any age, it is most common in women of childbearing age.

There is presently no cure for lupus. However, there are means of reducing and dealing with flares, including drugs such as anti-inflammatory drugs, alternative medicines or life-style changes. The treatment regimes will depend on the severity of the disease, and the responsiveness of the patient. Disease-modifying antirheumatic drugs can be used preventively to reduce the incidence of flares. When flares occur, they are often treated with corticosteroids.

As mentioned above, lupus patients can present with a variety of diverse symptoms, and is therefore difficult to diagnose. A number of tests are used in order to make a diagnosis, including antinuclear antibody test (ANA), tests for other auto-antibodies such as anti-DNA antibodies, tests for serum complement levels, urine analysis, looking for elevated protein levels in urine, and sometimes biopsies of an affected organ are used. However, these tests do not necessarily give a definitive diagnosis; for example, a positive ANA test can occur due to infections or rheumatic diseases, and even healthy people without lupus can test positive. The sensitivity and specificity for the ANA test are 93%

and 57% (Habash-Bseiso,D Clin Med Res. 2005 August; 3(3): 190–193). Lupus
diagnosis often takes 7 to 10 years using current methods.

There is therefore a need for new or improved biomarkers for lupus that can be used in

5   methods of diagnosing lupus, for the early detection of lupus, subclinical or
presymptomatic lupus or a predisposition to lupus, or for monitoring the progression of
lupus or the likelihood to transition from remission to flare or vice versa, or the efficacy of
a therapeutic treatment thereof.

It has now been found that the presence of certain auto-antibodies or combinations of

10   auto-antibodies in a patient, or a significant increase in the number or concentration of
such auto-antibodies or combinations of auto-antibodies provide useful biomarkers for
lupus.   Auto-antibodies for which a decrease in concentration indicates that a patient
suffers from or is presymptomatic or predisposed to develop lupus are also envisaged, and
may also provide useful biomarkers for lupus.   Early results indicate that the disclosed

15   invention will also allow to distinguish lupus from closely related diseases such as
rheumatoid arthritis, and therefore will provide methods for the differential diagnosis of
lupus over related diseases.

According to one aspect of the present invention therefore there is provided a method
of diagnosing a patient as having or being predisposed to developing lupus, said method

20   comprising detecting in a sample taken from said patient one or more auto-antibodies that
bind one or more auto-antigens; wherein the presence of said one or more auto-antibodies,
or an increase in the concentration of said one or more auto-antibodies, indicates that the
patient suffers or is predisposed to develop lupus, wherein said one or more auto-antigens
are selected from BANK1 (B-cell scaffold protein with ankyrin repeats 1); BPY2IP1

25   (MAP1S protein); CEBPG (CCAAT/enhancer binding protein (C/EBP), gamma); E1B-
AP5 (E1B-55kDa-associated protein 5); FUS (Fusion (involved in t(12;16) in malignant
liposarcoma)); HAGH (Hydroxyacylglutathione hydrolase); HMG20B (High-mobility
group 20B); HOXB6 (Homeo box B6, transcript variant 2); KRT8 (Keratin 8); LIN28
(Lin-28 homolog (C. elegans));   NDUFV3   (NADH dehydrogenase (ubiquinone)

30   flavoprotein 3, 10kDa); PABPC1 (Poly(A) binding protein, cytoplasmic 1); PHLDA1

(Pleckstrin homology-like domain, family A, member 1); PIAS2 (Msx-interacting-zinc finger, transcript variant alpha); RAB11FIP3 (RAB11 family interacting protein 3 (class II)); RAN (RAN, member RAS oncogene family); RARA (Retinoic acid receptor alpha); RBMS1 (RNA binding motif, single stranded interacting protein 1, transcript variant); RDBP (RD RNA binding protein); RNF12 (Ring finger protein 12, transcript variant 1); RUFY1 (RUN and FYVE domain containing 1); SMN1 (Survival of motor neuron 1, telomeric); SRPK1 (SFRS protein kinase 1); SSNA1 (Sjogren's syndrome nuclear autoantigen 1); STAU (Staufen, RNA binding protein (Drosophila), transcript variant T3); STK11 (Serine/threonine kinase 11 (Peutz-Jeghers syndrome)); TOM1 (Target of myb1 (chicken)); TXNL2 (Thioredoxin-like, clone MGC:12349); TXNRD1 (Thioredoxin reductase 1, transcript variant 5); ZMAT2 (Zinc finger, matrin type 2); or ZNF38 (Zinc finger protein 38).

Preferably, the method comprises detecting auto-antibodies to one or more of the auto-antigens listed above, in combination with one or more of the auto-antigens IFI16 (Interferon, gamma-inducible protein 16); PSME3 (Proteasome (prosome, macropain) activator subunit 3 (PA28 gamma; Ki), transc), SSA2 (Sjogren syndrome antigen A2 (60kDa, ribonucleoprotein autoantigen SS-A/Ro)), and RALBP1 (RalA-binding protein 1).

The sample may be a body fluid, and is preferably selected from blood, serum, plasma saliva, lymphatic fluid, wound secretion, urine, faeces, mucus and cerebrospinal fluid (CSF). Typically, the sample may be serum or plasma.

By "bind" or "bound" used herein is meant a specific interaction strong enough to withstand standard washing procedures in typical protein binding assays.

Accordingly, the present invention comprehends the use of one or more auto-antibodies that bind to said one or more auto-antigens as biomarkers for lupus.

By an increase in the concentration of said one or more auto-antibodies is meant a statistically significant increase in the concentration of the one or more auto-antibodies in said sample taken from the patient as compared with a predetermined mean concentration found in non-diseased, normal patients. A surrogate normal sample can also be used, for

example a pooled normal control measured in relation to normal and diseased samples. Typically, the mean concentration of auto-antibodies found in normal patients may be calculated by reference to samples taken from a population of non-diseased individuals or from pooled samples from such individuals. In some embodiments, an increase of more than 10%, and suitably more than 15% or 20%, as compared with such mean may be taken as statistically significant. It will be appreciated that the units are arbitrary. In the case of clinical drug evaluation, there will be a change in auto-antibody concentration upon drug treatment; it is anticipated that the levels of most of these auto-antibodies may decrease post drug treatment.

Suitably, said method may comprise the detection of auto-antibodies that bind two or more of said auto-antigens, preferably 3 or more of said auto-antigens, even more preferably 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or 31 of these auto-antigens. In some embodiments, the patient may be pre-symptomatic.

In another aspect of the present invention there is provided a method of diagnosing a patient as having or being predisposed to developing lupus, said method comprising detecting in a sample from said patient auto-antibodies to a panel comprising 5 or more, preferably 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 or more, even more preferably to a panel comprising all 35, auto-antigens selected from the group consisting of those auto-antigens listed in Table 1 below; wherein the presence or separately a significant increase in the concentration of said auto-antibodies in the sample to substantially all members of the panel indicates that the patient suffers or is predisposed to develop lupus. In some embodiments, said panel may comprise additional auto-antigens (or other antigens) to those specified.

A significant increase in the concentration of the auto-antibodies means a statistically significant increase as determined by methods for statistically analysing experimental results of the kind commonly used in the art. By "substantially all members of the panel" is meant that the sample contains auto-antibodies, or an elevated concentration of auto-antibodies, to at least 50%, preferably at least 75% or 80% and typically at least 90% or 95%, of the auto-antigens included in the panel. It is envisaged that in some embodiments, the presence of auto-antibodies to one or more members of the panel and an increase in the

respective concentrations of auto-antibodies to one or more other members of the panel will be indicative of lupus, it being unnecessary in most cases for there to be an increased concentration for all auto-antibodies detected. However, in some cases, the mere presence of auto-antibodies to substantially all members of the panel will be significant.

5   It is expected that such a panel of biomarkers according to the present invention will provide greatly improved sensitivity or improved specificity in such methods of diagnosis as compared with known biomarkers for lupus.

As mentioned above, the auto-antigens for lupus are listed in Table 1 below. References to gene sequences and uniprot accession numbers for such auto-antigens are also given. However, it will be appreciated that variants of the sequences given, for example allelic variants, orthologues from other mammalian species or other variants that would be bound by the respective auto-antibodies may also be used. In particular, the present invention comprehends the use of auto-antigens being encoded from respective gene sequences that individually have at least an 80% identity to the corresponding sequences listed in Table 1. Suitably, such sequences have at least an 85% or 90% identity to the corresponding sequences in Table 1, and preferably they have at least a 95% identity to such sequences, e.g., more than a 96%, 97%, 98% or 99% identity.

Table 1

Gene symbol	Name	Synonyms	Entrez Gene ID and protein sequence link
BANK1	B-cell scaffold protein with ankyrin repeats 1	BANK; FLJ20706; FLJ34204	55024; http://ca.expasy.org/uniprot/Q8NDB2
BPY2IP1	MAP1S protein	MAP8; C19orf5; VCY2IP1; FLJ10669; VCY2IP-1; MGC133087	55201; http://www.ncbi.nlm.nih.gov/ entrez/viewer.fcgi?db=protein&val=AAH063
CEBPG	CCAAT/enhancer binding protein (C/EBP), gamma		1054; http://ca.expasy.org/uniprot/P53567
E1B-AP5	E1B-55kDa-associated protein 5	E1BAP5; E1B-AP5; HNRPUL1; FLJ12944	11100; http://ca.expasy.org/uniprot/O76022
FUS	Fusion (involved in t(12;16) in malignant liposarcoma)	TLS; CHOP; FUS1; FUS-CHOP; TLS/CHOP; hnRNP-P2	2521; http://ca.expasy.org/uniprot/P35637

HAGH	Hydroxyacylglutathione hydrolase	GLO2; GLX2; GLXII; HAGH1	3029; http://ca.expasy.org/uniprot/Q16775
HMG20B	High-mobility group 20B	SOXL; HMGX2; BRAF25; BRAF35; PP7706; pp8857; FLJ26127; SMARCE1r	10362; http://ca.expasy.org/uniprot/Q9P0W2
HOXB6	Homeo box B6, transcript variant 2	HOX2; HU-2; HOX2B; Hox-2.2	3216; http://ca.expasy.org/uniprot/P17509
IFI16	Interferon, gamma-inducible protein 16	PYHIN2; IFNGIP1; MGC9466	3428; http://ca.expasy.org/uniprot/Q16666
KRT8	Keratin 8	K8; KO; CK8; CYK8; K2C8; CARD2	3856; http://ca.expasy.org/uniprot/P05787
LIN28	Lin-28 homolog (C. elegans)	CSDD1; LIN-28; LIN28A; ZCCHC1; FLJ12457	79727; http://ca.expasy.org/uniprot/Q9H9Z2
NDUFV3	NADH dehydrogenase (ubiquinone) flavoprotein 3, 10kDa	CI-9KD	4731; http://ca.expasy.org/uniprot/P56181
PABPC1	Poly(A) binding protein, cytoplasmic 1	PAB1; PABP; PABP1; PABPC2; PABPL1	26986; http://ca.expasy.org/uniprot/P11940
PHLDA1	Pleckstrin homology-like domain, family A, member 1	PHRIP; TDAG51; DT1P1B11; MGC131738	22822; http://ca.expasy.org/uniprot/Q8WV24
PIAS2	Msx-interacting-zinc finger, transcript variant alpha	miz; MIZ1; SIZ2; PIASX; ZMIZ4; MGC102682; PIASX-BETA; PIASX-ALPHA	9063; http://ca.expasy.org/uniprot/Q13105
PSME3	Proteasome (prosome, macropain) activator subunit 3 (PA28 gamma; Ki), transc	Ki; PA28G; REG-GAMMA; PA28-gamma	10197; http://ca.expasy.org/uniprot/Q12920
RAB11FIP3	RAB11 family interacting protein 3 (class II)	KIAA0665; Rab11-FIP3	9727; http://ca.expasy.org/uniprot/O75154
RALBP1	RalA binding protein 1	RIP; RIP1; RLIP76	10928; http://ca.expasy.org/uniprot/Q15311
RAN	RAN, member RAS oncogene family,	TC4; Gsp1; ARA24	5901; http://ca.expasy.org/uniprot/P17080
RARA	Retinoic acid receptor, alpha	RAR; NR1B1	5914; http://ca.expasy.org/uniprot/P10276
RBMS1	RNA binding motif, single stranded interacting protein 1, transcript variant	YC1; MSSP; SCR2; MSSP-1; MSSP-2; MSSP-3;	5937; http://ca.expasy.org/uniprot/P29558

		MGC3331; MGC15146	
RDBP	RD RNA binding protein	RD; RDP; D6S45; NELF-E	7936; http://ca.expasy.org/uniprot/P18615
RNF12	Ring finger protein 12, transcript variant 1	RLIM; MGC15161; NY-REN-43	51132; http://ca.expasy.org/uniprot/Q9NVW
RUFY1	RUN and FYVE domain containing 1	RABIP4; ZFYVE12; FLJ22251	80230; http://ca.expasy.org/uniprot/Q96T51
SMN1	Survival of motor neuron 1, telomeric	SMA; SMN; SMA1; SMA2; SMA3; SMA4; SMA@; SMNT; BCD541; T-BCD541	6606; http://ca.expasy.org/uniprot/Q16637
SRPK1	SFRS protein kinase 1	SFRSK1	6732; http://ca.expasy.org/uniprot/Q96SB4
SSA2	Sjogren syndrome antigen A2 (60kDa, ribonucleoprotein autoantigen SS-A/Ro)	RO60; SSA2; TROVE2	6738; http://ca.expasy.org/uniprot/P10155
SSNA1	Sjogren's syndrome nuclear autoantigen 1	N14; NA14; NA-14	8636; http://ca.expasy.org/uniprot/O43805
STAU	Staufen, RNA binding protein (Drosophila), transcript variant T3	STAU; FLJ25010	6780; http://ca.expasy.org/uniprot/O95793
STK11	Serine/threonine kinase 11 (Peutz-Jeghers syndrome)	PJS; LKB1	6794; http://ca.expasy.org/uniprot/Q15831
TOM1	Target of myb1 (chicken)	FLJ33404	10043; http://ca.expasy.org/uniprot/O60784
TXNL2	Thioredoxin-like, clone MGC:12349	GRX3; GRX4; GLRX4; PICOT; TXNL2; TXNL3; FLJ11864; bA500G10.4; GLRX3	10539; http://ca.expasy.org/uniprot/O76003
TXNRD1	Thioredoxin reductase 1, transcript variant 5	TR; TR1; TXNR; TRXR1; GRIM-12; MGC9145	7296; http://ca.expasy.org/uniprot/Q16881
ZMAT2	Zinc finger, matrin type 2	FLJ31121	153527; http://ca.expasy.org/uniprot/Q96NC
ZNF38	Zinc finger protein 38	ZNF38; Zipro1; NY-REN-21; DKFZp434L134; DKFZp686H10254	7589; http://ca.expasy.org/uniprot/Q9Y5A6

The term "sequence identity" refers to the degree to which two polynucleotide or polypeptide sequences are identical on a residue-by-residue basis over a particular region of comparison. The term "percentage of sequence identity" is calculated by comparing

two optimally aligned sequences over that region of comparison, determining the number of positions at which the identical nucleotide base or amino acid occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the region of comparison (*i.e.*, the window size), and

5   multiplying the result by 100 to yield the percentage of sequence identity. The identity between two DNA or protein sequences may be determined using computer programs known in the art, such as GAP software provided in the GCG program package with default parameters (Needleman and Wunsch 1970 *J Mol Biol* 48: 443-453). The identity between two protein sequences is also displayed after searching a protein database with

10   Blast2 (Altschul et al (1997) Nucleic Acids Res. 25 (17) 3389-3402).

It is anticipated that the methods of the present invention may provide greater sensitivity or specificity for the diagnosis of lupus than previously available in the art.

By "sensitivity" is meant the proportion of persons with the disease phenotype who test positive.

15   By "specificity" is meant the proportion of persons without the disease phenotype who test negative.

Advantageously, the methods of the invention may have a sensitivity of >67%, preferably >70%, even more preferably >75%, 80%, 85%, 90%, and most preferably >95%, or a specificity of >85%, preferably >90%, even more preferably >95%, and most

20   preferably > 98%.

According to yet another aspect of the invention there is provided a method of monitoring the progression or regression of lupus in a patient, said method comprising detecting in a first sample obtained from a patient one or more auto-antibodies that bind one or more auto-antigens selected from the group consisting of BANK1 (B-cell scaffold

25   protein with ankyrin repeats 1); BPY2IP1 (MAP1S protein); CEBPG (CCAAT/enhancer binding protein (C/EBP), gamma); E1B-AP5 (E1B-55kDa-associated protein 5); FUS (Fusion (involved in t(12;16) in malignant liposarcoma)); HAGH (Hydroxyacylglutathione hydrolase); HMG20B (High-mobility group 20B); HOXB6 (Homeo box B6, transcript variant 2); KRT8 (Keratin 8); LIN28 (Lin-28 homolog (C.

elegans)); NDUFV3 (NADH dehydrogenase (ubiquinone) flavoprotein 3, 10kDa); PABPC1 (Poly(A) binding protein, cytoplasmic 1); PHLDA1 (Pleckstrin homology-like domain, family A, member 1); PIAS2 (Msx-interacting-zinc finger, transcript variant alpha); RAB11FIP3 (RAB11 family interacting protein 3 (class II)); RAN (RAN, member RAS oncogene family); RARA (Retinoic acid receptor alpha); RBMS1 (RNA binding motif, single stranded interacting protein 1, transcript variant); RDBP (RD RNA binding protein); RNF12 (Ring finger protein 12, transcript variant 1); RUFY1 (RUN and FYVE domain containing 1); SMN1 (Survival of motor neuron 1, telomeric); SRPK1 (SFRS protein kinase 1); SSNA1 (Sjogren's syndrome nuclear autoantigen 1); STAU (Staufen, RNA binding protein (Drosophila), transcript variant T3); STK11 (Serine/threonine kinase 11 (Peutz-Jeghers syndrome)); TOM1 (Target of myb1 (chicken)); TXNL2 (Thioredoxin-like, clone MGC:12349); TXNRD1 (Thioredoxin reductase 1, transcript variant 5); ZMAT2 (Zinc finger, matrin type 2); or ZNF38 (Zinc finger protein 38) , or auto-antibodies that bind to a panel comprising 5 or more, preferably 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 or more, even more preferably to a panel comprising all 35, auto-antigens selected from the group consisting of those auto-antigens listed in Table 1; detecting said auto-antibodies in a second sample obtained from the patient at a time later than the first sample; wherein a change in the auto-antibodies detected in the first and second samples is indicative of the progression or regression of lupus.

According to yet another aspect of the invention there is provided a method of monitoring the progression towards a flare of the disease, or the transition from a flare to remission, of lupus in a patient, said method comprising detecting in a first sample obtained from a patient one or more auto-antibodies that bind one or more auto-antigens selected from the group consisting of BANK1 (B-cell scaffold protein with ankyrin repeats 1); BPY2IP1 (MAP1S protein); CEBPG (CCAAT/enhancer binding protein (C/EBP), gamma); E1B-AP5 (E1B-55kDa-associated protein 5); FUS (Fusion (involved in t(12;16) in malignant liposarcoma)); HAGH (Hydroxyacylglutathione hydrolase); HMG20B (High-mobility group 20B); HOXB6 (Homeo box B6, transcript variant 2); KRT8 (Keratin 8); LIN28 (Lin-28 homolog (C. elegans)); NDUFV3 (NADH dehydrogenase (ubiquinone) flavoprotein 3, 10kDa); PABPC1 (Poly(A) binding protein, cytoplasmic 1); PHLDA1 (Pleckstrin homology-like domain, family A, member 1); PIAS2 (Msx-

interacting-zinc finger, transcript variant alpha); RAB11FIP3 (RAB11 family interacting protein 3 (class II)); RAN (RAN, member RAS oncogene family); RARA (Retinoic acid receptor alpha); RBMS1 (RNA binding motif, single stranded interacting protein 1, transcript variant); RDBP (RD RNA binding protein); RNF12 (Ring finger protein 12, transcript variant 1); RUFY1 (RUN and FYVE domain containing 1); SMN1 (Survival of motor neuron 1, telomeric); SRPK1 (SFRS protein kinase 1); SSNA1 (Sjogren's syndrome nuclear autoantigen 1); STAU (Staufen, RNA binding protein (Drosophila), transcript variant T3); STK11_(Serine/threonine kinase 11 (Peutz-Jeghers syndrome)); TOM1 (Target of myb1 (chicken)); TXNL2 (Thioredoxin-like, clone MGC:12349); TXNRD1 (Thioredoxin reductase 1, transcript variant 5); ZMAT2 (Zinc finger, matrin type 2); or ZNF38 (Zinc finger protein 38), or auto-antibodies that bind to a panel comprising 5 or more, preferably 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 or more, even more preferably to a panel comprising all 35, auto-antigens selected from the group consisting of those auto-antigens listed in Table 1; detecting said auto-antibodies in a second sample obtained from the patient at a time later than the first sample; wherein a change in the auto-antibodies detected in the first and second samples is indicative of the progression towards a flare, or transition from a flare to a period of remission of the disease.

Said change in the auto-antibodies may be a change in the number of different auto-antigens to which auto-antibodies are detected in the first as compared to the second sample, or it may be a change in the concentration of an auto-antibody in the first as compared to the second sample, or it may be a combination of both.

According to yet another aspect of the present invention there is provided a method of monitoring the efficacy of a therapeutic agent to lupus, comprising detecting in a first sample obtained from a patient the presence of one or more auto-antibodies that bind one or more auto-antigens selected from the group consisting of BANK1 (B-cell scaffold protein with ankyrin repeats 1); BPY2IP1 (MAP1S protein); CEBPG (CCAAT/enhancer binding protein (C/EBP), gamma); E1B-AP5 (E1B-55kDa-associated protein 5); FUS (Fusion (involved in t(12;16) in malignant liposarcoma)); HAGH (Hydroxyacylglutathione hydrolase); HMG20B (High-mobility group 20B); HOXB6 (Homeo box B6, transcript variant 2); KRT8 (Keratin 8); LIN28 (Lin-28 homolog (C.

elegans)); NDUFV3 (NADH dehydrogenase (ubiquinone) flavoprotein 3, 10kDa); PABPC1 (Poly(A) binding protein, cytoplasmic 1); PHLDA1 (Pleckstrin homology-like domain, family A, member 1); PIAS2 (Msx-interacting-zinc finger, transcript variant alpha); RAB11FIP3 (RAB11 family interacting protein 3 (class II)); RAN (RAN, member RAS oncogene family); RARA (Retinoic acid receptor alpha); RBMS1 (RNA binding motif, single stranded interacting protein 1, transcript variant); RDBP (RD RNA binding protein); RNF12 (Ring finger protein 12, transcript variant 1); RUFY1 (RUN and FYVE domain containing 1); SMN1 (Survival of motor neuron 1, telomeric); SRPK1 (SFRS protein kinase 1); SSNA1 (Sjogren's syndrome nuclear autoantigen 1); STAU (Staufen, RNA binding protein (Drosophila), transcript variant T3); STK11 (Serine/threonine kinase 11 (Peutz-Jeghers syndrome)); TOM1 (Target of myb1 (chicken)); TXNL2 (Thioredoxin-like, clone MGC:12349); TXNRD1 (Thioredoxin reductase 1, transcript variant 5); ZMAT2 (Zinc finger, matrin type 2); or ZNF38 (Zinc finger protein 38), or auto-antibodies that bind to a panel comprising 5 or more, preferably 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 or more, even more preferably to a panel comprising all 35, auto-antigens selected from the group consisting of those auto-antigens listed in Table 1; administering the therapeutic agent to the patient; detecting said one or more auto-antibodies in a second sample taken from the patient taken after the administration of the therapeutic agent; and comparing the one or more auto-antibodies in said first and second samples respectively, wherein a change in the auto-antibodies in the first and second sample is indicative of the efficacy of the therapeutic agent to lupus.

The first sample may be taken prior to the first administration of a therapeutic agent, or after treatment has already commenced. The method can be used to monitor the efficacy of treatment over time, e.g., over days, weeks, months or years.

The therapeutic agent may be any therapeutic agent that can be beneficial for the treatment of lupus, for example anti-rheumatic agents or anti-inflammatory agents.

According to yet another aspect of the present invention, there is provided a method of predicting whether a patient suffering from lupus will respond to treatment with a therapeutic agent, comprising detecting the presence or absence of one or more auto-antibodies that bind to one or more auto-antigens selected from the group consisting of

BANK1 (B-cell scaffold protein with ankyrin repeats 1); BPY2IP1 (MAP1S protein); CEBPG (CCAAT/enhancer binding protein (C/EBP), gamma); E1B-AP5 (E1B-55kDa-associated protein 5); FUS (Fusion (involved in t(12;16) in malignant liposarcoma)); HAGH (Hydroxyacylglutathione hydrolase); HMG20B (High-mobility group 20B); HOXB6 (Homeo box B6, transcript variant 2); KRT8 (Keratin 8); LIN28 (Lin-28 homolog (C. elegans)); NDUFV3 (NADH dehydrogenase (ubiquinone) flavoprotein 3, 10kDa); PABPC1 (Poly(A) binding protein, cytoplasmic 1); PHLDA1 (Pleckstrin homology-like domain, family A, member 1); PIAS2 (Msx-interacting-zinc finger, transcript variant alpha); RAB11FIP3 (RAB11 family interacting protein 3 (class II)); RAN (RAN, member RAS oncogene family); RARA (Retinoic acid receptor alpha); RBMS1 (RNA binding motif, single stranded interacting protein 1, transcript variant); RDBP (RD RNA binding protein); RNF12 (Ring finger protein 12, transcript variant 1); RUFY1 (RUN and FYVE domain containing 1); SMN1 (Survival of motor neuron 1, telomeric); SRPK1 (SFRS protein kinase 1); SSNA1 (Sjogren's syndrome nuclear autoantigen 1); STAU (Staufen, RNA binding protein (Drosophila), transcript variant T3); STK11_(Serine/threonine kinase 11 (Peutz-Jeghers syndrome)); TOM1 (Target of myb1 (chicken)); TXNL2 (Thioredoxin-like, clone MGC:12349); TXNRD1 (Thioredoxin reductase 1, transcript variant 5); ZMAT2 (Zinc finger, matrin type 2); or ZNF38 (Zinc finger protein 38), or auto-antibodies that bind to a panel comprising 5 or more, preferably 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 or more, even more preferably to a panel comprising all 35, auto-antigens selected from the group consisting of those auto-antigens listed in Table 1; wherein the presence or absence of said one or more auto-antibodies indicates that the patient will or will not respond to treatment with said therapeutic agent.

The auto-antibodies may be detected by any means known in the art. This includes the use of protein arrays, bead-based assays, Western blots, immunoprecipitation, silver staining, dot blots, etc. The binding of the auto-antibodies can be detected by any means, including enzyme-linked assays such as ELISAs, radioimmunoassays (RIAs), DELFIA assays, fluorescence-based detection methods, Luminex bead assays, etc.

In some embodiments, arrays of said auto-antigens may be employed for detecting said auto-antibodies, for example arrays prepared according to the methods disclosed by WO-

A-01/57198 or WO-A-2003/064656; preferred detection methods are fluorescence-based detection methods.

The patient may have or be predisposed to develop the disease naturally. However, a patient may also develop lupus as a result of environmental factors, for example as a side effect of treatment with therapeutic agents or as a result of exposure to toxic materials, or as a result of manipulation, e.g., in animal models used in pharmaceutical research and development. The patient may be a human being, but can also be a non-human animal such, for example, as a mouse, rat, guinea pig, cat, dog, horse, pig, cow, monkey, or any other animal used for research and drug testing.

According to a different aspect of the present invention there is provided a diagnostic kit for use in diagnosing a patient as having or being predisposed to developing lupus, or having presymptomatic lupus, said kit comprising one or more auto-antigens selected from the group consisting of BANK1 (B-cell scaffold protein with ankyrin repeats 1); BPY2IP1 (MAP1S protein); CEBPG (CCAAT/enhancer binding protein (C/EBP), gamma); E1B-AP5 (E1B-55kDa-associated protein 5); FUS (Fusion (involved in t(12;16) in malignant liposarcoma)); HAGH (Hydroxyacylglutathione hydrolase); HMG20B (High-mobility group 20B); HOXB6 (Homeo box B6, transcript variant 2); KRT8 (Keratin 8); LIN28 (Lin-28 homolog (C. elegans)); NDUFV3 (NADH dehydrogenase (ubiquinone) flavoprotein 3, 10kDa); PABPC1 (Poly(A) binding protein, cytoplasmic 1); PHLDA1 (Pleckstrin homology-like domain, family A, member 1); PIAS2 (Msx-interacting-zinc finger, transcript variant alpha); RAB11FIP3 (RAB11 family interacting protein 3 (class II)); RAN (RAN, member RAS oncogene family); RARA (Retinoic acid receptor alpha); RBMS1 (RNA binding motif, single stranded interacting protein 1, transcript variant); RDBP (RD RNA binding protein); RNF12 (Ring finger protein 12, transcript variant 1); RUFY1 (RUN and FYVE domain containing 1); SMN1 (Survival of motor neuron 1, telomeric); SRPK1 (SFRS protein kinase 1); SSNA1 (Sjogren's syndrome nuclear autoantigen 1); STAU (Staufen, RNA binding protein (Drosophila), transcript variant T3); STK11_(Serine/threonine kinase 11 (Peutz-Jeghers syndrome)); TOM1 (Target of myb1 (chicken)); TXNL2 (Thioredoxin-like, clone MGC:12349); TXNRD1 (Thioredoxin reductase 1, transcript variant 5); ZMAT2 (Zinc finger, matrin type 2); or ZNF38 (Zinc finger protein 38), which one or more auto-antigens are immobilised on a

substrate. In addition to one or more of the auto-antigens above, the kit may comprise one or more of the auto-antigens IFI16 (Interferon, gamma-inducible protein 16); PSME3 (Proteasome (prosome, macropain) activator subunit 3 (PA28 gamma; Ki), transc); RALBP1 (RalA-binding protein 1); and SSA2 (Sjogren syndrome antigen A2 (60kDa, ribonucleoprotein autoantigen SS-A/Ro)).

Alternatively, said kit may comprise a panel comprising 5 or more, preferably 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 or more, even more preferably to a panel comprising all 35, auto-antigens selected from the auto-antigens listed in Table 1, wherein the auto-antigens are immobilised on a substrate.

In some embodiments, said substrate may comprise a strip, a slide, beads or a well of a microtitre plate or a "chip" for use with Surface Plasmon Resonance (Biacore) or planar waveguide technology. Suitably, the substrate may comprise a disposable test strip of the kind widely used in diagnostic kits for use at the point of care.

The panel of auto-antigens may be carried on said substrate in the form of an addressable array. Said array may be a microarray. The auto-antigens of the panel may constitute at least 50%, typically at least 70% or 80%, of the total number of proteins present on the array. For example, the auto-antigens of the panel may constitute more than 85%, 90% or 95% of the total number of proteins on the array.

Alternatively, said diagnostic kit may comprise said panel of auto-antigens carried on a substrate in the form of an array, which array includes one or more other proteins or polypeptides, and means for automatically analysing the results of tests performed using said array, said analysing means including computer-readable instructions in which results from said one or more other proteins or polypeptides are blinded during the analysis.

In addition to the auto-antigens (and optional one or more other proteins or polypeptides) forming said panel, the array may further comprise one or more control proteins or polypeptides. Suitably, the array may comprise one or more replicates of the auto-antigens or controls, preferably replicates of each auto-antigen and each control.

Advantageously, the array comprises an array of correctly folded auto-antigens, which may be prepared in accordance with the methods disclosed by WO-A-01/57198 or WO-A-02/27327, for example, the contents of which are incorporated herein by reference. Preferably, the array comprises an array of correctly folded tagged auto-antigens that are attached to said substrate *via* the tag, which array may be prepared, for example, in accordance with the methods disclosed by WO 03/064656 or WO-A-2004/046730, the contents of which are also incorporated herein by reference.

Said kit may further comprise an anti-human immunoglobulin antibody or other protein that specifically binds immunoglobulin, and means for detecting the anti-human immunoglobulin antibody or other immunoglobulin-binding protein.

Said detection means may comprise a colorimetric or fluorescent assay or utilise label-free methods such as surface plasmon resonance or planar waveguide technology.

The anti-human immunoglobulin antibody may be an anti-human IgM antibody, an anti-human IgG antibody, an anti-human IgA antibody, an anti-human IgE antibody or a combination of any two or more of such anti-human antibodies; preferably, the anti-human immunoglobulin antibody is an anti-human IgG antibody. Alternatively, or additionally, a protein that specifically binds an antibody is included in the kit. Preferably, this protein is protein A or protein G, or a recombinant form thereof.

According to yet another aspect of the present invention there is provided the use of one or more auto-antigens selected from the aforementioned group as shown in Table 1, or antigenic fragments or antigenic epitopes thereof, or a nucleic acid encoding the auto-antigen or fragment or epitope thereof, as an active agent in the manufacture of a vaccine for the treatment or prevention of lupus.

Said active agent may be formulated in a pharmaceutical composition in accordance with method is well known in the art. Said composition may further comprise a pharmaceutically acceptable vehicle. Said vehicle may comprise one or more excipients, carriers, buffers, stabilisers or other materials well known to those skilled in the art. Such materials should be non-toxic and should not interfere with the efficacy of the active agent. The precise nature of the carrier or other material may depend on the route of

administration, *e.g.,* oral, intravenous, cutaneous or subcutaneous, nasal, intramuscular, intraperitoneal or patch routes.

Pharmaceutical compositions for oral administration may be in tablet, capsule, powder or liquid form. A tablet may include a solid carrier such as gelatin or an adjuvant. Liquid pharmaceutical compositions generally include a liquid carrier such as water, petroleum, animal or vegetable oils, mineral oil or synthetic oil. Physiological saline solution, dextrose or other saccharide solution or glycols such as ethylene glycol, propylene glycol or polyethylene glycol may be included.

For intravenous, cutaneous or subcutaneous injection, or injection at the site of affliction, the active ingredient will be in the form of a parenterally acceptable aqueous solution which is pyrogen-free and has suitable pH, isotonicity and stability. Those of relevant skill in the art are well able to prepare suitable solutions using, for example, isotonic vehicles such as Sodium Chloride Injection, Ringer's Injection, Lactated Ringer's Injection. Preservatives, stabilisers, buffers, antioxidants and/or other additives may be included, as required.

Whether it is a polypeptide, peptide, or nucleic acid molecule, other pharmaceutically useful compound according to the present invention that is to be given to an individual, administration is preferably in a "prophylactically effective amount" or a "therapeutically effective amount" (as the case may be, although prophylaxis may be considered therapy), this being sufficient to show benefit to the individual. The actual amount administered, and rate and time-course of administration, will depend on the nature and severity of what is being treated. Prescription of treatment, *e.g.* decisions on dosage etc, is within the responsibility of general practitioners and other medical doctors, and typically takes account of the disorder to be treated, the condition of the individual patient, the site of delivery, the method of administration and other factors known to practitioners. Examples of the techniques and protocols mentioned above can be found in REMINGTON'S PHARMACEUTICAL SCIENCES, 16th edition, Osol, A. (ed), 1980.

Alternatively, targeting therapies may be used to deliver the active agent more specifically to certain types of cell, by the use of targeting systems such as antibody or cell specific ligands. Targeting may be desirable for a variety of reasons; for example if the

agent is unacceptably toxic, or if it would otherwise require too high a dosage, or if it would not otherwise be able to enter the target cells.

Instead of administering these agents directly, they could be produced in the target cells by expression from an encoding gene introduced into the cells, *e.g.* in a viral vector (a

5   variant of the VDEPT technique - *see* below). The vector could be targeted to the specific cells to be treated, or it could contain regulatory elements, which are switched on more or less selectively by the target cells.

Alternatively, the agent could be administered in a precursor form, for conversion to the active form by an activating agent produced in, or targeted to, the cells to be treated.

10   This type of approach is sometimes known as ADEPT or VDEPT; the former involving targeting the activating agent to the cells by conjugation to a cell-specific antibody, while the latter involves producing the activating agent, *e.g.* a vaccine or fusion protein, in a vector by expression from encoding DNA in a viral vector (*see* for example, EP-A-415731 and WO 90/07936).

15   The vaccines of the present invention also include one or more adjuvant compounds, i.e. a compound or compounds that increases an immunogenic response or the immunogenicity of an antigen or vaccine. Adjuvant compounds useful in the present invention include Complete Freund's Adjuvant (CFA); Incomplete Freund's Adjuvant (IFA); Montanide ISA (incomplete seppic adjuvant); Ribi Adjuvant System (RAS);

20   TiterMax; Syntex Adjuvant Formulation (SAF); Aluminum Salt Adjuvants; Nitrocellulose-adsorbed antigen; Encapsulated or entrapped antigens; Immune-stimulating complexes (ISCOMs); and Gerbu^R adjuvant.

Following is a description by way of example only with reference to the accompanying drawings of embodiments of the present invention.

25   In the drawings, Figure 1 is a cartoon illustrating a method for the detection of auto-antibodies in serum.

Example 1: A microarray of antigens for discovering auto-antibody biomarkers in patient serum that are diagnostic for lupus

Full-length open reading frames for 330 target genes covering a wide range of protein classes were cloned in-frame with a sequence encoding a C-terminal *E. coli* BCCP-myc tag (WO 03/064656; Boutell) in a baculovirus transfer vector and sequence-verified. Recombinant baculoviruses were generated, amplified and expressed in Sf9 cells grown in suspension using standard methods adapted for 24-well deep well plates (Whatman, Maidstone, UK; Chambers, Zhao). Recombinant protein expression was analyzed for protein integrity and biotinylation by Western blotting. Cells harboring recombinant protein were lysed and lysates were spotted in quadruplicate using a QArray2 Microarrayer (Genetix, New Milton, UK) equipped with 300 μm solid pins on to streptavidin-coated glass slides (Schott Nexterion, Jena, Germany). Slides were stored at -20°C. A range of protein activities have been demonstrated to be stable under these conditions for over 12 months. Proteins spotted on the array project into an aqueous environment and orient away from the surface of the glass chips, exposing them for binding by auto-antibodies.

In addition to the 330 proteins, four control proteins for the BCCP-myc tag (BCCP, BCCP-myc, β-galactosidase-BCCP-myc and β-galactosidase-BCCP) were arrayed, along with Cy3/Cy5-labeled biotin-BSA, dilution series of biotinylated-IgG and biotinylated IgM, a biotinylated-myc peptide dilution series and buffer only spots.

Example 2: Detection of auto-antibodies in serum samples from patients suffering from lupus

Serum samples of 169 individuals diagnosed with lupus (median age 47 years, 95% females) were obtained from 2 commercial sources. As controls, samples of 170 individuals with no known history of lupus (median age 50 years, 95% females) were obtained. These samples were assayed for the presence of auto-antibodies using the protein arrays described in Example 1.

Serum samples were clarified (centrifuged at 10-13K rcf for 3 minutes at 4°C to remove particulates, including lipids) and diluted 200-fold in 0.1% v/v Triton/0.1% w/v

BSA in 1X PBS (Triton-BSA-PBS buffer) and then applied to the arrays. Each diluted serum sample (2.0 mL) was applied to a separate array and incubated in a dish for 2 hours at room temperature (RT, 20°C) with gentle orbital shaking (~ 50 rpm). Arrays were then carefully removed from the dish and any excess probing solution was removed by blotting

5 the sides of the array onto lint-free tissue. Probed arrays were washed twice in fresh Triton-BSA-PBS buffer in a new plastic container with screw-on lid at room temperature for 5 minutes with rapid orbital shaking. The washed slides were then blotted onto lint-free tissue to remove excess wash buffer and were incubated in a secondary antibody solution (prepared just prior to use) at room temperature for 2 hours, with gentle orbital

10 shaking and protected from light. The secondary antibody solution was a Cy3 labeled rabbit anti-human IgG antibody diluted in Triton-BSA-PBS buffer to give optimal detection of bound serum IgG. The slides were washed three times in Triton-BSA-PBS buffer for 5 minutes at RT with gentle orbital shaking, rinsed briefly (5-10 seconds) in distilled water, and centrifuged for 2 minutes at 240 rcf in a container suitable for

15 centrifugation. To help wick away excess liquid on the arrays, a lint-free tissue was placed at the bottom of the arrays during centrifugation.

The probed and dried arrays were then scanned using a microarray scanner capable of using an excitation wavelength suitable for the detection of the secondary staining solution, such as the Molecular Devices Genepix 4000B microarray scanner, to identify auto-

20 antibodies bound by the array and to determine intensity of auto-antibody binding. The microarray scans produced TIFF images for each array that were used to determine the intensity of fluorescence bound to each protein spot.

Raw median signal intensity (also referred to as the relative fluorescent unit, RFU) of each protein feature (also referred to as a spot or antigen) on the array was determined

25 from the TIFF images using GenePix Pro microarray data analysis software or similar software and subtracted from the local median background intensity. It is also possible to use other measures of spot intensity such as the mean fluorescence, total fluorescence, as is known by anyone skilled in the art.

The resulting net fluorescent intensities of all protein features on each array were then

30 normalized using the quantile normalization (Bolstad) method. Data for sera from two lupus patients and one control serum were excluded because many of their antigen binding

signals were masked by high background binding. Other methods for data normalization suitable for the current data include, amongst others, multiplication of net fluorescent intensities by a normalization factor consisting of the product of the 1st quartile of all intensities of a sample and the mean of the 1st quartiles of all samples and the vsn method (Huber). Such normalization methods are known to anyone skilled in the art of microarray analysis. The normalized fluorescent intensities were then averaged for each protein feature.

The normalized mean data were used to select biomarkers for lupus. The data were randomly split into training and test sets 10 times (10 fold cross-validation)

Proteins were selected and ranked with the significance analysis of microarrays (SAM) algorithm (Tusher) using the data in the training subset. Similar methods that are applicable to this task include, amongst others, Prediction Analysis of Microarrays (PAM, Tibshirani), MP test (Fox), fold change (Witten) and principal component analysis (Bessant). Such data analysis methods are known to anyone skilled in the art of microarray analysis. The data were analysed as two class –unpaired (diseased vs control) using a non-parametric test statistic. The protein features selected in each cross-validation fold with the highest scores and a q-value of ≤30% were selected and those with q-values ≤30% in most of the 10 cross-validation iterations were selected and are shown in Table 2 alongside the frequency of selection in cross-validation.

Table 2 lists the protein features for which the q-value was ≤30% by SAM analysis alongside the selection frequency in cross-validation

Gene symbol	cross-validation selection frequency
ZNF38	10
BPY2IP1	10
CEBPG	10
E1B-AP5	10
FUS	10
HMG20B	10
LIN28	10
NDUFV3	10

PABPC1	10
PHLDA1	10
PIAS2	10
PSME3	10
RAB11FIP3	10
RALBP1	10
RAN	10
RARA	10
RBMS1	10
RNF12	10
SSA2	10
STAU	10
STK11	10
TOM1	10
ZMAT2	10
SRPK1	9
TXNRD1	9
BANK1	8
IFI16	8
KRT8	8
RUFY1	8
SMN1	8
SSNA1	8
HAGH	7
HOXB6	7
RDBP	7
TXNL2	7

The proteins listed in Table 2 are more frequently bound by sera from diseased than control subjects. The proteins listed in Table 2 are considered as biomarkers useful in diagnosis of lupus.

The normalized data of the training set for the proteins selected in each cross-validation fold were used to construct a classification algorithm. An algorithm was established based

on using logistic regression (Liao) algorithms. Similar methods that can be used to construct classification algorithms for the sample data of this study include, amongst others, PAM (Tibshirani), artificial neural networks (Bessant), support vector machine (SVM) (Brown) and regularized discriminant analysis (Guo). Such data analysis methods

5   are known to anyone skilled in the art of microarray analysis. The constructed algorithm was then used to classify the samples of the test set into diseased and control samples in order to check its validity as classification method for lupus. In order to get an estimate of the classification errors the mean specificity and sensitivity were examined using 10 fold cross validation. The cross validation mean sensitivity was 67 %, the mean specificity was

10   85 %. The cross validation approach ensured robust measurements of classification errors.

References

Bessant, C. Microarray Analysis Software and its Applications, in The Handbook of Biosensors and Biochips (Eds. R. Marks, D. Cullen, I. Karube, H. Weetall, C. Lowe),

15   Wiley, 2007.

Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics. 2003 Jan 22;19(2):185-93.

Boutell, J.M., Hart, D.J., Godber, B.L., Kozlowski, R.Z. and Blackburn, J.M.

20   Functional protein microarrays for parallel characterisation of p53 mutants. Proteomics 2004 4(7), 1950-1958

Brown MP, Grundy WN, Lin D, Cristianini N, Sugnet CW, Furey TS, Ares M Jr, Haussler D. Knowledge-based analysis of microarray gene expression data by using support vector machines. Proc Natl Acad Sci U S A. 2000 Jan 4;97(1):262-7.

25   Chambers S.P., Austen D.A., Fulghum J.R., Kim W.M. High-throughput screening for soluble recombinant expressed kinases in Escherichia coli and insect cells. Protein Expr Purif. 2004 36(1):40-7

Fox RJ, Dimmic MW. A two-sample Bayesian t-test for microarray data.BMC Bioinformatics. 2006 Mar 10;7:126.

Guo, Y. et al. (2004) Regularized Discriminant Analysis and Its Application in Microarrays, Technical Report, Department of Statistics, Stanford University.

Huber W, von Heydebreck A, Scultmann H, Poustka A, and Vingron M. Variance stablization applied to microarray data calibration and to quantification of differential

5   expression. Bioinformatics, 18:S96{S104, 2002.

Liao JG, Chin KV. Logistic regression for disease classification using microarray data: model selection in a large p and small n case.Bioinformatics. 2007 Aug 1;23(15):1945-51.

Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc Natl Acad Sci U S A. 2002 May

10   14;99(10):6567-72.

Tusher, Tibshirani and Chu (2001). Significance analysis of microarrays applied to the ionizing radiation response. PNAS 2001 98: 5116-5121.

Witten, D. and Tibshirani, R. A comparison of fold change and the t-statistic for microarray data analysis. Technical Report, Stanford University, 2007.

15   Zhao Y, Chapman DA, Jones IM. Improving baculovirus recombination. Nucleic Acids Res. 200331(2):E6-6

Appendix 1.

Gene symbol, Entrez Gene ID, DNA Accession number and sequence of antigens.

Gene symbol: BANK1
Entrez Gene ID: 55024
dna/ac: BC032241.1
sequence:

atgctgccagcagcgccaggcaaggggcttgggagcccggaccccgccccctgcggcccagcgcccccaggaaatacaaa
agatataataatgatatatgaagaagatgctgaggaatgggctctgtacttgacagaagtatttttacatgttgtgaaaagggaagccatcctgt
tatatcgcttggagaatttctcttttcggcatttggagttgctgaacttaacgtcttacaaatgtaaactttttgatattatcaaatagcctgcttagag
acctaactccaaagaaatgtcagtttctggaaaagatacttcattcaccaaaaagtgtagttactttgctttgtggagtgaagagttcagatcag
ctctatgaattactaaatatctctcaaagcagatgggagatctcaactgaacaggaacctgaagactacatctctgtaatccagagtatcatatt
caaagattctgaagactactttgaggtcaacattccaacagacctacgagcaaaacattctggggaaataagtgagagagaaaggaaattgaa
gaactatcagaagcttcaagaaacaccataccactagcagtggtgcttcccactgaaattccatgtgagaatcctggtgaaatattcataattt
gagagatgaagtaattggtgatactgtagaggttgaatttacatcaagtaataagcgcattagaacacggccagccctttggaataagaaagt
ctggtgcatgaaagctttagagtttcctgctggttcagtccatgtcaatgtctactgtgatggaatcgttaaagctacaaccaaaattaagtacta
cccaacagcaaaggcaaaggaatgcctattcagaatggcagattcaggagagagtttgtgccagaatagcattgaagaacttgatggtgtc
cttacatccatattcaaacatgagataccatattatgagttccagtctcttcaaactgaaatttgttctcaaaacaaatatactcatttcaaagaactt
ccaactcttctccactgtgcagcaaaatttggcttaaagaacctggctattcatttgcttcaatgttcaggagcaacctgggcatctaagatgaa
aaatatggagggttcagaccccgcacatattgctgaaaggcatggtcacaaagaactcaagaaaatcttcgaagactttttcaatccaagaaat
tgacataaataatgagcaagaaaatgattatgaagaggatattgcctcattttccacatatattccttccacacagaacccagcatttcatcatga
aagcagaaagacatacgggcagagtgcagatggagctgaggcaaatgaaatggaagggggaaggaaaacagaatggatcaggcatgga
gaccaaacacagcccactagaggttggcagtgagagttctgaagaccagtatgatgacttgtatgtgttcattcctggtgctgatccagaaaa
taattcacaagagccactcatgagcagcagacctcctctccccccgccgcgacctgtagctaatgccttccaactggaaagacctcacttca
ccttaccagggacaatggtggaaggccaaatggaaagaagtcaaaactggggtcatcctggtgttagacaagaaacaggagatgaaccc
aaaggagaaaaagagaagaaagaagaggaaaaagagcaggaggaggaagaagacccatatactttgctgagattgatgacagtgaat
atgacatgatattggccaatctgagtataaagaaaaaaactgggagtcggtctttcattataaatagacctcctgcccccacaccccgaccca
caagtatacctccaaaagaggaaactacaccttacatagctcaagtgtttcaacaaaagacagccagaagacaatctgatgatgacaagttc
cgtggtcttcctaagaaacaagacagagctcggatagagagtccagcctttctactctcaggggctgtctaactgatggtcaggaagaactc
atcctcctgcaggagaaagtaaagaatgggaaaatgtctatggatgaagctctggagaaatttaaacactggcagatgggaaaaagtggcc
tggaaatgattcagcaggagaaattacgacaactacgagactgcattattgggaaaaggccagaagaagaaatgtctataataaactcacc
attgtgcaccatccaggtggtaaggaaactgcccacaatgaaaataagtttttataatgtacacttcagcaataagcttcctgctcgaccccaag
ttgaaaaggaatttggtttctgttgcaagaaagatcattaa

Gene symbol: BPY2IP1
Entrez Gene ID: 55201
dna/ac: BC006358.1
sequence:

atgggcgtgggccggctggacatgtatgtgctgcacccgccctccgccggcgccgagcgcacgctggcctgtgtgtgcgccc
tgctggtgtggcaccccgccggccccggcgagaaggtggtgcgcgtgctgttccccggttgcaccccgcccgcctgcctcctggacggc
ctggtccgcctgcagcacttgaggttcctgcgagagcccgtggtgacgccccaggacctggagggggccggggcgagccgagagc

Gene symbol: CEBPG
Entrez Gene ID: 1054
dna/ac: BC007582.1

sequence:

atgagcaagatatcgcagcaaaacagcactccaggggtgaacggaattagtgttatccatacccaggcacatgccagcggctta
cagcaggttcctcagctggtgcctgctggccctggggggaggaggcaaagctgtggctcccagcaagcagagcaaaaagagttcgcccat
ggatcgaaacagtgacgagtatcggcaacgccgagagaggaacaacatggctgtgaaaaagagccggttgaaaagcaagcagaaagc
acaagacacactgcagagagtcaatcagctcaaagaagagaatgaacggttggaagcaaaaatcaaattgctgaccaaggaattaagtgt
actcaaagatttgtttcttgagcatgcacacaaccttgcagacaacgtacagtccattagcactgaaaatacgacagcagatggcgacaatgc
aggacagtag

Gene symbol: E1B-AP5
Entrez Gene ID: 11100
dna/ac: BC009988.2
sequence:

atggacaatattaccaggcagaaccaattctacgatacccaagtcatcaaacaagaaaacgagtcaggctacgagaggagacc
actggaaatggagcagcagcaggcctatcgtccagaaatgaagacagagatgaagcaaggagcacccaccagcttcctcccgcctgaa
gcttctcaactcaagccagacaggcagcaattccagagtcgaaagaggccttatgaagaaaaccggggacgggggtactttgagcaccg
agaggataggaggggccgctctcctcagcctcctgctgaagaggatgaagatgactttgatgatacccttgttgctattgacacctataactg
cgacctccacttcaaggtggcccgagatcggagtagtggctatccgctcacaattgagggctttgcatacctgtggtcaggagcccgtgcc
agctatggggtcagaagggggccgtgtatgcttcgagatgaagatcaatgaggaaatctccgtgaagcaccttccgtctacagagcctgacc
cccacgtggtccgtatcggctggtccctggactcctgcagcacccagctaggcgaagagcctttctcctatggctatggaggcactgggaa
gaagtccaccaatagccggtttgaaaactacggagacaagtttgcagagaacgatgtgattggctgctttgcggattttgaatgtggaaatga
cgtggaactgtcttttaccaagaatggaaagtggatgggcattgctttccgaatccagaaggaagccttggggggtcaggccctctatcctca
tgtcctggtgaagaattgcgcagtggagttcaacttcggacagagagcagagccctactgttctgtcctcccggggtttaccttcatccagca
ccttccccttagtgagcgtatccggggcaccgttggaccaaagagcaaggcagaatgtgagattctgatgatggtgggcctgcctgctgct
ggcaagaccacatgggccatcaaacatgcagcctccaaccccttccaagaagtacaacatcctgggtaccaatgccatcatggataagatgc
gggtgatgggcctacgccggcagcggaactatgctggccgctgggatgtcctgatccagcaggccacccagtgcctcaaccgcctcatc
cagattgctgcccgcaagaaacgcaactatatcctagatcagacaaatgtttatgggtcagcccagagacgaaaaatgagaccatttgaag
gcttccagcgcaaagctattgtaatttgtcccactgacgaggacctaaaagaccgaacaataaagcgaaccgacgaggaaggggaaggatg
tcccagatcatgcggtcttagaaatgaaagccaacttcacgttgccagatgttggggacttcctggatgaggttctgttcattgagctgcagcg
ggaggaagcggacaagctagtgaggcagtacaacgaggaaggccgcaaggctgggccacccccctgaaaagcgctttgacaaccgag
gtggtggtggcttccgggggccgcgggggtggtggtggcttccagcgctatgaaaaccgaggacccccctggaggcaaccgtggcggctt
ccagaaccgagggggaggcagcggtggaggaggcaactaccgaggaggtttcaaccgcagcggaggtggtggctatagccagaacc
gctgggggtaacaacaaccgggataacaacaactccaacaacagaggcagctacaaccgggctccccagcaacagccgccaccacagc
agcctccgccaccacagccaccaccccagcagccaccgccaccacccagctacagccctgctcggaacccccccaggggccagcacct
acaataagaacagcaacatccctggctcaagcgccaataccagcacccccaccgtcagcagctacagccctccacagccgagttacagc
cagccaccctacaaccagggaggttacagccagggctacacagccccaccgcctccacctccaccaccacctgcctacaactatgggag
ctacggcggttacaacccggcccctataccccaccgccacccccaccgcacagacctaccctcagcccagctataaccagtatcagca
gtatgcccagcagtggaaccagtactatcagaaccagggccagtggccgccatactacgggaactacgactacgggagctactccggga
acacacagggtggcacaagtacacagtag

Gene symbol: FUS
Entrez Gene ID: 2521
dna/ac: BC000402.2
sequence:

atggcctcaaacgattatacccaacaagcaacccaaagctatggggcctaccccacccagcccgggcagggctattcccagca
gagcagtcagccctacggacagcagagttacagtggttatagccagtccacggacacttcaggctatggccagagcagctattcttcttatg

gccagagccagaacacaggctatggaactcagtcaactccccagggatatggctcgactggcggctatggcagtagccagagctcccaa
tcgtcttacgggcagcagtcctcctatcctggctatggccagcagccagctcccagcagcacctcgggaagttacggtagcagttctcaga
gcagcagctatgggcagccccagagtgggagctacagccagcagcctagctatggtggacagcagcaaagctatggacagcagcaaa
gctataatccccctcagggctatggacagcagaaccagtacaacagcagcagtggtggtggaggtggaggtggaggtggaggtaactat
ggccaagatcaatcctccatgagtagtggtggtggcagtggtggcggttatggcaatcaagaccagagtggtggaggtggcagcggtgg
ctatggacagcaggaccgtggaggccgcggcaggggtggcagtggtggcggcggcggcggcggcggtggtggttacaaccgcagca
gtggtggctatgaacccagaggtcgtggaggtggccgtggaggcagaggtggcatgggcggaagtgaccgtggtggcttcaataaattt
ggtggccctcgggaccaaggatcacgtcatgactccgaacaggataattcagacaacaacaccatctttgtgcaaggcctgggtgagaatg
ttacaattgagtctgtggctgattacttcaagcagattggtattattaagacaaacaagaaaacgggacagcccatgattaatttgtacacagac
agggaaactggcaagctgaagggagaggcaacggtctcttttgatgacccaccttcagctaaagcagctattgactggtttgatggtaaaga
attctccggaaatcctatcaaggtctcatttgctactcgccgggcagactttaatcggggtggtggcaatggtcgtggaggccgagggcgag
gaggacccatgggccgtggaggctatggaggtggtggcagtggtggtggtggccgaggaggatttcccagtggaggtggtggcggtgg
aggacagcagcgagctggtgactggaagtgtcctaatcccacctgtgagaatatgaacttctcttggaggaatgaatgcaaccagtgtaag
gcccctaaaccagatggcccaggaggggggaccaggtggctctcacatggggggtaactacggggatgatcgtcgtggtggcagaggag
gctatgatcgaggcggctaccggggccgcggcggcggggaccgtggaggcttccgagggggccgggggtggtggggacagaggtggcttt
ggccctggcaagatggattccaggggtgagcacagacaggatcgcagggagaggccgtattaa

Gene symbol: HAGH
Entrez Gene ID: 3029
dna/ac: BC000840.1
sequence:

atgaaggtagaggtgctgcctgccctgaccgacaactacatgtacctggtcattgatgatgagaccaaggaggctgccattgtg
gatccggtgcagccccagaaggtcgtggacgcggcgagaaagcacggggtgaaactgaccacagtgctcaccacccaccaccactgg
gaccatgctggcgggaatgagaaactggtcaagctggagtcgggactgaaggtgtacggggggtgacgaccgtatcggggccctgactca
caagatcactcacctgtctacactgcaggtggggtctctgaacgtcaagtgcctggcgacccccgtgccacacttcaggacacatttgttactt
cgtgagcaagcccggaggctcggagcccccctgccgtgttcacaggtgacaccttgtttgtggctggctgcgggaagttctatgaagggact
gcggatgagatgtgtaaagctctgctggaggtcttgggccggctcccccccggacacaagagtctactgtggccacgagtacaccatcaac
aacctcaagtttgcacgccacgtggagcccggcaatgccgccatccgggagaagctggcctgggccaaggagaagtacagcatcggggg
agcccacagtgccatccaccctggcagaggagtttacctacaaccccttcatgagagtgagggagaagacggtgcagcagcacgcaggt
gagacggacccggtgaccaccatgcggggccgtgcgcagggagaaggaccagttcaagatgccccgggactga

Gene symbol: HMG20B
Entrez Gene ID: 10362
dna/ac: BC002552.2
sequence:

atgtcccacggccccaagcagcccggcgcggccgccgcgccggcgggcggcaaggctccgggccagcatggggggcttcg
tggtgactgtcaagcaagagcgcggcgagggtccacgcgcgcgggcgagaaggggtcccacgaggaggagccggtgaagaaacgcgg
ctggcccaagggcaagaagcggaagaagattctgccgaatgggcccaaggcaccggtcacgggctacgtgcgcttcctgaacgagcgg
cgcgagcagatccgcacgcgccacccggatctgcccttttcccgagatcaccaagatgctgggcgccgagtggagcaagctgcagccaa
cggaaaagcagcggtacctggatgaggccgagagagagaagcagcagtacatgaaggagctgcgggcgtaccagcagtctgaagcct
ataagatgtgcacggagaagatccaggagaagaagatcaagaaagaagactcgagctctgggctcatgaacactctcctgaatggacaca
agggtggggactgcgatggcttctccaccttcgatgttcccatcttcactgaagagttcttggaccaaaacaaagcgcgtgaggcggagctt
cggcgcttgcggaagatgaatgtggccttcgaggagcagaacgcggtactgcagaggcacacgcagagcatgagcagcgcgcgcgag
cgtctggagcaggagctggcgctggaggagcggaggacgctggcgctgcagcagcagctccaggccgtgcgccagcgctcaccgc
cagcttcgcctcactgccggtgccgggcacgggcgaaacgcccacgctgggcactctggacttctacatggcccggcttcacggagccat
cgagcgcgacccccgcccagcacgagaagctcatcgtccgcatcaaggaaatcctggcccaggtcgccagcgagcacctgtga

Gene symbol: HOXB6
Entrez Gene ID: 3216
dna/ac: BC014651.1
sequence:

atgagttcctatttcgtgaactccaccttccccgtcactctggccagcgggcaggagtccttcctgggccagctaccgctctattcg
tcgggctatgcggacccgctgagacattaccccgcgccctacgggccagggccgggccaggacaagggctttgccacttcctcctattac
ccgccggcgggcggtggctacggccgagcggcgccctgcgactacgggccggcgccggccttctaccgcgagaaagagtcggcctg
cgcactctccggcgccgacgagcagcccccgttccaccccgagccgcggaagtcggactgcgcgcaggacaagagcgtgttcggcga
gacagaagagcagaagtgctccactccggtctacccgtggatgcagcggatgaattcgtgcaacagttcctcctttgggcccagcggccg
gcgaggccgccagacatacacacgttaccagacgctggagctggagaaggagtttcactacaatcgctacctgacgcggcggcggcgc
atcgagatcgcgcacgccctgtgcctgacggagaggcagatcaagatatggttccagaaccgacgcatgaagtggaaaaaggagagca
aactgctcagcgcgtctcagctcagtgccgaggaggaggaagaaaaacaggccgagtga

Gene symbol: IFI16
Entrez Gene ID: 3428
dna/ac: BC017059.1
sequence:

atgggaaaaaaatacaagaacattgttctactaaaaggattagaggtcatcaatgattatcattttagaatggttaagtccttactgag
caacgatttaaaacttaatttaaaaatgagagaagagtatgacaaaattcagattgctgacttgatggaagaaaaagttccgaggtgatgctggt
ttgggcaaactaataaaaattttcgaagatataccaacgcttgaagacctggctgaaactcttaaaaaagaaaagttaaaagtaaaaggacca
gccctatcaagaaagaggaagaaggaagtggatgctacttcacctgcaccctccacaagcagcactgtcaaaactgaaggagcagaggc
aactcctggagctcagaaaagaaaaaaaatcaaccaaagaaaaggctggacccaaaggagtaaggtgtccgaggaacagactcagcct
ccctctcctgcaggagccggcatgtccacagccatgggccgttccccatctcccaagacctcattgtcagctccacccaacacttcttcaact
gagaacccgaaaacagtggccaaatgtcaggtaactcccagaagaaatgttctccaaaaacgcccagtgatagtgaaggtactgagtaca
acaaagccatttgaatatgagacccccagaaatggagaaaaaaataatgtttcatgctacagtggctacacagacacagttcttccatgtgaag
gttttaaacaccagcttgaaggagaaattcaatggaaagaaaatcatcatcatatcagattatttggaatatgatagtctcctagaggtcaatga
agaatctactgtatctgaagctggtcctaaccaaacgtttgaggttccaaataaaatcatcaacagagcaaaggaaactctgaagattgatatt
cttcacaaacaagcttcaggaaatattgtatatggggtatttatgctacataagaaaacagtaaatcagaagaccacaatctacgaaattcagg
atgatagaggaaaaatggatgtagtgggggacaggacaatgtcacaatatcccctgtgaagaaggagataagctccaacttttctgctttcgac
ttagaaaaaagaaccagatgtcaaaactgatttcagaaatgcatagtttttatccagataaagaaaaaaaacaaacccgagaaacaatgacccc
aagagcatgaagctaccccaggaacagagtcagcttccaaatccttcagaggccagcacaaccttccctgagagccatcttcggactcctc
agatgccaccaacaactccatccagcagtttcttcaccaagaaaagtgaagacacaatctccaaaatgaatgacttcatgaggatgcagata
ctgaaggaagggagtcattttccaggaccgttcatgaccagcataggcccagctgagagccatcccccacactcctcagatgcctccatcaa
caccaagcagcagtttcttaaccacgttgaaaccaagactgaagactgaacctgaagaagtttccatagaagacagtgcccagagtgacct
caaagaagtgatggtgctgaacgcaacagaatcatttgtatatgagcccaaagagcagaagaaaatgtttcatgccacagtggcaactgag
aatgaagtcttccgagtgaaggtttttaatattgacctaaaggagaagttcaccccaaagaagatcattgccatagcaaattatgtttgccgcaa
tgggttcctggaggtatatcctttcacacttgtggctgatgtgaatgctgaccgaaacatggagatcccaaaaggattgattagaagtgccag
cgtaactcctaaaatcaatcagctttgctcacaaactaaaggaagttttgtgaatggggtgtttgaggtacataagaaaaatgtaaggggtgaa
ttcacttattatgaaatacaagataatacagggaagatggaagtggtggtgcatggacgactgaccacaatcaactgtgaggaaggagataa
actgaaactcacctgctttgaattggcaccgaaaagtgggaataccggggagttgagatctgtaattcatagtcacatcaaggtcatcaagac
caggaaaaacaagaaagacatactcaatcctgattcaagtatggaaacttcaccagactttttcttctaa

Gene symbol: KRT8
Entrez Gene ID: 3856

dna/ac: BC008200.1

sequence:

atgaacaaggtagagctggagtctcgcctggaaggggctgaccgacgagatcaacttcctcaggcagctatatgaagaggagat
ccggggagctgcagtcccagatctcggacacatctgtggtgctgtccatggacaacagccgctccctggacatggacagcatcattgctgag
gtcaaggcacagtacgaggatattgccaaccgcagccgggctgaggctgagagcatgtaccagatcaagtatgaggagctgcagagcct
ggctgggaagcacggggatgacctgcggcgcacaaagactgagatctctgagatgaaccggaacatcagccggctccaggctgagatt
gagggcctcaaaggccagaggggcttccctggaggccgccattgcagatgccgagcagcgtggagagctggccattaaggatgccaacg
ccaagttgtccgagctggaggccgccctgcagcgggccaagcaggacatggcgcggcagctgcgtgagtaccaggagctgatgaacgt
caagctggccctggacatcgagatcgccacctacaggaagctgctggagggcgaggagagccggctggagtctgggatgcagaacatg
agtattcatacgaagaccaccagcggctatgcaggtggtctgagctcggcctatgggggcctcacaagccccggcctcagctacagcctg
ggctccagctttggctctggcgcgggctccagctccttcagccgcaccagctcctccagggccgtggttgtgaagaagatcgagacacgt
gatgggaagctggtgtctgagtcctctgacgtcctgcccaagtga

Gene symbol: LIN28

Entrez Gene ID: 79727

dna/ac: BC028566.2

sequence:

atgggctccgtgtccaaccagcagtttgcaggtggctgcgccaaggcggcagaagaggcgccccgaggaggcgccggagga
cgcggcccgggcggcggacgagcctcagctgctgcacggtgcgggcatctgtaagtggttcaacgtgcgcatggggttcggcttcctgtc
catgaccgcccgcgccgggggtcgcgctcgaccccccagtggatgtctttgtgcaccagagtaagctgcacatggaagggttccggagctt
gaaggagggtgaggcagtggagttcacctttaagaagtcagccaagggtctggaatccatccgtgtcaccggacctggtggagtattctgt
attgggagtgagaggcggccaaaaggaaagagcatgcagaagcgcagatcaaaaggagacaggtgctacaactgtggaggtctagatc
atcatgccaaggaatgcaagctgccacccagcccaagaagtgccacttctgccagagcatcagccatggtagcctcatgtccgctgaa
ggcccagcagggccctagtgcacagggaaagccaacctactttcgagaggaagaagaagaaatccacagccctaccctgctcccggag
gcacagaattga

Gene symbol: NDUFV3

Entrez Gene ID: 4731

dna/ac: BC021217.2

sequence:

atggctgccccgtgtgtttgctgcggcaaggacgagccggggcgctgaagactatgctccaggaagcccaggtgtttcgaggact
tgcttctacggtttctttgtctgcggaatcagggaagagtgaaaagggtcagccacagaattccaagaagcaaagtccaccaaaaaatgtag
tggaaccaaaggagagggggcaagctcctagccacccagacagcagctgaattgtctaaaaacttatcttcacccagttcttacccgccagct
gtgaataagggcaggaaggtagctagtcccagtcccagtggcagcgtgctattcacagatgaagggggttccgaaatttttgtcaagaaaga
ctttggtagagtttccacagaaagttctgtctccattcagaaaacagggctctgattcagaagctcgtcaggtgggtcggaaagtgacgtcgc
cttcgtcttcatcctcgtccagctcctctgattctgaatctgatgatgaggctgacgtttcagaggtcactcctcgagtggtgagcaaaggcag
aggggggcttcgaaaaccagaggcctctcattcctttgaaaacagagccccccgagttacagtatcagcaaaagagaaaaccttgctgcag
aagccgcatgtggacattactgatccagagaagccccaccagccaaagaagaaagggtcccctgctaagccatcagaaggcagggaaa
atgcgagaccaaaaaccacaatgcccagatctcaagtagatgaagagttttttgaagcaaagtttaaaggaaaaacaattgcagaaaacattt
agattaaatgaaatagataaagaaagccaaaagccatttgaagttaaaggacccttacctgtccacacaaaatcagggttgtctgcgccacc
gaagggcagcccagcgcctgctgtgttggcagaagaggccagagcagagggggcagctgcaagccagtcctcctggggcggcagagg
ggcatctggaaaaacccgtgccagagccccagcgcaaggcggcccctcccctgcccagaaaggaaacctcagggacgcagggaatag
aaggccacctgaagggtggacaggcaatcgtggaagatcagataccaccaagcaatttggagacagttcctgttgagaataaccacggtt
ccatgaaaagacagcagcgctgaagcttgaggccgagggcgaggccatggaagatgcagccgcgccagggaacgaccgaggcggc
acacaggagccagccccagtgcctgctgagccgtttgacaacactacctacaagaacctgcagcatcatgactacagcacgtacaccttctt
agacctcaacctcgaactctcaaaattcaggatgcctcagccctcctcaggccgggagtcacctcgacactga

Gene symbol: PABPC1
Entrez Gene ID: 26986
dna/ac: BC015958.2
sequence:

atgaaccccagtgcccccagctaccccatggcctcgctctacgtgggggacctccaccccgacgtgaccgaggcgatgctcta
cgagaagttcagcccggccgggcccatcctctccatccgggtctgcagggacatgatcacccgccgctccttgggctacgcgtatgtgaa
cttccagcagccggcggacgcggagcgtgctttggacaccatgaattttgatgttataaagggcaagccagtacgcatcatgtggtctcagc
gtgatccatcacttcgcaaaagtggagtaggcaacatattcattaaaaatctggacaaatccattgataataaagcactgtatgatacattttctg
cttttggtaacatcctttcatgtaaggtggtttgtgatgaaaatggttccaagggctatggatttgtacactttgagacgcaggaagcagctgaa
agagctattgaaaaaatgaatggaatgctcctaaatgatcgcaaagtatttgttggacgatttaagtctcgtaaagaacgagaagctgaacttg
gagctagggcaaaagaattcaccaatgtttacatcaagaattttggagaagacatggatgatgagcgccttaaggatctctttggcaagtttgg
gcctgccttaagtgtgaaagtaatgactgatgaaagtggaaaatccaaaggatttggatttgtaagctttgaaaggcatgaagatgcacagaa
agctgtggatgagatgaacggaaaggagctcaatggaaaacaaatttatgttggtcgagctcagaaaaaggtggaacggcagacggaact
taagcgcaaatttgaacagatgaaacaagataggatcaccagataccagggtgttaatctttatgtgaaaaatcttgatgatggtattgatgatg
aacgtctccggaaagagttttctccatttggtacaatcactagtgcaaaggttatgatggagggtggtcgcagcaaagggtttggttttgtatgt
ttctcctccccagaagaagccactaaagcagttacagaaatgaacggtagaattgtggccacaaagccattgtatgtagctttagctcagcgc
aaagaagagcgccaggctcacctcactaaccagtatatgcagagaatggcaagtgtacgagctgttcccaaccctgtaatcaacccctacc
agccagcacctccttcaggttacttcatggcagctatcccacagactcagaaccgtgctgcatactatcctcctagccaaattgctcaactaag
accaagtcctcgctggactgctcagggtgccagacctcatccattccaaaatatgcccggtgctatccgcccagctgctcctagaccaccatt
tagtactatgagaccagcttcttcacaggttccacgagtcatgtcaacacagcgtgttgctaacacatcaacacagacaatgggtccacgtcc
tgcagctgcagccgctgcagctactcctgctgtccgcaccgttccacagtataaatatgctgcaggagttcgcaatcctcagcaacatcttaat
gcacagccacaagttacaatgcaacagcctgctgttcatgtacaaggtcaggaacctttgactgcttccatgttggcatctgcccctcctcaag
agcaaaagcaaatgttgggtgaacggctgtttcctcttattcaagccatgcaccctactcttgctggtaaaatcactggcatgttgttggagatt
gataattcagaacttcttcatatgctcgagtctccagagtcactccgttctaaggttgatgaagctgtagctgtactacaagcccaccaagctaa
agaggctgcccagaaagcagttaacagtgccaccggtgttccaactgtttaa

Gene symbol: PHLDA1
Entrez Gene ID: 22822
dna/ac: BC018929.2
sequence:

atgctggagagtagcggctgcaaagcgctgaaggagggcgtgctggagaagcgcagcgacgggttgttgcagctctggaag
aaaaagtgttgcatcctcaccgaggaagggctgctgcttatcccgcccaagcagctgcaacaccagcagcagcagcaacagcagcagca
gcagcagcaacaacagcccgggcaggggccggccgagccgtcccaacccagtggccccgctgtcgccagcctcgagccgccggtca
agctcaaggaactgcacttctccaacatgaagaccgtggactgtgtggagcgcaagggcaagtacatgtacttcactgtggtgatggcaga
gggcaaggagatcgactttcggtgcccgcaagaccagggctggaacgccgagatcacgctgcagatggtgcagtacaagaatcgtcag
gccatcctggcggtcaaatccacgcggcagaagcagcagcacctggtccagcagcagcccccctcgcagccgcagccgcagccgcag
ctccagccccaaccccagcctcagcctcagccgcaaccccagccccaatcacaaccccagcctcagccccaacccaagcctcagcccc
agcagctccacccgtatccgcatccacatccacatccacactctcatcctcactcgcacccacaccctcacccgcacccgcatccgcacca
ataccgcacccacacccacagccgcactcgcagccgcacgggcaccggcttctccgcagcacctccaactctgcctga

Gene symbol: PIAS2
Entrez Gene ID: 9063
dna/ac: BC015190.1

sequence:

atggcggatttcgaagagttgaggaatatggtttctagtttttagggtttctgaactacaagtattactaggctttgctggacggaataa
aagtggacgcaagcatgacctcctgatgagggcgctgcatttattgaagagcggctgcagccctgcggttcagattaaaatccgagaattgt
atagacgccgatatccacgaactcttgaaggactttctgatttatccacaatcaaatcatcggttttcagtttggatggtggctcatcacctgtag
aacctgacttggccgtggctggaatccactcgttgccttccacttcagttacacctcactcaccatcctctcctgttggttctgtgctgcttcaag
atactaagcccacatttgagatgcagcagccatctcccccaattcctcctgtccatcctgatgtgcagttaaaaaatctgccctttttatgatgtcc
ttgatgttctcatcaagcccacgagtttagttcaaagcagtattcagcgatttcaagagaagttttttatttttgctttgacacctcaacaagttaga
gagatatgcatatccagggattttttgccaggtggtaggagagattatacagtccaagttcagttgagactttgcctggcagagacaagttgcc
ctcaagaagataactatccaaatagtctatgtataaaagtaaatgggaagctatttcctttgcctggctatgcaccaccgcctaaaaatgggatt
gaacagaagcgccctggacgcccccttgaatattacatctttagttaggttatcttcagctgtgccaaaccaaatttccatttcttgggcatcagaa
attgggaagaattactctatgtctgtatatcttgtacggcagcttacatcagccatgttattacagagattaaaaatgaaaggtattagaaaccct
gatcattccagagcactaattaaagaaaaacttactgcagatcctgatagtgaaattgctacaactagccttcgggtatccttgatgtgccccttta
ggaaaaatgaggctgacaatcccatgccgtgcagtgacttgtacacatctgcagtgttttgatgctgccctctatctacaaatgaatgagaaaa
agcccacctggatttgtcctgtgtgtgacaaaaaagctgcctatgaaagtctaatattagatgggctttttatggaaattctcaatgactgttctga
tgtagatgagatcaaattccaagaagatggttcttggtgtccaatgagaccgaagaaagaagctatgaaagtatccagccaaccgtgtacaa
aaatagaaagttcaagcgtcctcagtaagccttgttcagtgactgtagccagtgaggcaagcaagaagaaagtagatgttattgatcttacaat
agaaagctcttctgacgaagaggaagaccctcctgccaaaaggaaatgcatctttatgtcagaaacacaaagcagcccaaccaaaggggt
tctcatgtatcagccatcttctgtaagggtgcccagtgtgacttcggttgatcctgctgctattccgccttcattaacagactactcagtaccattc
caccatacgccaatatcaagcatgtcatcagatttgccaggagaacaaagaagaaatgatattaataatgaactgaagcttggaacatcttct
gatactgtgcaacagtga

Gene symbol: PSME3
Entrez Gene ID: 10197
dna/ac: BC001423.2
sequence:

atggcctcgttgctgaaggtggatcaggaagtgaagctcaaggttgattctttcagggagcggatcacaagtgaggcagaagac
ttggtggcaaatttttttcccaaagaagttattagaacttgatagtttttctgaaggaaccaatcttaaacatccatgacctaactcagatccactctg
acatgaatctcccagtccctgaccccattcttctcaccaatagccatgatggactggatggtcccacttataagaagcgaaggttggatgagt
gtgaagaagccttccaaggaaccaaggtgtttgtgatgcccaatgggatgctgaaaagcaaccagcagctggtggacattattgagaaagt
gaaacctgagatccggctgttgattgagaaatgtaacacggtcaaaatgtgggtacagctcctgattcccaggatagaagatggaaacaact
ttggggtgtccattcaggaggaaacagttgcagagctaagaactgttgagagtgaagctgcatcttatctggaccagatttctagatattatatt
acaagagccaaattggtttctaaaatagctaaatatccccatgtggaggactatcgccgcaccgtgacagagattgatgagaaagaatatatc
agccttcggctcatcatatcagagctgaggaatcaatatgtcactctacatgacatgatcctgaaaaatatcgagaagatcaaacggccccg
gagcagcaatgcagagactctgtactga

Gene symbol: RAB11FIP3
Entrez Gene ID: 9727
dna/ac: BC051360.1
sequence:

atggcgtcggccccgccggcctcgcccccgggctcggagccgccggggcccgacccggagccgggcgggccggacggg
ccgggggcggcgcaactggctccgggccctgcggagctacgcctcggagcgcccgtcggcggccccgacccgcagtccccgggcct
ggatgagcctgcgcccggggccgctgcagatggcggggcgcgttggagcgccgggccggccccggggctggagggaggcccgcg
agaccccgggccgtccgccccgccgccgcgctccggcccgcgggggcagcttgcgagccccgacgccccgggcccagggccgcgc
tccgaagcgccgcttccagaactcgacccgttgttctcctggactgaggagcccgaggagtgtggcccccgcgagctgcccggagagcgc
gcctttccgcttgcaggggtccagcagcagccaccgagcgcggggcgaggtcgacgtcttctctcccttccccgcgcccacggcgggcg
agctggcgctggagcaaggtcccgggtccccgccgcagccctcggacctcagccagacccaccccccttccgagcgagcccgtggggga

gtcaggaggacggcccccgcctccgagccgtgttcgatgccctggacggggatggggacggtttcgtccgcatcgaggacttcatccagt
ttgctacggtctacggggcagagcaggtgaaggacttaactaagtacttggatcccagtgggctcggcgtgatcagctttgaagacttctac
caagggatcacagccatcagaaacggagatcctgatggccagtgctacggtggtgtcgcttctgcccaagatgaggagcccctggcctgc
ccggacgagttcgatgacttcgtcacctatgaggccaacgaggtgacggacagcgcgtacatgggctccgagagcacctacagtgagtgt
gagaccttcacggacgaggacaccagcaccctggtgcaccctgagctgcaacctgaaggggacgcagacagtgccggcggctcggcc
gtgccctctgagtgcctggacgccatggaggagcccgaccatggtgccctgctgctgctcccaggcaggcctcaccccatggccagtct
gtcatcacggtgatcggggggcgaggagcactttgaggactacggtgaaggcagtgaggcggagctgtccccagagaccctatgcaacg
ggcagctgggctgcagtgaccccgctttcctcacgcccagtccgacaaagcggctctccagcaagaaggtggcaaggtacctgcaccag
tcaggggccctgaccatggaggccctggaggacccttcccccgagctcatggagggcccagaggaggacattgctgacaaggttgtcttc
ctggaaaggcgtgtgctggagctggaaaaggacacggcagccaccggtgagcaacacagccgcctgaggcaggagaacctgcagctg
gtgcacagagcaaacgccctggaggagcagctgaaggagcaggagctgagagcctgcgagatggtcctggaagagacccggcgtca
gaaggagctcctgtgcaagatggagagggagaagagcattgagatcgagaacctgcagaccaggctacagcaactggacgaggagaa
cagtgaactccggtcctgcacgccctgtctgaaggccaacattgagcgtctggaggaggagaagcagaagctgttggatgagatagagtc
gctgacgctgcggctcagtgaagagcaggagaacaagaggagaatgggggacaggctgagtcacgagaggcaccagttccagaggg
acaaggaggccacccaggagctgatcgaggacctccgaaagcagctggagcacctgcagctcctcaagctggaggccgagcagcggc
gggggccgcagcagcagcatgggcctgcaggagtaccacagccgcgcccgggagagcgagctggagcaggaggtccgcaggctgaa
gcaggacaaccgcaacctgaaggagcagaacgaggagctgaacgggcagatcattaccctcagcatccagggcgccaagagcctcttc
tccacagccttctctgagtccctggctgcagagatcagctccgtctcccgagatgagctcatggaggcgattcagaagcaggaggagatca
acttccgcctgcaggactacatcgacaggatcatcgtggccatcatggagaccaacccgtccatcctggaggtcaagtag

Gene symbol: RALBP1
Entrez Gene ID: 10928
dna/ac: BC013126.1
sequence:

atgactgagtgcttcctgcccccccaccagcagccccagtgaacaccgcagggtggagcatggcagcgggcttacccggaccc
ccagctctgaagagatcagccctactaagtttcctggattgtaccgcactggcgagccctcacctccccatgacatcctccatgagcctcctg
atgtagtgtctgatgatgagaaagatcatgggaagaaaaaagggaaatttaagaaaaaggaaaagaggactgaaggctatgcagcctttca
ggaagatagctctggagatgaggcagaaagtccttctaaaatgaagaggtccaagggaatccatgttttcaagaagcccagcttttctaaaa
agaaggaaaaggattttaaaataaaagagaaacccaaagaagaaaagcataaagaagaaaagcacaaagaagaaaaacataaagagaa
gaagtcaaaagacttgacagcagctgatgttgttaaacagtggaaggaaaagaagaaaaagaaaaagccaattcaggagccagaggtgc
ctcagattgatgttccaaatctcaaacccatttttggaattcctttggctgatgcagtagagaggaccatgatgtatgatggcattcggctgcca
gccgttttccgtgaatgtatagattacgtagagaagtatggcatgaagtgtgaaggcatctacagagtatcaggaattaaatcaaaggtggat
gagctaaaagcagcctatgaccgggaggagtctacaaacttggaagactatgagcctaacactgtagccagtttgctgaagcagtatttgcg
agaccttccagagaatttgcttaccaaagagcttatgcccagatttgaagaggcttgtgggaggaccacggagactgagaaagtgcaggaa
ttccagcgtttactcaaagaactgccagaatgtaactatcttctgatttcttggctcattgtgcacatggaccatgtcattgcaaaggaactggaa
acaaaaatgaatatacagaacatttctatagtgctcagcccaactgtgcagatcagcaatcgagtcctgtatgtgttttcacacatgtgcaaga
actctttggaaatgtggtactaaagcaagtgatgaaacctctgcgatggtctaacatggccacgatgcccacgctgccagagacccaggcg
ggcatcaaggaggagatcaggagacaggagtttcttttgaattgtttacatcgagatctgcaggtgtgggataaaggatttgtctaaagaagaa
agattatgggaagtacaaagaattttgacagccctcaaaagaaaactgagagaagctaaaagacaggagtgtgaaaccaagattgcacaa
gagatagccagtctttcaaaagaggatgtttccaaagaagagatgaatgaaaatgaagaagttataaatattctccttgctcaggagaatgag
atcctgactgaacaggaggagctcctggccatggagcagtttctgcgccggcagattgcctcagaaaaagaagagattgaacgcctcaga
gctgagattgctgaaattcagagtcgccagcagcacggccgaagtgagactgaggagtactcctccgagagcgagagcgagagtgagg
atgaggaggagctgcagatcattctggaagacttacagagacagaacgaagagctggaaataaagaacaatcatttgaatcaagcaattca
tgaggagcgcgaggccatcatcgagctgcgcgtgcagctgcggctgctccagatgcagcgagccaaggccgagcagcaggcgcagg
aggacgaggagcctgagtggcgcggggggtgccgtccagccgcccagagacggcgtccttgagccaaaagcagctaaagagcagcca
aaggcaggcaaggagccggcaaagccatcgcccagcagggataggaaggagacgtccatctga

Gene symbol: RAN
Entrez Gene ID: 5901
dna/ac: BC014518.2
sequence:

atggctgcgcagggagagccccaggtccagttcaaacttgtattggttggtgatggtggtactggaaaaacgaccttcgtgaaac
gtcatttgactggtgaatttgagaagaagtatgtagccaccttgggtgttgaggttcatcccctagtgttccacaccaacagaggacctattaa
gttcaatgtatgggacacagccggccaggagaaattcggtggactgagagatggctattatatccaagcccagtgtgccatcataatgtttga
tgtaacatcgagagttacttacaagaatgtgcctaactggcatagagatctggtacgagtgtgtgaaaacatccccattgtgttgtgtggcaac
aaagtggatattaaggacaggaaagtgaaggcgaaatccattgtcttccaccgaaagaagaatcttcagtactacgacatttctgccaaaagt
aactacaactttgaaaagcccttcctctggcttgctaggaagctcattggagaccctaacttggaatttgttgccatgcctgctctcgccccacc
agaagttgtcatggacccagctttggcagcacagtatgagcacgacttagaggttgctcagacaactgctctcccggatgaggatgatgacc
tgtga

Gene symbol: RARA
Entrez Gene ID: 5914
dna/ac: BC008727.2
sequence:

atggccagcaacagcagctcctgcccgacacctgggggcgggcacctcaatgggtacccggtgcctccctacgccttcttcttc
cccctatgctgggtggactctccccgccaggcgctctgaccactctccagcaccagcttccagttagtggatatagcacaccatccccagc
caccattgagacccagagcagcagttctgaagagatagtgcccagccctccctcgccaccccctctaccccgcatctacaagccttgctttg
tctgtcaggacaagtcctcaggctaccactatggggtcagcgcctgtgagggctgcaagggcttcttccgccgcagcatccagaagaacat
ggtgtacacgtgtcaccgggacaagaactgcatcatcaacaaggtgacccggaaccgctgccagtactgccgactgcagaagtgctttga
agtgggcatgtccaaggagtctgtgagaaacgaccgaaacaagaagaagaaggaggtgcccaagcccgagtgctctgagagctacacg
ctgacgccggaggtggggggagctcattgagaaggtgcgcaaagcgcaccaggaaaccttccctgccctctgccagctgggcaaatacac
tacgaacaacagctcagaacaacgtgtctctctggacattgacctctgggacaagttcagtgaactctccaccaagtgcatcattaagactgt
ggagttcgccaagcagctgcccggcttcaccaccctcaccatcgccgaccagatcaccctcctcaaggctgcctgcctggacatcctgatc
ctgcggatctgcacgcggtacacgcccgagcaggacaccatgaccttctcggacgggctgaccctgaaccggacccagatgcacaacg
ctggcttcggccccctcaccgacctggtctttgccttcgccaaccagctgctgcccctggagatggatgatgcggagacggggctgctcag
cgccatctgcctcatctgcggagaccgccaggacctggagcagccggaccgggtggacatgctgcaggagccgctgctggaggcgcta
aaggtctacgtgcggaagcgggaggcccagccgcccccacatgttccccaagatgctaatgaagattactgacctgcgaagcatcagcgcc
aaggggggctgagcgggtgatcacgctgaagatggagatcccgggctccatgccgcctctcatccaggaaatgttggagaactcagaggg
cctggacactctgagcggacagccggggggtgggggcgggacggggtggcctggcccccccgccaggcagctgtagccccagcc
tcagccccagctccaacagaagcagcccggccacccactccccgtga

Gene symbol: RBMS1
Entrez Gene ID: 5937
dna/ac: BC012992.2
sequence:

atgggcaaagtgtggaaacagcagatgtaccctcagtacgccacctactattaccccccagtatctgcaagccaagcagtctctgg
tcccagcccacccatggcccctcccagtcccagcaccaccagcagtaataacaacagtagcagcagtagcaactcaggatgggatcag
ctcagcaaaacgaacctctatatccgaggactgcctccccacaccaccgaccaggacctggtgaagctctgtcaaccatatgggaaaatag
tctccacaaaggcaattttggataagacaacaaacaaatgcaaaggttatggttttgtcgactttgacagccctgcagcagctcaaaaagctgt
gtctgccctgaaggccagtggggttcaagctcaaatggcaaagcaacaggaacaagatcctaccaacctctacatttctaatttgccactctc
catggatgagcaagaactagaaaatatgctcaaaccatttggacaagttatttctacaaggatactacgtgattccagtggtacaagtcgtggt
gttggctttgctaggatggaatcaacagaaaaatgtgaagctgttattggtcattttaatggaaaatttattaagacaccaccaggagtttctgcc

cccacagaacctttattgtgtaagtttgctgatggaggacagaaaaagagacagaacccaaacaaatacatccctaatggaagaccatggc
atagagaaggagaggctggaatgacacttacttacgacccaactacagctgctatacagaacggattttatccttcaccatacagtattgctac
aaaccgaatgatcactcaaacttctattacaccctatattgcatctcctgtatctgcctaccaggtgcaaagtccttcgtggatgcaacctcaac
catatattctacagcaccctggtgccgtgttaactccctcaatggagcacaccatgtcactacagcccgcatcaatgatcagccctctggccc
agcagatgagtcatctgtcactaggcagcaccggaacatacatgcctgcaacgtcagctatgcaaggagcctacttgccacagtatgcaca
tatgcagacgacagcggttcctgttgaggaggcaagtggtcaacagcaggtggctgtcgagacgtctaatgaccattctccatatacctttca
acctaataagtaa

Gene symbol: RDBP
Entrez Gene ID: 7936
dna/ac: BC050617.2
sequence:

atgttggtgatacccccggactgagcgaggaagaggaggctctgcagaagaaattcaacaagctcaagaaaaagaaaaagg
cattgctggctctgaagaagcaaagtagcagcagcacaaccagccaaggtggtgtcaaacgctcactatcagagcagcctgtcatggaca
cagccacagcaacagagcaggcaaagcagctggtgaagtcaggagccatcagtgccatcaaggctgagaccaagaactcaggcttcaa
gcgttctcgaacccttgaggggaagttaaaggaccccgagaagggaccagtccccactttccagccgttccagaggagcatatctgctgat
gatgacctgcaagagtcatccagacgtccccagaggaaatctctgtatgagagctttgtgtcttctagtgatcgacttcgagaactaggacca
gatggagaagaggcagagggcccaggggctggtgatggtcccctcgaagctttgactggggctatgaagaacgcagtggtgcccactc
ctcagcctcccctccccgaagccgcagccgggaccgcagccatgagaggaaccgggacagagaccgagatcgggagcgggatcgag
accgggatcgagacagagacagagagcgggacaggatcgggatcgggatcgagatcgagaccgggaacgggacaggatcggga
gcgggatcgagaccgagaccgagagggtcctttccgcaggtcggattcattccctgaacggcgagcccctaggaaagggaatactctcta
tgtatatggagaagacatgacacccacccttctccgtggggccttctctccttttggaaacatcattgacctctccatggacccacccagaaac
tgtgccttcgtcacctatgaaaagatggagtcagcagatcaggccgttgctgagctcaacgggacccaggtggagtctgtacagctcaaag
tcaacatagcccgaaaacagcccatgctggatgccgctactggcaagtctgtctggggctccctcgctgtccagaacagccctaagggttg
ccaccgggacaagaggacccagattgtctacagtgatgacgtctacaaggaaaaccttgtggatggcttctag

Gene symbol: RNF12
Entrez Gene ID: 51132
dna/ac: BC013357.2
sequence:

atggaaaactcagattccaatgacaaaggaagtggtgatcagtctgcagcacagcgcagaagtcagatggaccgattggatcg
agaagaagctttctatcaatttgtaaataacctgagtgaagaagattataggcttatgagagataacaatttgctaggcaccccaggtgaaagt
actgaggaagagttgctgagacgactacagcaaattaaagaaggcccaccaccgcaaaactcagatgaaaatagaggaggagactcttc
agatgatgtgtctaatggtgactctataatagactggcttaactctgtcagacaaactggaaatacaacaagaagtgggcaaagaggaaacc
aatcttggagagcagtgagtcggactaatccaaacagtggtgatttcagattcagtttagagatataatgttaaccgtaataatgggagccaaa
attcagagaatgaaaatgagccatctgcaagacgttctagtggagaaaatgtggaaaacaacagccaaaggcaagtggaaaacccacgat
ctgaatcaacatctgcaaggccatctagatcagaacgaaattcaactgaagcgttaacagaggtcccacctaccagaggtcagaggaggg
caagaagcaggagcccagaccatcggagaaccagagcaagagctgaaagaagtaggtcacctctgcatccaatgagtgaaattccacga
agatctcatcatagtatctcatctcagactttgaacatcctttggtaaatgagacggagggaagttctagaacccggcaccatgtgacattga
ggcagcaaatatctgggcctgagttgctaagtagaggtctttttgcagcttctggaacaagaaatgcttctcaaggagcaggttcttcagacac
agctgccagtggtgaatctacaggatcaggacagagacctccaaccatagtccttgatcttcaagtaagaagagttcgtcctggagaatatc
ggcagagagatagcatagccagcagaactcggtctaggtctcagacaccaaacaacactgtcacctatgaaagtgaacgaggaggtttta
ggcgtacattttcacgttctgagcgggcaggtgtgagaacctatgtcagtaccatcagaattcccattcgtagaatcttaaatactggtttaagt
gagactacatctgttgcaattcagaccatgttaaggcagataatgacaggttttggtgagttaagctattttatgtacagtgatagcgactcaga
gcctactggctcagtctcaaatcgaaatatggaaagggcagagtcacggagtggaagaggaggttctggtggtggtagtagttctggttcca
gttcgagttccagttccagttcgagttccagttccagttcaagttccagttccagtcctagttccagttccggtggtgaaagttcagaaactagct

cagatttatttgaaggcagtaatgaaggaagctcatcatcaggctcatcaggtgccaggcgagagggtcgacatagggccccagtcacattt
gatgaaagtggctctttgcccttccttagcctggctcagtttttcctcttaaatgaggatgatgatgaccaacctagaggactcaccaaagaaca
gattgacaacttggcaatgagaagttttggtgaaaatgatgcattaaaaacctgtagtgtttgcattacagaatatacagaaggcaacaaacttc
gtaaactaccttgttcccatgagtaccatgtccactgcatcgatcgctggttatctgagaattctacctgtcctatttgtcgcagagcagtcttagc
ttctggtaacagagaaagtgttgtgtaa

Gene symbol: RUFY1
Entrez Gene ID: 80230
dna/ac: BC032571.1
sequence:

atgatggaggagcgtgccaacctgatgcacatgatgaaactcagcatcaaggtgttgctccagtcggctctgagcctgggccgc
agcctggatgcggaccatgccccccttgcagcagttctttgtagtgatggagcactgcctcaaacatgggctgaaagttaagaagagtttt att
ggccaaaataaatcattctttggtcctttggagctggtggagaaacttt gtccagaagcatcagatatagcgactagtgtcagaaatcttccag
aattaaagacagctgtgggaagaggccgagcgtggctttatcttgcactcatgcaaaagaaactggcagattatctgaaagtgcttatagaca
ataaacatctcttaagcgagttctatgagcctgaggctttaatgatggaggaagaagggatggtgattgttggtctgctggtgggactcaatgt
tctcgatgccaatctctgcttgaaaggagaagacttggattctcaggttggagtaatagattttccctctaccttaaggatgtgcaggatcttga
tggtggcaaggagcatgaaagaattactgatgtccttgatcaaaaaaattatgtggaagaacttaaccggcacttgagctgcacagttgggg
atcttcaaaccaagatagatggcttggaaaagactaactcaaagcttcaagaagagctttcagctgcaacagaccgaatttgctcacttcaag
aagaacagcagcagttaagagaacaaaatgaattaattcgagaaagaagtgaaaagagtgtagagataacaaaacaggataccaaagttg
agctggagacttacaagcaaactcggcaaggtctggatgaaatgtacagtgatgtgtggaagcagctaaaagaggagaagaaagtccggt
tggaactggaaaaagaactggagttacaaattggaatgaaaaccgaaatggaaattgcaatgaagttactggaaaaggacacccacgaga
agcaggacacactagttgccctccgccagcagctggaagaagtcaaagcgattaatttacagatgtttcacaaagctcagaatgcagagag
cagtttgcagcagaagaatgaagccatcacatcctttgaaggaaaaaccaaccaagttatgtccagcatgaaacaaatggaagaaaggttg
cagcactcggagcgggcgaggcagggagctgaggagcggagccacaagctgcagcaggagctgggcgggaggatcggcgccctgc
agctgcagctctcccagctgcacgagcaatgctcaagcctggagaaagaattgaaatcagaaaaagagcaaagacaggctcttcagcgc
gaattacagcacgagaaagacacttcctctctactcaggatggagctgcaacaagtggaaggactgaaaaaggagttgcgggagcttcag
gacgagaaggcagagctgcagaagatctgcgaggagcaggaacaagccctccaggaaatgggcctgcacctcagccagtccaagctg
aagatggaagatataaaagaagtgaaccaggcactgaagggccacgcctggctgaaagatgacgaagcgacacactgtaggcagtgtg
agaaggagttctccatttcccggagaaagcaccactgccggaactgtggccacatcttctgcaacacctgctccagcaacgagctggccct
gccctcctaccccaagccggtgcgagtgtgcgacagctgccacaccctgctcctgcagcgctgctcctccacggcctcctga

Gene symbol: SMN1
Entrez Gene ID: 6606
dna/ac: BC000908.2
sequence:

atggcgatgagcagcggcggcagtggtggcggcgtcccggagcaggaggattccgtgctgttccggcgcggcacaggcca
gagcgatgattctgacatttgggatgatacagcactgataaaagcatatgataaagctgtggcttcatttaagcatgctctaaagaatggtgac
atttgtgaaacttcgggtaaaccaaaaaccacacctaaaagaaaacctgctaagaagaataaaagccaaaagaagaatactgcagcttcctt
acaacagtggaaagttggggacaaatgttctgccatttggtcagaagacggttgcatttacccagctaccattgcttcaattgattttaagagag
aaacctgtgttgtggtttacactggatatggaaatagagaggagcaaaatctgtccgatctactttccccaatctgtgaagtagctaataatata
gaacagaatgctcaagagaatgaaaatgaaagccaagtttcaacagatgaaagtgagaactccaggtctcctggaaataaatcagataaca
tcaagcccaaatctgctccatggaactctttctccctccaccacccccatgccagggccaagactgggaccaggaaagccaggtctaaa
attcaatggcccaccaccgccaccgccaccaccaccacccacttactatcatgctggctgcctccatttccttctggaccaccaataattccc
ccaccacctcccatatgtccagattctcttgatgatgctgatgctttgggaagtatgttaatttcatggtacatgagtggctatcatactggctatta
tatggaaatgctggcatag

Gene symbol: SRPK1
Entrez Gene ID: 6732
dna/ac: BC038292.1
sequence:

atggagcggaaagtgcttgcgctccaggcccgaaagaaaaggaccaaggccaagaaggacaaagcccaaaggaaatctga
aactcagcaccgaggctctgctccccactctgagagtgatctaccagagcaggaagaggagattctgggatctgatgatgatgagcaaga
agatcctaatgattattgtaaaggaggttatcatcttgtgaaaattggagatctattcaatgggagataccatgtgatccgaaagttaggctggg
gacactttcaacagtatggttatcatgggatattcaggggaagaaatttgtggcaatgaaagtagttaaaagtgctgaacattacactgaaac
agcactagatgaaatccggttgctgaagtcagttcgcaattcagaccctaatgatccaaatagagaaatggttgttcaactactagatgacttta
aaatatcaggagttaatggaacacatatctgcatggtatttgaagttttggggcatcatctgctcaagtggatcatcaaatccaattatcagggg
cttccactgccttgtgtcaaaaaaattattcagcaagtgttacagggtcttgattatttacataccaagtgccgtatcacccacactgacattaaa
ccagagaacatcttattgtcagtgaatgagcagtacattcggaggctggctgcagaagcaacagaatggcagcgatctggagctcctccgc
cttccggatctgcagtcagtactgctcccccagcctaaaccagctgacaaaatgtcaaagaataagaagaagaaattgaagaagaagcagaa
gcgccaggcagaattactagagaagcgaatgcaggaaattgaggaaatggagaaagagtcgggccctgggcaaaaaagaccaaacaa
gcaagaagaatcagagagtcctgttgaaagacccttgaaagagaacccacctaataaaatgacccaagaaaaacttgaagagtcaagtac
cattggccaggatcaaacgcttatggaacgtgatacagagggtggtgcagcagaaattaattgcaatggagtgattgaagtcattaattatact
cagaacagtaataatgaaacattgagacataaagaggatctacataatgctaatgactgtgatgtccaaaatttgaatcaggaatctagtttcct
aagctcccaaaatggagacagcagcacatctcaagaaacagactcttgtacacctataacatctgaggtgtcagacaccatggtgtgccagt
cttcctcaactgtaggtcagtcattcagtgaacaacacattagccaacttcaagaaagcattcgggcagagatacctgtgaagatgaacaa
gagcaagaacataacggaccactggacaacaaaggaaaatccacggctggaaattttcttgttaatccccttgagccaaaaaatgcagaaa
agctcaaggtgaagattgctgaccttggaaatgcttgttgggtgcacaaacatttcactgaagatattcaaacaaggcaatatcgttccttgga
agttctaatcggatctggctataatacccctgctgacatttggagcacggcatgcatggcctttgaactggccacaggtgactatttgtttgaac
ctcattcaggggaagagtacactcgagatgaagatcacattgcattgatcatagaacttctggggaaggtgcctcgcaagctcattgtggca
ggaaaatattccaaggaatttttcaccaaaaaaggtgacctgaaacatatcacgaagctgaaaccttggggcctttttgaggttctagtggaga
agtatgagtggtcgcaggaagaggcagctggcttcacagatttcttactgcccatgttggagctgatccctgagaagagagccactgccgc
cgagtgtctccggcacccttggcttaactcctaa

Gene symbol: SSA2
Entrez Gene ID: 6738
dna/ac: BC036658.2
sequence:

atggaggaatctgtaaaccaaatgcagccactgaatgagaagcagatagccaattctcaggatggatatgtatggcaagtcactg
acatgaatcgactacaccggttcttatgtttcggttctgaaggtggggacttattatatcaaagaacagaagttgggccttgaaaatgctgaagct
ttaattagattgattgaagatggcagaggatgtgaagtgatacaagaaataaagtcatttagtcaagaaggcagaaccacaaagcaagagcc
tatgctctttgcacttgccatttgttcccagtgctccgacataagcacaaaacaagcagcatttaaagctgtttctgaagtttgtcgcattcctacc
catctctttactttatccagtttaagaaagatctgaaggaaagcatgaaatgtggcatgtggggtcgtgccctccggaaggctatagcggact
ggtacaatgagaaaggtggcatggcccttgctctggcagttacaaaatataaacagagaaatggctggtctcacaaagatcattaagattgt
cacatcttaaaccttccagtgaaggacttgcaattgtgaccaaatatattacaaagggctggaaagaagttcatgaattgtataaagaaaaagc
actctctgtggagactgaaaaattattaaagtatctggaggctgtagagaaagtgaagcgcacaagagatgagctagaagtcattcatctaat
agaagaacatagattagttagagaacatcttttaacaaatcacttaaagtctaaagaggtatggaaggctttgttacaagaaatgccgcttactg
cattactaaggaatctaggaaagatgactgctaattcagtacttgaaccaggaaattcagaagtatctttagtatgtgaaaaactgtgtaatgaa
aaactattaaaaaaggctcgtatacatccatttcatattttgatcgcattagaaacttacaagacaggtcatggtctcagagggaaactgaagtg
gcgccctgatgaagaaattttgaaagcattggatgctgcttttttataaaacatttaagacagttgaaccaactggaaaacgtttcttactagctgtt
gatgtcagtgcttctatgaaccaaagagttttgggtagtatactcaacgctagtacagttgctgcagcaatgtgcatggttgtcacacgaacag
aaaaagattcttatgtagttgctttttccgatgaaatggtaccatgtccagtgactacagatatgaccttacaacaggttttaatggctatgagtca
gatcccagcaggtggaactgattgctctcttccaatgatctgggctcagaagacaaacacacctgctgatgtcttcattgtattcactgataatg

agacctttgctggaggtgtccatcctgctattgctctgagggagtatcgaaagaaaatggatattccagctaaattgattgtttgtggaatgaca
tcaaatggtttcaccattgcagacccagatgatagaggcatgttggatatgtgcggctttgatactggagctctggatgtaattcgaaatttcac
attagatatgatttaa

Gene symbol: SSNA1
Entrez Gene ID: 8636
dna/ac: BC000864.1
sequence:

atgacccagcagggcgcggcgctgcagaactacaacaacgagctggtcaagtgcatagaggagctgtgccagaagcggga
ggagctgtgccggcagatccaggaggaggaggacgagaagcagcggctgcagaatgaggtgaggcagctgacagagaagctggccc
gcgtcaacgagaacctggcacgcaagattgcctctcgcaacgagttcgaccggaccatcgcggagacggaggccgcctacctcaagatc
ctggagagctcccagactttgctcagcgttctcaagagggaagctgggaacctgaccaaggctacagccccagaccagaaaagtagcgg
cggcagggacagctga

Gene symbol: STAU
Entrez Gene ID: 6780
dna/ac: BC050432.1
sequence:

atgtctcaagttcaagtgcaagttcagaacccatctgctgctctctcagggagccaaatactgaacaagaaccagtctcttctctca
cagcctttgatgagtattccttctactactagctctctgccctctgaaaatgcaggtagacccattcaaaactctgctttaccctctgcatctattac
atccaccagtgcagctgcagaaagcataacccctactgtagaactaaatgcactgtgcatgaaacttgga

Gene symbol: STK11
Entrez Gene ID: 6794
dna/ac: BC007981.2
sequence:

atggaggtggtggacccgcagcagctgggcatgttcacggagggcgagctgatgtcggtgggtatggacacgttcatccacc
gcatcgactccaccgaggtcatctaccagccgcgccgcaagcgggccaagctcatcggcaagtacctgatgggggacctgctgggggga
aggctcttacggcaaggtgaaggaggtgctggactcggagacgctgtgcaggagggccgtcaagatcctcaagaagaagaagttgcga
aggatccccaacggggaggccaacgtgaagaaggaaattcaactactgaggaggttacggcacaaaaatgtcatccagctggtggatgt
gttatacaacgaagagaagcagaaaatgtatatggtgatggagtactgcgtgtgtggcatgcaggaaatgctggacagcgtgccggagaa
gcgtttcccagtgtgccaggcccacgggtacttctgtcagctgattgacggcctggagtacctgcatagccagggcattgtgcacaaggac
atcaagccgggggaacctgctgctcaccaccggtggcaccctcaaaatctccgacctgggcgtggccgaggcactgcacccgttcgcggc
ggacgacacctgccggaccagccagggctccccggctttccagccgcccgagattgccaacggcctggacaccttctccggcttcaagg
tggacatctggtcggctggggtcaccctctacaacatcaccacgggtctgtacccccttcgaaggggacaacatctacaagttgtttgagaac
atcgggaaggggagctacgccatcccgggcgactgtggcccccccgctctctgacctgctgaaagggatgcttgagtacgaaccggccaa
gaggttctccatccggcagatccggcagcacagctggttccggaagaaacatcctccggctgaagcaccagtgcccatcccaccgagcc
cagacaccaaggaccggtggcgcagcatgactgtggtgccgtacttggaggacctgcacggcgcgggacgaggacgaggacctcttcga
catcgaggatgacatcatctacactcaggacttcacggtgcccggacaggtcccagaagaggaggccagtcacaatggacagcgccgg
ggcctccccaaggccgtgtgtatgaacggcacagaggcggcgcagctgagcaccaaatccagggcggagggccgggcccccaaccct
gcccgcaaggcctgctccgccagcagcaagatccgccggctgtcggcctgcaagcagcagtga

Gene symbol: TOM1
Entrez Gene ID: 10043

dna/ac: BC046151.1
sequence:

atggactttctcctggggaacccgttcagctctccagtgggacagcgcatcgagaaagccacagatggctccctgcagagcga ggactgggccctcaacatggagatctgcgacatcatcaacgagacggaggaaggtcccaaagatgccctccgagcagtaaagaagaga atcgtggggaataagaacttccacgaggtgatgctggctctcacagtcttagaaacctgtgtcaagaactgcgggcaccgcttccacgtgct ggtggccagccaggacttcgtggagagtgtgctggtgaggaccatcctgcccaagaacaacccacccaccatcgtgcatgacaaagtgct caacctcatccagtcctgggctgacgcgttccgcagctcgcccgatctgacaggtgtggtcaccatctatgaggacctgcggaggaaagg cctggagttcccccatgactgacctggacatgctgtcacccatccacacaccccagaggaccgtgttcaactcagagacacaatcaggacag gattctgtgggcactgactccagccagcaagaggactctggccagcatgctgcccctctgcccgccccgcccatactctccggtgacacg cccatagcaccaaccccggaacagattgggaagctgcgcagtgagctggagatggtgagtgggaacgtgagggtgatgtcggagatgct gacggagctggtgcccacccaggccgagcccgcagacctggagctgctgcaggagctcaaccgcacgtgccgagccatgcagcagcg ggtcctggagctcatccctcagatcgccaatgagcagctgacagaggagctgctcatcgtcaatgacaatctcaacaatgtgttcctgcgcc atgaacggtttgaacggttccgaacaggccagaccaccaaggccccaagtgaggccgagccggcagctgacctgatcgacatgggccc tgacccagcagccaccggcaacctctcatcccagctggcaggaatgaacctgggctccagcagtgtgagagctggcctgcagtctctgga ggcctctggtcgactggaagatgagtttgacatgtttgcgctgacacgggggcagctcactggctgaccaacggaaagaggtaaaatacga agccccccaagcaacagacggcctggctggagccctggacgcccggcagcagagcactggcgcgatcccagtcacccaggcctgcct catggaggacatcgagcagtggctgtccactgacgtgggtaatgatgcggaagagcctaaggggggtcaccagcgaaggtaaatttgaca aattcctggaagaacgggccaaagccgcggaccgattgcccaacctctccagcccctcagctgaggggcccccgggtcccccatctggc ccagcgccccggaagaagacccaggagaaagatgatgacatgctgtttgccttatga

Gene symbol: TXNL2
Entrez Gene ID: 10539
dna/ac: BC005289
sequence:

atggcggcggggggcggctgaggcagctgtagcggccgtggaggaggtcggctcagccgggcagtttgaggagctgctgcg cctcaaagccaagtccctccttgtggtccatttctgggcaccatgggctccacagtgtgcacagatgaacgaagttatggcagagttagctaa agaactccctcaagtttcatttgtgaagttggaagctgaaggtgttcctgaagtatctgaaaaatatgaaattagctctgttcccactttctgtttt caagaattctcagaaaatcgaccgattagatggtgcacatgccccagagttgaccaaaaaagttcagcgacatgcatctagtggctccttcct acccagcgctaatgaacatcttaaagaagatctcaaccttcgcttgaagaaattgactcatgctgcccccctgcatgctgtttatgaaaggaact cctcaagaaccacgctgtggtttcagcaagcagatggtggaaattcttcacaaacataatattcagtttagcagttttgatatcttctcagatgaa gaggttcgacagggactcaaagcctattccagttggcctacctatcctcagctctatgtttctggagagctcataggaggacttgatataattaa ggagctagaagcatctgaagaactagatacaatttgtcccaaagctcccaaattagaggaaaaggctcaaagtgctgacaaataaagcttctg tgatgctctttatgaaaggaaacaaacaggaagcaaaatgtggattcagcaaacaaattctggaaatactaaatagtactggtgttgaatatga aacattcgatatattggaggatgaagaagttcggcaaggattaaaaagcttactcaaattggccaacatacccctcagctgtatgtgaaagggga gctggtggaggattggatattgtgaaggaactgaaagaaatggtgaattgctgcctatactgagaggagaaaattaa

Gene symbol: TXNRD1
Entrez Gene ID: 7296
dna/ac: BC018122.1
sequence:

atgaacggccctgaagatcttcccaagtcctatgactatgaccttatcatcattggaggtggctcaggaggtctggcagctgctaa ggaggcagcccaatatggcaagaaggtgatggtcctggactttgtcactcccaccctcttggaactagatggggtcttggaggaacatgtg tgaatgtggggttgcatacctaaaaaactgatgcatcaagcagctttgttaggacaagccctgcaagactctcgaaattatggatggaaagtcg aggagacagttaagcatgattgggacagaatgatagaagctgtacagaatcacattggctctttgaattggggctaccgagtagctctgcgg gagaaaaaagtcgtctatgagaatgcttatgggcaatttattggtcctcacaggattaaggcaacaaataataaaggcaaagaaaaaatttatt cagcagagagatttctcattgccactggtgaaagaccacgttacttgggcatccctggtgacaaagaatactgcatcagcagtgatgatcttt

ctccttgccttactgcccgggtaagaccctggttgttggagcatcctatgtcgctttggagtgcgctggatttcttgctggtattggtttagacgt
cactgttatggttaggtccattcttcttagaggatttgaccaggacatggccaacaaaattggtgaacacatggaagaacatggcatcaagttt
ataagacagttcgtaccaattaaagttgaacaaattgaagcagggacaccaggccgactcagagtagtagctcagtccaccaatagtgagg
aaatcattgaaggagaatataatacggtgatgctggcaataggaagagatgcttgcacaagaaaaattggcttagaaaccgtagggggtgaa
gataaatgaaaagactggaaaaatacctgtcacagatgaagaacagaccaatgtgccttacatctatgccattggcgatatattggaggataa
ggtggagctcaccccagttgcaatccaggcaggaagattgctggctcagaggctctatgcaggttccactgtcaagtgtgactatgaaaatg
ttccaaccactgtatttactcctttggaatatggtgcttgtggcctttctgaggagaaagctgtggagaagtttggggaagaaaatattgaggttt
accatagttacttttggccattggaatggacgattccgtcaagagataacaacaaatgttatgcaaaaataatctgtaatactaaagacaatgaa
cgtgttgtgggctttcacgtactgggtccaaatgctggagaagttacacaaggctttgcagctgcgctcaaatgtggactgaccaaaaagca
gctggacagcacaattggaatccaccctgtctgtgcagaggtattcacaacattgtctgtgaccaagcgctctggggcaagcatcctccagg
ctggctgctgaggttaa

Gene symbol: ZMAT2
Entrez Gene ID: 153527
dna/ac: BC056668.1
sequence:

atggcgtcgggcagcgggacaaaaaacttggactttcgccgaaagtgggacaaagatgaatatgagaaactcgccgagaaga
ggctcacggaagagagagaaaagaaagatggaaaaccagtgcagcctgtcaagcgagagctttacggcatagggactacaaggtgga
cttggaatccaagcttgggaagacaattgtcattaccaagacaacccctcaatctgagatgggaggatattactgcaatgtctgtgactgtgtg
gtgaaggactccatcaactttctggatcacattaatggaaagaaacatcagagaaacctgggcatgtctatgcgtgtggaacgttccaccctg
gatcaggtgaagaaacgtttgaggtcaacaagaagaagatggaagagaagcagaaggattatgattttgaggaaaggatgaaggagctc
agagaagaggaggaaaaggccaaagcgtacaagaaagagaaacagaaggagaagaaaaggagggctgaggaggacttgacatttga
ggaggacgatgagatggcagctgtgatgggcttctctggctttggttccaccaagaagagttactga

Gene symbol: ZNF38
Entrez Gene ID: 7589
dna/ac: BC047309.1
sequence:

atgatgaccaaggtactaggcatggccccagttctgggccctaggcctccacaggagcaggtggggcctctgatggtaaaagt
cgaggagaaagaagagaaaggcaagtaccttcctagcctggagatgttccgccagcgcttcaggcagtttgggtaccatgatacccctgg
accccgagaggccctgagccaactccgggtgctctgctgtgagtggctgaggcccgagatccacaccaaggagcagatcctggagctac
tggtgctggagcagttcctgaccatcctgccccaggagctccaggcctgggtgcaggagcattgcccggagagcgctgaagaggctgtc
actctcctcgaagatctggagcgggaactggatgagccaggacaccaggtctcaactcctccaaacgaacagaaaccggtgtgggagaa
gatatcctcctcaggaactgcaaaggaatccccgagcagcatgcagccacagcccttggagaccagtcacaaatacgagtcttgggggcc
cctgtacatccaagagtctggtgaggagcaggagttcgctcaagatccaagaaaggtccgagattgcagattgagtacccagcacgagga
atcagcagatgagcagaaaggttctgaagcagaggggctcaaaggggatataatttctgtgattatcgccaataaacctgaggccagcttag
agaggcagtgcgtaaaccttgaaaatgaaaaaggaacaaaaccccctcttcaagaggcaggctccaagaaaggtagagaatcagttccta
ctaaacctaccccaggagagagacgttatatatgtgctgaatgtggcaaagcctttagtaatagctcaaatctcaccaaacacaggagaaca
cacactggggagaaaccttacgtgtgcaccaagtgtgggaaagctttcagccacagctcaaacctcaccctccactacagaacacacttgg
tggaccggccctatgactgtaagtgtggaaaagcttttgggcagagctcagaccttcttaaacatcagagaatgcacacagaagaggcgcc
atatcagtgcaaagattgtggcaaggctttcagcgggaaaggcagcctcattcgtcactatcggatccacactggggagaagccttatcagt
gtaacgaatgtgggaagagcttcagtcagcatgcgggcctcagctcccaccagagactccacaccggagagaagccatataagtgtaag
gagtgtgggaaagccttcaaccacagctccaacttcaataaacaccacagaatccacaccggggaaaagccctactggtgtcatcactgtg
gaaagaccttctgtagcaagtccaatctttccaaacatcagcgagtccacactggagagggagaagcaccgtaa

Claims

1. A method of diagnosing a patient as having or being predisposed to developing lupus, comprising detecting in a sample taken from said patient one or more auto-antibodies that bind an auto-antigen selected from the group consisting of BANK1 (B-cell scaffold protein with ankyrin repeats 1); BPY2IP1 (MAP1S protein); CEBPG (CCAAT/enhancer binding protein (C/EBP), gamma); E1B-AP5 (E1B-55kDa-associated protein 5); FUS (Fusion (involved in t(12;16) in malignant liposarcoma)); HAGH (Hydroxyacylglutathione hydrolase); HMG20B (High-mobility group 20B); HOXB6 (Homeo box B6, transcript variant 2); KRT8 (Keratin 8); LIN28 (Lin-28 homolog (C. elegans)); NDUFV3 (NADH dehydrogenase (ubiquinone) flavoprotein 3, 10kDa); PABPC1 (Poly(A) binding protein, cytoplasmic 1); PHLDA1 (Pleckstrin homology-like domain, family A, member 1); PIAS2 (Msx-interacting-zinc finger, transcript variant alpha); RAB11FIP3 (RAB11 family interacting protein 3 (class II)); RAN (RAN, member RAS oncogene family); RARA (Retinoic acid receptor alpha); RBMS1 (RNA binding motif, single stranded interacting protein 1, transcript variant); RDBP (RD RNA binding protein); RNF12 (Ring finger protein 12, transcript variant 1); RUFY1 (RUN and FYVE domain containing 1); SMN1 (Survival of motor neuron 1, telomeric); SRPK1 (SFRS protein kinase 1); SSNA1 (Sjogren's syndrome nuclear autoantigen 1); STAU (Staufen, RNA binding protein (Drosophila), transcript variant T3); STK11 (Serine/threonine kinase 11 (Peutz-Jeghers syndrome)); TOM1 (Target of myb1 (chicken)); TXNL2 (Thioredoxin-like, clone MGC:12349); TXNRD1 (Thioredoxin reductase 1, transcript variant 5); ZMAT2 (Zinc finger, matrin type 2); or ZNF38 (Zinc finger protein 38),

wherein the presence of said one or more auto-antibodies, or an increase in the concentration of said one or more auto-antibodies, indicates that the patient suffers or is predisposed to develop lupus.

2. The method of claim 1, wherein the presence of auto-antibodies to at least two of the auto-antigens are detected.

3. A method of diagnosing a patient as having or being predisposed to developing lupus, comprising detecting in a sample from said patient auto-antibodies to a panel comprising 5 or more auto-antigens selected from the group consisting of BANK1 (B-cell scaffold protein with ankyrin repeats 1); BPY2IP1 (MAP1S protein); CEBPG (CCAAT/enhancer binding protein

(C/EBP), gamma); E1B-AP5 (E1B-55kDa-associated protein 5); FUS (Fusion (involved in t(12;16) in malignant liposarcoma)); HAGH (Hydroxyacylglutathione hydrolase); HMG20B (High-mobility group 20B); HOXB6 (Homeo box B6, transcript variant 2); IFI16 (Interferon, gamma-inducible protein 16); KRT8 (Keratin 8); LIN28 (Lin-28 homolog (C. elegans)); NDUFV3 (NADH dehydrogenase (ubiquinone) flavoprotein 3, 10kDa); PABPC1 (Poly(A) binding protein, cytoplasmic 1); PHLDA1 (Pleckstrin homology-like domain, family A, member 1); PIAS2 (Msx-interacting-zinc finger, transcript variant alpha); RAB11FIP3 (RAB11 family interacting protein 3 (class II)); RALBP1 (RalA-binding protein 1); RAN (RAN, member RAS oncogene family); RARA (Retinoic acid receptor alpha); RBMS1 (RNA binding motif, single stranded interacting protein 1, transcript variant); RDBP (RD RNA binding protein); RNF12 (Ring finger protein 12, transcript variant 1); RUFY1 (RUN and FYVE domain containing 1); SMN1 (Survival of motor neuron 1, telomeric); SRPK1 (SFRS protein kinase 1); SSNA1 (Sjogren's syndrome nuclear autoantigen 1); STAU (Staufen, RNA binding protein (Drosophila), transcript variant T3); STK11 (Serine/threonine kinase 11 (Peutz-Jeghers syndrome)); TOM1 (Target of myb1 (chicken)); TXNL2 (Thioredoxin-like, clone MGC:12349); TXNRD1 (Thioredoxin reductase 1, transcript variant 5); ZMAT2 (Zinc finger, matrin type 2); or ZNF38 (Zinc finger protein 38), PSME3 (Proteasome (prosome, macropain) activator subunit 3 (PA28 gamma; Ki), transc), and SSA2 (Sjogren syndrome antigen A2 (60kDa, ribonucleoprotein autoantigen SS-A/Ro)), wherein the presence or separately a significant increase in the concentration of said auto-antibodies in the sample to substantially all members of the panel indicates that the patient suffers or is predisposed to develop lupus.

4.   The method of claim 3, wherein the panel comprises 10 or more of said auto-antigens.

5.   The method of any previous claim, wherein the patient is pre-symptomatic.

6.   A method of monitoring the progression or regression of lupus, or monitoring the progression to a flare of the disease or transition from a flare into remission, comprising

(a)   detecting in a first sample from a patient the presence of one or more auto-antibodies that bind an auto-antigen selected from the group consisting of BANK1 (B-cell scaffold protein with ankyrin repeats 1); BPY2IP1 (MAP1S protein); CEBPG (CCAAT/enhancer binding protein (C/EBP), gamma); E1B-AP5 (E1B-55kDa-associated protein 5); FUS (Fusion (involved in t(12;16) in malignant liposarcoma)); HAGH

(Hydroxyacylglutathione hydrolase); HMG20B (High-mobility group 20B); HOXB6 (Homeo box B6, transcript variant 2); KRT8 (Keratin 8); LIN28 (Lin-28 homolog (C. elegans)); NDUFV3 (NADH dehydrogenase (ubiquinone) flavoprotein 3, 10kDa); PABPC1 (Poly(A) binding protein, cytoplasmic 1); PHLDA1 (Pleckstrin homology-like domain, family A, member 1); PIAS2 (Msx-interacting-zinc finger, transcript variant alpha); RAB11FIP3 (RAB11 family interacting protein 3 (class II)); RAN (RAN, member RAS oncogene family); RARA (Retinoic acid receptor alpha); RBMS1 (RNA binding motif, single stranded interacting protein 1, transcript variant); RDBP (RD RNA binding protein); RNF12 (Ring finger protein 12, transcript variant 1); RUFY1 (RUN and FYVE domain containing 1); SMN1 (Survival of motor neuron 1, telomeric); SRPK1 (SFRS protein kinase 1); SSNA1 (Sjogren's syndrome nuclear autoantigen 1); STAU (Staufen, RNA binding protein (Drosophila), transcript variant T3); STK11 (Serine/threonine kinase 11 (Peutz-Jeghers syndrome)); TOM1 (Target of myb1 (chicken)); TXNL2 (Thioredoxin-like, clone MGC:12349); TXNRD1 (Thioredoxin reductase 1, transcript variant 5); ZMAT2 (Zinc finger, matrin type 2); or ZNF38 (Zinc finger protein 38),

(b)　detecting said auto-antibodies in a second sample obtained from the patient at a time later than the first sample;

(c)　comparing the auto-antibodies in said first and second samples;

wherein a change in the auto-antibodies in the first and second sample indicates the progression or regression of lupus, or the progression to a flare, or transition from a flare into remission.

7.　The method of claim 6, wherein a panel comprising 3 or more of said auto-antigens is used for the detection of auto-antibodies.

8.　The method of claim 7, wherein the panel of auto-antigens may also comprise one or more of IFI16 (Interferon, gamma-inducible protein 16); PSME3 (Proteasome (prosome, macropain) activator subunit 3 (PA28 gamma; Ki), transc); RALBP1 (RalA-binding protein 1) and SSA2 (Sjogren syndrome antigen A2 (60kDa, ribonucleoprotein autoantigen SS-A/Ro)).

9.　The method of any of claims 6 to 8, wherein the change is an increase in the number of auto-antigens to which auto-antibodies are detected in the second sample as compared to the first sample, and said change indicates that lupus has become more severe.

10.   The method of any of claims 6 to 8, wherein the change is a decrease in the number of auto-antigens to which auto-antibodies are detected in the second sample as compared to the first sample, and said change indicates that lupus has become less severe.

11.   The method of any of claims 6 to 8, wherein the change is a change in the concentration of one or more of the auto-antibodies in the second sample as compared to the first sample.

12.   The method of claim 11, wherein the change is an increase in the concentration of one or more of the auto-antibodies in the second sample as compared to the first sample, wherein the increase indicates that lupus has become more severe.

13.   The method of claim 11, wherein the change is a decrease in the concentration of one or more of the auto-antibodies in the second sample as compared to the first sample, wherein the decrease indicates that lupus has become less severe.

14.   A method of monitoring the efficacy of a therapeutic agent to lupus, comprising:

    (a)   detecting in a first sample from a patient the presence of one or more auto-antibodies that bind an auto-antigen selected from the group consisting of BANK1 (B-cell scaffold protein with ankyrin repeats 1); BPY2IP1 (MAP1S protein); CEBPG (CCAAT/enhancer binding protein (C/EBP), gamma); E1B-AP5 (E1B-55kDa-associated protein 5); FUS (Fusion (involved in t(12;16) in malignant liposarcoma)); HAGH (Hydroxyacylglutathione hydrolase); HMG20B (High-mobility group 20B); HOXB6 (Homeo box B6, transcript variant 2); KRT8 (Keratin 8); LIN28 (Lin-28 homolog (C. elegans)); NDUFV3 (NADH dehydrogenase (ubiquinone) flavoprotein 3, 10kDa); PABPC1 (Poly(A) binding protein, cytoplasmic 1); PHLDA1 (Pleckstrin homology-like domain, family A, member 1); PIAS2 (Msx-interacting-zinc finger, transcript variant alpha); RAB11FIP3 (RAB11 family interacting protein 3 (class II)); RAN (RAN, member RAS oncogene family); RARA (Retinoic acid receptor alpha); RBMS1 (RNA binding motif, single stranded interacting protein 1, transcript variant); RDBP (RD RNA binding protein); RNF12 (Ring finger protein 12, transcript variant 1); RUFY1 (RUN and FYVE domain containing 1); SMN1 (Survival of motor neuron 1, telomeric); SRPK1 (SFRS protein kinase 1); SSNA1 (Sjogren's syndrome nuclear autoantigen 1); STAU (Staufen, RNA binding protein (Drosophila), transcript variant T3); STK11 (Serine/threonine kinase 11 (Peutz-Jeghers syndrome)); TOM1 (Target of myb1

(chicken)); TXNL2 (Thioredoxin-like, clone MGC:12349); TXNRD1 (Thioredoxin reductase 1, transcript variant 5); ZMAT2 (Zinc finger, matrin type 2); or ZNF38 (Zinc finger protein 38),

(b)   administering the therapeutic agent to the patient;

(c)   detecting said one or more auto-antibodies in a second sample from the patient taken after the administration of the therapeutic agent;

(d)   comparing the one or more auto-antibodies in said first and second samples;

wherein a change in the auto-antibodies in the first and second sample is indicative of the efficacy of the therapeutic agent to lupus.

15.   The method of claim 14, wherein the change in the one or more auto-antibodies is a decrease in the number of auto-antigens to which auto-antibodies are detected, or the concentration of auto-antibodies, in the second sample as compared to the first sample, and wherein the decrease indicates that the therapeutic agent is efficacious.

16.   The method of claim 15, wherein the change in the one or more auto-antibodies is a decrease in the number of auto-antigens to which auto-antibodies are detected in the second sample as compared to the first sample.

17.   The method of claim 15, wherein the change is a decrease in the concentration of the auto-antibody.

18.   The method of claim 17, wherein the decrease in concentration is by at least 20%.

19.   The method of any previous claim, wherein the sample is a body fluid.

20.   The method of claim 19, wherein the body fluid is selected from blood, serum, plasma, saliva, lymphatic fluid, wound secretion, urine, faeces, mucus or cerebrospinal fluid (CSF).

21.   The method of claim 20, wherein the sample is serum or plasma.

22.   A diagnostic kit for use in diagnosing a patient as having or being predisposed to developing lupus, said kit comprising one or more auto-antigens selected from the group

consisting of BANK1 (B-cell scaffold protein with ankyrin repeats 1); BPY2IP1 (MAP1S protein); CEBPG (CCAAT/enhancer binding protein (C/EBP), gamma); E1B-AP5 (E1B-55kDa-associated protein 5); FUS (Fusion (involved in t(12;16) in malignant liposarcoma)); HAGH (Hydroxyacylglutathione hydrolase); HMG20B (High-mobility group 20B); HOXB6 (Homeo box B6, transcript variant 2); KRT8 (Keratin 8); LIN28 (Lin-28 homolog (C. elegans)); NDUFV3 (NADH dehydrogenase (ubiquinone) flavoprotein 3, 10kDa); PABPC1 (Poly(A) binding protein, cytoplasmic 1); PHLDA1 (Pleckstrin homology-like domain, family A, member 1); PIAS2 (Msx-interacting-zinc finger, transcript variant alpha); RAB11FIP3 (RAB11 family interacting protein 3 (class II)); RAN (RAN, member RAS oncogene family); RARA (Retinoic acid receptor alpha); RBMS1 (RNA binding motif, single stranded interacting protein 1, transcript variant); RDBP (RD RNA binding protein); RNF12 (Ring finger protein 12, transcript variant 1); RUFY1 (RUN and FYVE domain containing 1); SMN1 (Survival of motor neuron 1, telomeric); SRPK1 (SFRS protein kinase 1); SSNA1 (Sjogren's syndrome nuclear autoantigen 1); STAU (Staufen, RNA binding protein (Drosophila), transcript variant T3); STK11 (Serine/threonine kinase 11 (Peutz-Jeghers syndrome)); TOM1 (Target of myb1 (chicken)); TXNL2 (Thioredoxin-like, clone MGC:12349); TXNRD1 (Thioredoxin reductase 1, transcript variant 5); ZMAT2 (Zinc finger, matrin type 2); or ZNF38 (Zinc finger protein 38), which one or more auto-antigens are immobilised on a substrate.

23. A diagnostic kit for use in diagnosing a patient as having or being predisposed to developing lupus, said kit comprising a panel comprising 5 or more auto-antigens selected from the group consisting of BANK1 (B-cell scaffold protein with ankyrin repeats 1); BPY2IP1 (MAP1S protein); CEBPG (CCAAT/enhancer binding protein (C/EBP), gamma); E1B-AP5 (E1B-55kDa-associated protein 5); FUS (Fusion (involved in t(12;16) in malignant liposarcoma)); HAGH (Hydroxyacylglutathione hydrolase); HMG20B (High-mobility group 20B); HOXB6 (Homeo box B6, transcript variant 2); IFI16 (Interferon, gamma-inducible protein 16); KRT8 (Keratin 8); LIN28 (Lin-28 homolog (C. elegans)); NDUFV3 (NADH dehydrogenase (ubiquinone) flavoprotein 3, 10kDa); PABPC1 (Poly(A) binding protein, cytoplasmic 1); PHLDA1 (Pleckstrin homology-like domain, family A, member 1); PIAS2 (Msx-interacting-zinc finger, transcript variant alpha); RAB11FIP3 (RAB11 family interacting protein 3 (class II)); RALBP1 (RalA-binding protein 1); RAN (RAN, member RAS oncogene family); RARA (Retinoic acid receptor alpha); RBMS1 (RNA binding motif, single stranded interacting protein 1, transcript variant); RDBP (RD RNA binding protein); RNF12 (Ring

finger protein 12, transcript variant 1); RUFY1 (RUN and FYVE domain containing 1); SMN1 (Survival of motor neuron 1, telomeric); SRPK1 (SFRS protein kinase 1); SSNA1 (Sjogren's syndrome nuclear autoantigen 1); STAU (Staufen, RNA binding protein (Drosophila), transcript variant T3); STK11 (Serine/threonine kinase 11 (Peutz-Jeghers syndrome)); TOM1 (Target of myb1 (chicken)); TXNL2 (Thioredoxin-like, clone MGC:12349); TXNRD1 (Thioredoxin reductase 1, transcript variant 5); ZMAT2 (Zinc finger, matrin type 2); or ZNF38 (Zinc finger protein 38), PSME3 (Proteasome (prosome, macropain) activator subunit 3 (PA28 gamma; Ki), transc) and SSA2 (Sjogren syndrome antigen A2 (60kDa, ribonucleoprotein autoantigen SS-A/Ro)), wherein the auto-antigens are immobilised on a substrate.

24.   The kit of claim 22 or 23, wherein said substrate comprises a strip, a slide, beads, SPR or planar waveguide chip or a well of a microtitre plate.

25.   The kit of any of claims 22 to 24, wherein the panel of auto-antigens is carried on said substrate in the form of an array.

26.   The kit of claim 25, wherein the auto-antigens constitute at least 50% of the total number of proteins present on the array.

27.   The kit of claim 25, wherein the array includes one or more other proteins or polypeptides, and means for automatically analysing the results of tests performed using said array, said analysing means including computer-readable instructions in which results from said one or more other proteins or polypeptides are blinded during the analysis.

28.   The kit of claim 25, wherein the array further comprises one or more control proteins or polypeptides.

29.   The kit of claim 25 or claim 28, wherein the array comprises one or more replicates of the auto-antigens or controls.

30.   The kit of any of claims 22 to 29, further comprising an anti-human immunoglobulin antibody and means for detecting said anti-human immunoglobulin antibody.

31.   The kit of claim 30, wherein the means for detecting comprise a colorimetric or a fluorescent assay or a label free detection method such as SPR or planar waveguide technology.

32.   Use of an auto-antigen selected from the group consisting of BANK1 (B-cell scaffold protein with ankyrin repeats 1); BPY2IP1 (MAP1S protein); CEBPG (CCAAT/enhancer binding protein (C/EBP), gamma); E1B-AP5 (E1B-55kDa-associated protein 5); FUS (Fusion (involved in t(12;16) in malignant liposarcoma)); HAGH (Hydroxyacylglutathione hydrolase); HMG20B (High-mobility group 20B); HOXB6 (Homeo box B6, transcript variant 2); IFI16

(Interferon, gamma-inducible protein 16); KRT8 (Keratin 8); LIN28 (Lin-28 homolog (C. elegans)); NDUFV3 (NADH dehydrogenase (ubiquinone) flavoprotein 3, 10kDa); PABPC1 (Poly(A) binding protein, cytoplasmic 1); PHLDA1 (Pleckstrin homology-like domain, family A, member 1); PIAS2 (Msx-interacting-zinc finger, transcript variant alpha); RAB11FIP3 (RAB11 family interacting protein 3 (class II)); RALBP1 (RalA-binding protein 1); RAN (RAN, member RAS oncogene family); RARA (Retinoic acid receptor alpha); RBMS1 (RNA binding motif, single stranded interacting protein 1, transcript variant); RDBP (RD RNA binding protein); RNF12 (Ring finger protein 12, transcript variant 1); RUFY1 (RUN and FYVE domain containing 1); SMN1 (Survival of motor neuron 1, telomeric); SRPK1 (SFRS protein kinase 1); SSNA1 (Sjogren's syndrome nuclear autoantigen 1); STAU (Staufen, RNA binding protein (Drosophila), transcript variant T3); STK11 (Serine/threonine kinase 11 (Peutz-Jeghers syndrome)); TOM1 (Target of myb1 (chicken)); TXNL2 (Thioredoxin-like, clone MGC:12349); TXNRD1 (Thioredoxin reductase 1, transcript variant 5); ZMAT2 (Zinc finger, matrin type 2); or ZNF38 (Zinc finger protein 38), PSME3 (Proteasome (prosome, macropain) activator subunit 3 (PA28 gamma; Ki), transc) and SSA2 (Sjogren syndrome antigen A2 (60kDa, ribonucleoprotein autoantigen SS-A/Ro))

in the manufacture of a vaccine for the treatment or prevention of lupus.

33. The use of claim 32, wherein the vaccine comprises two or more of the antigens.

34. The use of claims 32 or claim 33, wherein the vaccine additionally comprises an adjuvant.
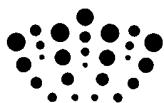
Application No:	GB0810709.6	**Examiner:**	Dr Rowena Dinham
Claims searched:	1-34 (all in part)	**Date of search:**	7 October 2009

Patents Act 1977: Search Report under Section 17

Documents considered to be relevant:

Category	Relevant to claims	Identity of document and passage or figure of particular relevance
A	-	Nature Genet; Vol 40, pp 211-216 (2008). Kozyrev et al. "Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus" See entire document
A	-	Arth Rheumat; Vol 58, pp S221- S222 (2008). Soejima et al. "Association of SLE with STAT4, BANK1, BLK and PXK: a replication study in a Japanese population" See Abstract
A	-	Arth Rheumat; Vol 58, pp S220 (2008). Guthridge et al. "Confirmation of the genetic association of BANK1 with systemic lupus erythematosus" See Abstract
A,P	-	Genes Immunity; Vol 10, pp 373-379 (2009). Moser et al. "Recent insights into the genetic basis of systemic lupus erythematosus" See entire document
A,P	-	WO 2009/068481 A (MERCK SERONA SA) See entire document
A	-	EMBO J; Vol 21, pp 83-92 (2002). Yokoyama et al. "BANK regulates BCR-induced calcium mobilization by promoting tyrosine phosphorylation of IP3 receptor" See entire document

Categories:

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.

Field of Search:

Search of GB, EP, WO & US patent documents classified in the following areas of the UKC^X :

Worldwide search of patent documents classified in the following areas of the IPC

A61K; G01N

The following online and other databases have been used in the preparation of this search report
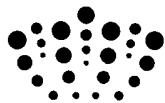
WPI, EPODOC, MEDLINE, BIOSIS

International Classification:

Subclass	Subgroup	Valid From
G01N	0033/564	01/01/2006
A61K	0039/00	01/01/2006
A61P	0037/00	01/01/2006