

BIOMARKERS FOR PREDICTING IMMUNOGENICITY AND THERAPEUTIC RESPONSES TO ADALIMUMAB IN RHEUMATOID ARTHRITIS PATIENTS

5 **Field of Invention**

The invention relates to the detection of immunological biomarkers, particularly autoantibodies, to predict immunogenicity and therapeutic responses to adalimumab in patients with Rheumatoid Arthritis (RA).

10 **Background**

Rheumatoid arthritis (RA), a chronic inflammatory articular disease, is characterized by persistent synovitis, cartilage degradation, and bone erosions [1], and tumor necrosis factor (TNF)- α is a crucial inflammatory mediator in RA-related synovitis and joint damage [2]. The importance of the role of TNF- α in RA pathogenesis is supported by the effectiveness of biologics targeting this cytokine [2-4], although the efficacy diminishes in some patients over time (secondary failure) [5]. Accumulating evidence indicates that the presence of anti-drug antibodies (ADAb) in certain patients may be associated with low or undetectable drug levels and ensuing reduction of therapeutic responsiveness to TNF- α inhibitors [6-10]. Such ADAb responses reflect the differential immunogenicity of the given biologic drug triggered in individual patients, which results in some patients developing a neutralising antibody response against the biologic drug and others not. In the face of such uncertainty about whether individual RA patients will show therapeutic responsiveness to TNF- α inhibitors or not [11], physicians hoping to optimize personalized and precision therapy are thus eager to find biomarkers which can predict the emergence of ADAb and the effectiveness of anti-TNF- α biologics.

Proteomics research has been increasingly applied to the identification of novel biomarkers that might be useful for monitoring therapeutic response in RA patients on specific treatments [12-14]. However, there is currently limited knowledge about circulating biomarkers that are able to predict the development of ADAb in RA patients receiving anti-TNF- α therapy.

Autoantibody biomarkers as described herein are autoantibodies to antigens, autoantibodies being antibodies which are produced by an individual which are directed against one or more of the individual's own proteins ('self' antigens).

- 5 The aim of the present invention therefore is to provide a novel panel of autoantibody biomarkers that are able to predict immunogenicity of adalimumab and therapeutic responses to adalimumab in individual RA patients, prior to treatment with adalimumab, a widely used TNF- α inhibitor marketed under the brand name HUMIRA® and commonly used for the treatment of autoimmune diseases, such as RA, Crohn's Disease
10 and Psoriasis.

Summary of Invention

In one aspect of the invention, there is provided a method for predicting a response to adalimumab from a sample extracted from a rheumatoid arthritis patient prior to treating
15 the patient with adalimumab, said response being classified as a good response corresponding to anti-drug antibody negative or a poor response corresponding to anti-drug antibody positive, comprising the steps of:

- (i) testing the sample for the presence of autoantibody biomarkers; and
- (ii) determining whether the patient will develop a good response or a bad
20 response to treatment with adalimumab, based on the detection of said autoantibody biomarkers;

characterised in that the autoantibody biomarkers are autoantibodies to antigens comprising SSB, TROVE2 and ZHX2, wherein ZHX2 is associated with the good response, and SSB and TROVE2 are associated with the poor response.

25

Advantageously the autoantibody biomarkers can be used to predict immunogenicity of adalimumab and therapeutic responses to adalimumab in individual rheumatoid arthritis (RA) patients at baseline (i.e. prior to treating the patient with adalimumab).

- 30 In one embodiment the sample is tested using a panel of antigens that correspond to the autoantibody biomarkers. Typically the antigens are biotinylated proteins. Advantageously the biotinylation ensures that the antigens are folded in their correct form to ensure accuracy of detection by the autoantibody biomarkers.

In one embodiment the antigens may include one or more additional antigens from the group comprising of PPARD, SPANXN2, HNRNPA2B, TRIB2, CEP55, SH3GL1, FN3K, PANK3, HPCAL1, THRA, AIFM1, ODC1, RPS6KA4, EEF1D, KLF10, EPHA2,
 5 PRKAR1A and EAPP.

It should be noted that not all human antigens generate an autoantibody response and it is not possible to predict *a priori* which human antigens will do so in a given patient cohort – of the 1622 antigens tested, only autoantibodies against the 21 antigens described above
 10 are suitable as biomarkers in predicting immunogenicity of adalimumab and therapeutic responses to adalimumab in RA patient at baseline.

In one embodiment each biotinylated protein is formed from a Biotin Carboxyl Carrier Protein (BCCP) folding marker which is fused in-frame with the protein.
 15

In one embodiment the biotinylated proteins are bound to a streptavidin-coated substrate.

Advantageously full-length proteins are expressed as fusions to the BCCP folding marker which itself becomes biotinylated *in vivo* when the fusion partner is correctly folded. By
 20 comparison misfolded fusion partners cause the BCCP to remain in the ‘apo’ (i.e. non-biotinylated) form such that it cannot attach to a streptavidin substrate. Thus only correctly folded fusion proteins become attached to the streptavidin substrate via the biotin moiety appended to the BCCP tag.

25 In one embodiment the substrate comprises a glass slide, biochip, strip, slide, bead, microtitre plate well, surface plasmon resonance support, microfluidic device, thin film polymer base layer, hydrogel-forming polymer base layer, or any other device or technology suitable for detection of antibody-antigen binding.

30 In one embodiment the substrate is exposed to a sample extracted from a person, such that autoantibody biomarkers from the sample may bind to the antigens.

Typically the sample comprises any or any combination of exosomes, blood, serum, plasma, urine, saliva, amniotic fluid, cerebrospinal fluid, breast milk, semen or bile.

Typically the sample is collected at baseline prior to administration of the first dose of
5 adalimumab.

In one embodiment following exposure to the sample, the substrate is exposed to a fluorescently-tagged secondary antibody to allow the amount of any autoantibodies from the sample bound to the antigens on the panel to be determined. Typically the secondary
10 antibody is anti-human IgG, but it will be appreciated that other secondary antibodies could be used, such as anti-IgM, anti-IgG1, anti-IgG2, anti-IgG3, anti-IgG4 or anti-IgA.

In one embodiment the patient's response to treatment with adalimumab (i.e. the immunogenic and/or therapeutic response outcome to adalimumab in RA patient at
15 baseline) corresponds to the relative or absolute amount of autoantibodies from the baseline sample specifically binding to the antigens.

In one embodiment the method is performed *in vitro*.

20 In a further aspect of the invention, there is provided a method for manufacturing a kit for predicting a response to adalimumab from a sample extracted from a rheumatoid arthritis patient prior to treating the patient with adalimumab, comprising the steps of:

for each antigen in a panel, cloning a biotin carboxyl carrier protein folding marker in-frame with a gene encoding the antigen and expressing the resulting
25 biotinylated antigen;

binding the biotinylated antigens to addressable locations on one or more streptavidin-coated substrates, thereby forming an antigen array;

such that the amount of autoantibodies from the sample binding to the antigens on the panel can be determined by exposing the substrate to the sample and measuring the
30 immunogenicity and response;

characterised in that the antigens comprise SSB, TROVE2 and ZHX2.

In one embodiment the antigens further comprise one or more of PPARD, SPANXN2, HNRNPA2B, TRIB2, CEP55, SH3GL1, FN3K, PANK3, HPCAL1, THRA, AIFM1, ODC1, RPS6KA4, EEF1D, KLF10, EPHA2, PRKAR1A and EAPP.

- 5 In a further aspect of the invention there is provided a method for predicting immunogenicity of adalimumab and therapeutic responses to adalimumab in RA patients at baseline by exposing a composition comprising a panel of antigens as herein described to a sample extracted from a person, and determining the level of autoantibodies from the sample binding to the antigens.

10

In a yet further aspect of the invention there is provided a method for predicting immunogenicity of adalimumab and therapeutic responses to adalimumab in RA patients at baseline by exposing a composition comprising a panel of antigens as herein described to a sample extracted from a person *in vitro*, and determining the level of autoantibodies from the sample binding to the antigens.

15

In further aspect of the invention, there is provided a composition comprising a panel of antigens for predicting an immunogenic and/or therapeutic response to adalimumab in a rheumatoid arthritis patient who has not previously been treated with adalimumab, characterised in that the antigens comprise SSB, TROVE2 and ZHX2.

20

In one embodiment the antigens further comprise one or more of PPARD, SPANXN2, HNRNPA2B, TRIB2, CEP55, SH3GL1, FN3K, PANK3, HPCAL1, THRA, AIFM1, ODC1, RPS6KA4, EEF1D, KLF10, EPHA2, PRKAR1A and EAPP.

25

In one embodiment the antigens are biotinylated proteins

30

In one embodiment the amount of one or more autoantibody biomarkers binding *in vitro* to the antigens in a sample from a patient can be measured to determine the immunogenicity and therapeutic response outcome to adalimumab in an RA patient at baseline.

In yet further aspect of the invention, there is provided a composition comprising a panel of autoantibody biomarkers for predicting an immunogenic and/or therapeutic response to adalimumab in a rheumatoid arthritis patient who has not previously been treated with adalimumab, wherein the level of the autoantibody biomarkers are measured in a sample
 5 collected from the patient;

characterised in that the autoantibody biomarkers are specific to antigens comprising SSB, TROVE2 and ZHX2.

Brief Description of Drawings

10 It will be convenient to further describe the present invention with respect to the accompanying drawings that illustrate possible arrangements of the invention. Other arrangements of the invention are possible, and consequently the particularity of the accompanying drawings is not to be understood as superseding the generality of the preceding description of the invention.

15

Figure 1 illustrates the structure of the *E. coli* Biotin Carboxyl Carrier Protein domain.

Figure 2 illustrates the pPRO9 plasmid used as a vector.

20 Figure 3 illustrates the distribution of the normalised RFU (i.e. autoantibody responses) for all 21 biomarkers between ADAAb-positive (Group A) and ADAAb-negative (Group B) RA patients collected at baseline.

Figure 4 illustrates the discriminatory performance represented as a receiver operating
 25 curve (ROC) with an area under the curve (AUC) of 0.835 for all 21 biomarkers (SSB, TROVE2, ZHX2, PPARD, SPANXN2, HNRNPA2B, TRIB2, CEP55, SH3GL1, FN3K, PANK3, HPCAL1, THRA, AIFM1, ODC1, RPS6KA4, EEF1D, KLF10, EPHA2, PRKAR1A and EAPP) between ADAAb-positive (Group A) and ADAAb-negative (Group B) RA patients collected at baseline.

30

Figure 5 illustrates the variable importance measure for each of the 21 biomarkers identified in the study.

Figure 6 illustrates the discriminatory performance represented as a receiver operating curve (ROC) with an area under the curve (AUC) of 0.761 for the core biomarkers (SSB, TROVE2 and ZHX2) between ADAb-positive (Group A) and ADAb-negative (Group B) RA patients collected at baseline.

5

Detailed Description

The invention utilises the Biotin Carboxyl Carrier Protein (BCCP) folding marker which is cloned in-frame with the gene encoding the protein of interest, as described above and in EP1470229. The structure of the *E. coli* BCCP domain is illustrated in Figure 1,
10 wherein residues 77-156 are drawn (coordinate file 1bdo) showing the N- and C- termini and the single biotin moiety that is attached to lysine-122 *in vivo* by biotin ligase.

BCCP acts not only as a protein folding marker but also as a protein solubility enhancer. BCCP can be fused to either the N- or C-terminal of a protein of interest. Full-length
15 proteins are expressed as fusions to the BCCP folding marker which becomes biotinylated *in vivo*, but only when the protein is correctly folded. Conversely, misfolded proteins drive the misfolding of BCCP such that it is unable to become biotinylated by host biotin ligases. Hence, misfolded proteins are unable to specifically attach to a streptavidin-coated solid support. Therefore only correctly folded proteins become attached to a solid
20 support via the BCCP tag.

The surface chemistry of the support is designed carefully and may use a three-dimensional thin film hydrogel layer (polyethylene glycol; PEG), which retains protein spot morphologies and ensures consistent spot sizes across the array. The PEG layer
25 inhibits non-specific macromolecule absorption, therefore reducing the high background observed using other platforms. The solid support used to immobilize the selected biomarkers thus provides excellent signal-to-noise ratios and low limits of detection (translating in to improved sensitivity). In addition the PEG hydrogel layer also aids preservation of the folded structure and functionality of arrayed proteins and protein
30 complexes post-immobilisation.

Retention of the correct folded structure of immobilised antigens during antibody binding assays ('immuno-assays') is particularly advantageous because human antibodies are

known in general to specifically recognise and bind to discontinuous, solvent-accessible epitopes on protein surfaces, yet are also known to bind non-specifically to exposed hydrophobic surfaces on unfolded proteins. Thus serological assays carried out on arrays of unfolded proteins typically give rise to many false positive results due to such non-specific binding events (which have no biological relevance), whilst at the same time also giving rise to many false negative results due to the absence of biologically-relevant discontinuous epitopes. By contrast, serological assays carried out on arrays of folded antigens result in detection of biologically meaningful antibody-antigen interactions that are not obscured by high rates of non-specific binding.

As biotinylated proteins bound to a streptavidin-coated surface show negligible dissociation, this interaction therefore provides a superior means for tethering proteins to a planar surface in a controlled orientation and is thus ideal for applications such as protein arrays, SPR and bead-based assays. The use of a compact, folded, biotinylated, 80 residue domain BCCP affords two significant advantages over for example the AviTag and intein-based tag. First, the BCCP domain is cross-recognised by eukaryotic biotin ligases enabling it to be biotinylated efficiently in yeast, insect, and mammalian cells without the need to co-express the *E. coli* biotin ligase. Second, the N- and C-termini of BCCP are physically separated from the site of biotinylation by 50Å (as shown in Figure 1), so the BCCP domain can be thought of as a stalk which presents the recombinant proteins away from the solid support surface, thus minimising any deleterious effects due to immobilisation.

The addition of BCCP permits the monitoring of fusion protein folding by measuring the extent of *in vivo* biotinylation. This can be measured by standard blotting procedures, using SDS-PAGE or *in situ* colony lysis and transfer of samples to a membrane, followed by detection of biotinylated proteins using a streptavidin conjugate such as streptavidin-horseradish peroxidase. Additionally, the fact that the BCCP domain is biotinylated *in vivo* is particularly useful when multiplexing protein purification for fabrication of protein arrays since the proteins can be simultaneously purified from cellular lysates and immobilised in a single step via the high affinity and specificity exhibited by a streptavidin surface.

Example 1

Materials and Methods

Gene synthesis and cloning. The pPRO9 plasmid (see Figure 2 below) was constructed
 5 by standard techniques and consists of genetic elements encoding a c-myc tag and a
 BCCP protein domain, preceded by a multiple-cloning site. Synthetic genes encoding
 individual human antigens were assembled from synthetic oligonucleotides and were
 cloned into pPRO9 using *SpeI* and *NcoI* cloning sites such that each resultant clonal
 ‘transfer vector’ encoded an in-frame fusion protein comprising a specific human antigen
 10 fused to the BCCP tag. The plasmid DNA was purified from transformed bacteria and
 verified by DNA sequencing. The required sequence congruence within the synthetic
 gene region was 100%.

Recombinant baculovirus was generated via co-transfection of Sf9 cells (a clonal isolate
 15 derived from the parental *Spodoptera frugiperda* cell line IPLB-Sf-21-AE) with a
 replication-deficient bacmid vector carrying the viral polyhedrin promoter and a transfer
 vector carrying a specific coding sequence for a specific antigen. Homologous
 recombination between the transfer vector and the bacmid within Sf9 cells resulted in
 formation of a replication competent baculoviral vector encoding the specific antigen
 20 fused to the BCCP tag. Successful homologous recombination between the transfer
 vector and the bacmid within Sf9 cells caused the transfected cells to show signs of viral
 cytopathic effect (CPE) within few days of culture incubation. The most common CPE
 observed was the significantly enlargement of average cell size, a consequences of viral
 progeny propagation. These baculoviruses known as P0 were then released into the
 25 culture medium, and viral amplification were done to generate a higher titre of P1 viruses.

Protein Expression. Expression of recombinant antigens was carried out in 24 well
 blocks using 3ml cultures containing 6×10^6 Sf9 cells per well. High titre, low passage,
 viral stocks of recombinant baculovirus ($>10^7$ pfu/ml) were used to infect Sf9 insect cells.
 30 The infected cells were then cultured for 72 hours to allow them to produce the
 recombinant protein of interest. The cells were washed with PBS, resuspended in buffer,
 and were frozen in aliquots at -80°C ready for lysis as required. Depending on the transfer
 vector construct and the nature of the antigen itself, the resultant recombinant protein

lysate can be recovered either from the cultured cell or the culture medium. Expression of recombinant proteins was confirmed by SDS-PAGE as well as by Western blot using streptavidin-HRP-based detection. In total, 1622 human antigens were cloned and expressed using this methodology.

5

Array fabrication. Hydrogel coated, streptavidin-derivatised slides were custom manufactured by Schott and used as substrates on to which the biotinylated proteins were then printed. A total of 9 nanoliters of crude protein lysate was printed on a HS slide in quadruplicate using non-contact piezo printing technology. Print buffer that have a pH
10 between 7.0 and 7.5 were used. The slides were dried by centrifugation (200 x g for 5 min) before starting the washing and blocking. The printed arrays were blocked with solutions containing BSA or casein (concentration: 0.1 mg/ml) in a phosphate buffer. The pH was adjusted to be between 7.0 and 7.5 and cold solutions were used (4 °C - 20 °C). Slides were not allowed to dry between washes, and were protected from light. In total,
15 each resultant ‘Immunome array’ comprised 1622 antigens, each printed in quadruplicate.

Experimental Procedure.

1. Study cohort

The study cohort comprised of a total of 62 plasma and serum samples collected from RA
20 patients at baseline (i.e. prior to treatment administration);

- i. Run 1: 6 ADAAb-positive (“poor response”) and 6 ADAAb-negative (“good response”)
- ii. Run 2: 24 ADAAb-positive (“poor response”) and 26 ADAAb-negative (“good response”)

25

Patients were administered with *adalimumab* at a dose of 40mg every other week. The immunogenicity and therapeutic response to *adalimumab* were evaluated at week 24, the latter by using the EULAR response criteria [15]. EULAR responders were defined as RA patients with good and moderate (“good response”) or poor (“poor response”)
30 EULAR therapeutic responses.

2. Sample Collection and Storage

Peripheral blood samples were collected immediately before the first adalimumab administration (the baseline sample) and also at week 24. After centrifugation at 1000 g for 10 min within 15 min of withdrawal, serum and plasma samples were stored at -70°C.

5 3. Sample Preparation and Dilution

For each run, samples were placed in a shaking incubator set at 20°C to allow thawing for 30 minutes. When completely thawed, each sample was vortexed vigorously three times and debris was pelleted by centrifugation for 3 minutes at 13,000 rpm. 11.25 µL of the sample was pipetted into 4.5 mL of Serum Assay Buffer (SAB) containing 0.1% v/v Triton, 0.1% w/v BSA in PBS (20°C) and vortexed to mix three times. The tube was tilted during aspiration to ensure that the sera was sampled from below the lipid layer at the top but did not touch the bottom of the tube in case of presence of any sediment. Batch records were marked accordingly to ensure that the correct samples were added to the correct tubes. Samples were then randomised prior to assay.

15

4. Biomarker Assay

Each Immunome array was removed from the storage buffer using forceps, placed in the slide box and rack containing 200 mL cold SAB and shaken on an orbital shaker at 50 rpm, for 5 minutes. After washing, each slide was scanned using a barcode scanner and then placed array side up in an individual slide hybridization chamber containing an individual diluted sera (Step 3 above). All slides were and incubated on a horizontal shaker at 50 rpm for 2 hours at 20°C.

20

5. Array Washing After Serum Binding

Each Immunome array slide was rinsed twice in individual “Pap jars” with 30 mL SAB, followed by 200 mL of SAB buffer in the slide staining box for 20 minutes on the shaker at 50 rpm at room temperature. All slides were transferred sequentially and in the same orientation.

25

30 6. Incubation with Cy3-anti-human IgG

Binding of autoantibodies to the arrayed antigens on the arrays was detected by incubation with Cy3-rabbit anti-human IgG (Dako Cytomation) labelled according to the manufacturer's recommended protocols (GE Healthcare). Arrays were immersed in

hybridization solution containing a mixture of Cy3- rabbit anti-human IgG solution diluted 1:1000 in SAB buffer for 2 hours at 50 rpm in 20°C.

7. Washing After Incubation with Cy3-anti-human IgG

- 5 After incubation, the slide was dipped in 200 mL of SAB buffer, 3 times for 5 minutes at 50 rpm at room temperature. Excess buffer was removed by immersing the slide in 200 mL of pure water for a few minutes. Slides were then dried for 2 min by centrifugation at 240g at room temperature. Slides were then stored at room temperature until scanning. Fluorescent hybridization signals were measured with excitation at 550nm and emission
10 at 570nm using a microarray laser scanner (Agilent) at 10µm resolution.

Bioinformatic analysis.

1. Image Analysis: Raw Data Extraction

- 15 The aim of an image analysis is to evaluate the amount of autoantibody present in the serum sample by measuring the median intensities of all the pixels within each probed spot. A raw .tiff format image file is generated for each slide, i.e. each sample. Automatic extraction and quantification of each spot on the array are performed using the GenePix Pro 7 software (*Molecular Devices*) which outputs the statistics for each probed spot on
20 the array. This includes the mean and median of the pixel intensities within a spot as well as in its surrounding local background area. A GAL (GenePix Array List) file for the array is generated to enable image analysis. This file contains the information of all probed spots and their positions on the array. Following data extraction, a GenePix Results (.GPR) file is generated for each slide which contains the information for each
25 spot: Protein ID, protein name, foreground intensities, background intensities etc. In the data sheet generated from the experiment, both foreground and background intensities of each spot are represented in relative fluorescence units (RFUs).

2. Data Handling and Pre-processing

- 30 For each slide, antigens and control probes are spotted in quadruplicate on each array. The following steps were performed to verify the quality of the antigen array data before proceeding with data analysis:

Step 1:

Calculate net intensities for each spot by subtracting background signal intensities from the foreground signal intensities of each spot. For each spot, the background signal intensity was calculated using a circular region with three times the diameter of the spot, centered on the spot.

Step 2:

Remove replica spots with net intensity ≤ 0 .

Step 3:

Zero net intensities if only 1 replica spot remaining.

Step 4:

Calculate the coefficient of variant (CV%) for the replica spots on each array.

$$CV\% = \frac{S. D.}{Mean} \times 100\% \quad \text{Equation 1}$$

Flag any replica spots with only 2 or less replica/s remaining and $CV\% > 20\%$ as “High CV”. The mean net intensity of such replica spots (i.e. antigens) is excluded from downstream analysis.

For antigens/controls with a $CV\% > 20\%$ and with 3 or more replica spots remaining, the replica spots which result in this high CV% value were filtered out. This was done by calculating the standard deviation between the median value of the net intensities and individual net intensities for each set of replica spots. The spot with the highest standard deviation was removed. CV% values were re-calculated and the process repeated.

Step 5:

Calculate the mean of the net intensities for the remaining replica spots.

Step 6:

Inspect signal intensities of two positive controls: IgG and Cy3-BSA.

Step 7:

Carry out a composite normalisation [16] using both quantile-based and total intensity-based modules for each dataset . This method assumes that different samples share a common underlying distribution of their control probes while taking into account the potential existence of flagged spots within them. The Immunome array uses Cy3-labelled biotinylated BSA (Cy3-BSA) replicates as the positive control spots across slides. Hence it is considered as a ‘housekeeping’ probe for normalisation of signal intensities for any given study.

- 10 The quantile module adopts the algorithm described by Bolstad *et al.*, 2003 [17]. This reorganisation enables the detection and handling of outliers or flagged spots in any of the Cy3BSA control probes. A total intensity-based module was then implemented to obtain a scaling factor for each sample. This method assumes that post-normalisation, the positive controls should have a common total intensity value across all samples. This composite method aims to normalise the protein array data from variations in their measurements whilst preserving the targeted biological activity across samples. The steps are as follows:

Quantile-Based Normalisation of all cy3BSA across all samples

- 20 ($i = \text{spot number and } j = \text{sample number}$)

1. Load all Cy3-BSA across all samples, j , into an $i \times j$ matrix X
2. Sort spot intensities in each column j of X to get X_{sort}
3. Take the mean across each row i of X_{sort} to get $\langle X_i \rangle$

- 25 Intensity-Based Normalisation

1. Calculate sum of the mean across each row i , $\sum \langle X_i \rangle$
2. For each sample, k , calculate the sum of all Cy3-BSA controls, $\sum X_k$
3. For each sample, k ,

30
$$\text{Scaling factor (k)} = \frac{\sum \langle X_i \rangle}{\sum X_k}$$

Equation 2

3. Data Analysis

Batch normalisation: The composite normalised data sets from the assays in the two runs were merged using a ComBat normalisation method [18]. For each protein, this method inputs the net intensity values across all the samples from the 2 data sets and adjusts for any possible batch effects between the two data sets using a parametric empirical Bayes frameworks.

Biomarker Panel Selection: A pipeline was developed which utilises a combination of feature selection and machine learning methodologies to determine the optimal combination of antigens eliciting autoantibody responses from the list of 1622 antigens which are able to provide the best stratification between ADAb-positive and ADAb-negative patients [19]. For feature selection, univariate statistical tests, random forest importance and mutual information metrics were the filter methods used to rank biomarkers.

Biomarker panels were generated by additively selecting the top-ranking biomarkers as inputs to machine learning models up to a total of top 160 biomarkers (top 10% of biomarkers). Any further addition of number of biomarkers did not lead to significant improvements of model performance and would lead to further increase of computational time. To estimate the biomarker panel performance, ROC, sensitivity and specificity was evaluated and the biomarker panel with the best sensitivity and specificity was deemed as the optimal panel to stratify ADAb status. For this analysis, machine learning models were built using Random Forests [20], under default settings with leave-one out cross validation (LOOCV). All analyses were performed using packages available in R. Feature selection was performed using ranger [21] package and all machine learning models were performed using the caret [22] package.

Table 1 shows top 4 best performing biomarker panel from the machine learning models using leave-one out cross validation. The lowest number of antigens with the highest sensitivity and specificity was deemed to be the top biomarker panel. This panel comprises SSB, TROVE2, ZHX2, PPARD, SPANXN2, HNRNPA2B, TRIB2, CEP55, SH3GL1, FN3K, PANK3, HPCAL1, THRA, AIFM1, ODC1, RPS6KA4, EEF1D, KLF10, EPHA2, PRKAR1A and EAPP. Figure 3 shows the distribution of the normalised net intensity (i.e. autoantibody responses) for each of these 21 individual biomarkers in

ADAb-positive (Group A) and ADAb-negative (Group B) RA patients at baseline. Figure 4 shows the discriminatory performance of the combined panel of 21 autoantibody biomarkers represented as a receiver operating curve (ROC), yielding an area under the curve (AUC) of 0.835.

5

The biomarkers were ranked based on Random Forests estimated variable importance measure [23] derived from each panel (Figure 5 and Table 2). This further identified a core set of biomarkers which are common across the top 4 biomarker panels, comprising SSB, TROVE2 and ZHX2, with an AUC performance of 0.761 (Figure 6).

10

References

1. **Choy EH**, Panayi GS. Cytokine pathways and joint inflammation in rheumatoid arthritis. *N Engl J Med* 2001;**344**:907-916.
- 15 2. **Furst DE**, Emery P. Rheumatoid arthritis pathophysiology: update on emerging cytokine and cytokine-associated cell targets. *Rheumatology (Oxford)* 2014;**53**:1560-9.
3. **Breedveld FC**, Weisman MH, Kavanaugh AF, *et al.* The PREMIER study: a multicenter, randomized, double-blind clinical trial of combination therapy with
20 adalimumab plus methotrexate versus methotrexate alone or adalimumab alone in patients with early, aggressive rheumatoid arthritis who had not had previous methotrexate treatment. *Arthritis Rheum* 2006;**54**:26-37.
4. **Kievit W**, Fransen J, Adang EMM, *et al.* Long-term effectiveness and safety of TNF-blocking agents in daily clinical practice: results from the Dutch rheumatoid arthritis
25 monitoring register. *Rheumatology (Oxford)* 2011;**50**:196-203.
5. **Bandres Ciga S**, Salvatierra J, Lopez-Sidro M, *et al.* An examination of the mechanisms involved in secondary clinical failure to adalimumab or etanercept in

- inflammatory arthropathies. *J Clin Rheumatol* 2015;**21**:115-9.
6. **Radstake TR**, Svenson M, Eijsbouts AM, *et al.* Formation of antibodies against infliximab and adalimumab strongly correlates with functional drug levels and clinical response in rheumatoid arthritis. *Ann Rheum Dis* 2009;**68**:1739-45.
 - 5 7. **van schouwenburg PA**, van de Stadt LA, de Jong RN, *et al.* Adalimumab elicits a restricted anti-idiotypic antibody response in autoimmune patients resulting in functional neutralization. *Ann Rheum Dis* 2013;**72**:104-109.
 8. **Bartelds GM**, Krieckaert CL, Nurmohamed MT, *et al.* Development of antidrug antibodies against adalimumab and association with disease activity and treatment
10 failure during long-term follow-up. *JAMA* 2011;**305**:1460-68.
 9. **Balsa A**, Sanmarti R, Rosas J, *et al.* Drug immunogenicity in patients with inflammatory arthritis and secondary failure to tumor necrosis factor inhibitor therapies: the REASON study. *Rheumatology (Oxford)* 2018;**57**:688-93.
 10. **Chen DY**, Chen YM, Tsai WC, *et al.* Significant associations of antidrug antibody
15 levels with serum drug trough levels and therapeutic response of adalimumab and etanercept treatment in rheumatoid arthritis. *Ann Rheum Dis* 2015;**74**:e16.
 11. **Smolen JS**, Breedveld FC, Burmester GR, *et al.* Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. *Ann Rheum Dis* 2016;**75**:3-15.
 - 20 12. **Ortea I**, Roschitzki B, López-Rodríguez R, *et al.* Independent candidate serum protein biomarkers of response to adalimumab and to infliximab in rheumatoid arthritis: an exploratory study. *PLoS One* 2016;**11**:e0153140.
 13. **Cuppen B**, Fritsch-Stork R, Eekhout I, *et al.* Proteomics to predict the response to tumor necrosis factor-alpha inhibitors in rheumatoid arthritis using a supervised

- cluster-analysis based protein score. *Scand J Rheumatol* 2018;**47**:12-21.
14. **Tasaki S**, Suzuki K, Kassai Y, et al. Multi-omics monitoring of drug response in rheumatoid arthritis in pursuit of molecular remission. *Nat Commun* 2018;**9**:2755.
 15. **Van Gestel AM**, Prevoo ML, van't Hof MA, *et al.* Development and validation of
5 the European League Against Rheumatism response criteria for rheumatoid arthritis. Comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism Criteria. *Arthritis Rheum* 1996;**39**:34-40.
 16. **Duarte J**, Serufuri J-M, Mulder N, Blackburn J. Protein function microarrays: design,
10 use and bioinformatic analysis in cancer biomarker discovery and quantitation. In: Wang X, editor. *Bioinformatics of Human Proteomics*. Dordrecht: Springer 2013;39–74.
 17. **Bolstad BM**, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on bias and variance.
15 *Bioinformatics* 2003;**19**:185-193.
 18. **Johnson, WE, Rabinovic, A, and Li, C** (2007). Adjusting batch effects in microarray expression data using Empirical Bayes methods. *Biostatistics* 8(1):118-127.
 19. **Bommert, A.**, Sun, X., Bischl, B., Rahnenführer, J., Lang, M., 2020. Benchmark for filter methods for feature selection in high-dimensional classification data. *Comput.*
20 *Stat. Data Anal.* 143, 106839.
 20. **Breiman, L.**, 2001. Random forests. *Mach. Learn.* 45 (1), 5–32.
 21. **Wright, M.N.**, Ziegler, A., 2017. Ranger: A fast implementation of random forests for high dimensional data in C++ and R. *J. Stat. Softw.* 77 (1), 1–17.
 22. **Kuhn, M.**. Caret package. *Journal of Statistical Software*, 2008; 28(5)

Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., Zeileis, A., 2008. Conditional variable importance for random forests. *BMC Bioinformatics* 9 (307).

Table 1

test	No of antigen	Biomarkers in panel	ROC	Sens	Spec
ranger_permutation	20	SSB,TROVE2,ZHX2,PPARD,SPANXN2,HNRNPA2B1,TRIB2,CEP55,SH3GL1,FN3K,PANK3,HPCAL1,THRA,AIFM1,ODC1,RPS6KA4,EEF1D,KLF10,EPHA2,PRKAR1A	0.821	0.806	0.774
ranger_permutation	21	SSB,TROVE2,ZHX2,PPARD,SPANXN2,HNRNPA2B1,TRIB2,CEP55,SH3GL1,FN3K,PANK3,HPCAL1,THRA,AIFM1,ODC1,RPS6KA4,EEF1D,KLF10,EPHA2,PRKAR1A,EAPP	0.835	0.839	0.774
ranger_permutation	28	SSB,TROVE2,ZHX2,PPARD,SPANXN2,HNRNPA2B1,TRIB2,CEP55,SH3GL1,FN3K,PANK3,HPCAL1,THRA,AIFM1,ODC1,RPS6KA4,EEF1D,KLF10,EPHA2,PRKAR1A,EAPP,ZNF331,GMP5,POTEE,ASPSCR1,ECI2,ETV7,BUD31	0.824	0.839	0.710
ranger_permutation	29	SSB,TROVE2,ZHX2,PPARD,SPANXN2,HNRNPA2B1,TRIB2,CEP55,SH3GL1,FN3K,PANK3,HPCAL1,THRA,AIFM1,ODC1,RPS6KA4,EEF1D,KLF10,EPHA2,PRKAR1A,EAPP,ZNF331,GMP5,POTEE,ASPSCR1,ECI2,ETV7,BUD31,ATF3	0.811	0.839	0.774

Table 2

test	Importance value	Ranking
SSB	0.003172	1
TROVE2	0.003142	2
ZHX2	0.002135	3
PPARD	0.001290	4
SPANXN2	0.000886	5
HNRNPA2B1	0.000870	6
TRIB2	0.000825	7
CEP55	0.000814	8
SH3GL1	0.000776	9
FN3K	0.000759	10
PANK3	0.000703	11
HPCAL1	0.000689	12
THRA	0.000657	13
AIFM1	0.000639	14
ODC1	0.000638	15
RPS6KA4	0.000613	16
EEF1D	0.000590	17
KLF10	0.000588	18
EPHA2	0.000587	19
PRKAR1A	0.000587	20
EAPP	0.000565	21

Table 3

Protein Name	UniprotID	Description
SSB	P05455	Lupus La protein

Nucleotide Sequence (Seq ID No. 1):

>P003196_Q311_Q311_tube_SSB/La_6741_0_NM_003142.3_0_P05455_0

ATGGCTGAGAACGGCGACAACGAGAAGATGGCTGCTCTCGAGGCTAAGATCTGCCACCAGATCGAGTACTA
CTTCGGCGACTTCAACCTGCCCCGTGACAAGTTCCTGAAGGAACAGATCAAGCTGGACGAGGGCTGGGTGC
CCCTCGAGATCATGATCAAGTTCAACCGTCTGAACCGCTGACCACCGACTTCAACGTGATCGTCGAGGCTC
TGTCCAAGTCCAAGGCTGAGCTGATGGAATCTCCGAGGACAAGACCAAGATCCGTCGTTCCCCATCCAAG
CCCCTGCCCCGAAGTGACCGACGAGTACAAGAACGACGTGAAGAACCGTTCCGTGTACATCAAGGGTTTCCC
CACCAGCGCTACCCTGGACGACATCAAGGAATGGCTCGAGGACAAGGGCCAGGTGCTGAACATCCAGATGC
GTCGTACCCTGCACAAGGCTTTCAAGGGTTCCATCTTCGTGGTGTTCGACTCCATCGAGTCCGCTAAGAAGT
TCGTGAGACTCCCGGCCAGAAGTACAAGGAAACCGACCTGCTGATCCTGTTCAAGGACGACTACTTCGCC
AAGAAGAACGAGGAACGCAAGCAGAACAAGGTGCGAGGCCAAGCTGCGTGCTAAGCAAGAGCAAGAGGCTA
AGCAGAAGCTGGAAGAGGACGCTGAGATGAAGTCCCTGGAAGAGAAGATCGGTTGCCTGCTGAAGTTCTCC
GGCGACCTGGACGACGACCTGCCGCGAGGACCTGCACATCCTGTTCTCCAACCACGGCGAGATCAAGT
GGATCGACTTCGTGCGTGGTGTCTAAGGAAGGCATCATCTCTTCAAGGAAAAGGCCAAGGAAGCTCTGGGC
AAGGCTAAGGACGCTAACACGGCAACCTGCAGCTGCGTAACAAGGAAGTGACCTGGGAGGTCCTCGAGG
GCGAGGTGCGAGAAGGAAGCCCTGAAGAAGATCATCGAGGACCAGCAAGAGTCCCTGAACAAGTGGAAGTCC
AAGGGTCGTGCTTTCAAGGGCAAGGGAAGGGCAACAAGGCTGCTCAGCCCGGTTCCGGAAGGGGAAAGG
TGCAGTTCCAGGGCAAGAAGACCAAGTTCGTTCCGACGACGAGCAGATGAGCAGCAGAGAACGGTGCT
ACCGGTCCCGTGAAGCGTGCTCGTGAAGAGACTGACAAGGAAGAACCCGCTTCCAAGCAGCAAAAGACCGA
GAACGGCGCTGGCGACCAG

Protein Sequence (Seq ID No. 2):

>sp|P05455|LA_HUMAN Lupus La protein OS=Homo sapiens OX=9606 GN=SSB PE=1 SV=2

MAENGDNKMAALEAKICHQIEYYFGDFNLPRDKFLKEQIKLDEGWVPLEIMIKFNRLNRLTTDFNVIVEALSLSKA
ELMEISEDKTKIRRSPLPEVTDEYKNDVKNRSVYIKGFPTDATTDDIKEWLEDKGQVLNIQMRRTLHKAFKGS
FVVFDSIESAKKFVETPGQYKETDLLILFKDDYFAKKNEERKQNKVEAKLRKQEQEAKQKLEEDAEMKSLEEK
GCLLKFSGLDDQTCREDLHILFSNHGEIKWIDFVRGAKEGIIKFKEKAKEALGKAKDANNGNLQLRNKEVTWEVL
EGEVEKEALKKIIEDQQESLNKWKSGRRFKGKGKGNKAAQPGSGKGKVQFQGGKTKFASDDEHDEHDENGAT
GPVKRAREETDKEEPASKQKKTENGAGDQ

TROVE2	P10155	HUMAN 60 kDa SS-A/Ro ribonucleoprotein
--------	--------	--

Nucleotide Sequence (Seq ID No. 2):

>P001236_CAG_CAGp1_SSA2_6738_Homo sapiens Sjogren syndrome antigen A2 (60kDa ribonucleoprotein autoantigen SS-A/Ro)_BC036658.2_AAH36658.1_P10155_0_0_1617_0_1614

ATGGAGGAATCTGTAACCAAATGCAGCCACTGAATGAGAAGCAGATAGCCAATTCTCAGGATGGATATGTA
TGGCAAGTCACTGACATGAATCGACTACACCGGTTCTTATGTTTCGGTTCTGAAGGTGGGACTTATTATATCA
AAGAACAGAAGTTGGGCCCTTGAAAATGCTGAAGCTTTAATTAGATTGATTGAAGATGGCAGAGGATGTGAAG
TGATACAAGAAATAAAGTCATTTAGTCAAGAAGGCAGAACCAAGCAAGAGCCTATGCTCTTTGCACTTGC
CATTTGTTCCAGTGCTCCGACATAAGCACAAAACAAGCAGCATTTAAAGCTGTTTCTGAAGTTTGTCGATT
CCTACCCATCTCTTTACTTTTATCCAGTTTAAGAAAGATCTGAAGGAAAGCATGAAATGTGGCATGTGGGGTC
GTGCCCTCCGGAAGGCTATAGCGGACTGGTACAATGAGAAAGGTGGCATGGCCCTTGCTCTGGCAGTTACA
AAATATAAACAGAGAAATGGCTGGTCTCACAAAGATCTATTAAGATTGTCACATCTTAAACCTTCCAGTGAAG
GACTTGCAATTGTGACCAAATATATTACAAAGGGCTGGAAAGAAGTTCATGAATTGTATAAGAAAAAGCACT
CTCTGTGGAGACTGAAAAATTATTAAGTATCTGGAGGCTGTAGAGAAAGTGAAGCGCACAAAGAGATGAGCT
AGAAGTCATTCATCTAATAGAAGAACATAGATTAGTTAGAGAACATCTTTTAAACAAATCACTTAAAGTCTAAAG
AGGTATGGAAGGCTTTGTTACAAGAAATGCCGCTTACTGCATTACTAAGGAATCTAGGAAAGATGACTGCTAA
TTCAGTACTTGAACCAGGAAATTCAGAAGTATCTTTAGTATGTGAAAACTGTGTAATGAAAACTATTAATAA

AGGCTCGTATACATCCATTTTCATATTTTGATCGCATTAGAACTTACAAGACAGGTCATGGTCTCAGAGGGAA
 ACTGAAGTGGCGCCCTGATGAAGAAATTTTGAAGCATTGGATGCTGCTTTTTATAAACATTTAAGACAGTT
 GAACCAACTGGAAAACGTTTCTTACTAGCTGTTGATGTCAGTGCTTCTATGAACCAAAGAGTTTGGGTAGTA
 TACTCAACGCTAGTACAGTTGCTGCAGCAATGTGCATGGTTGTACACGAACAGAAAAAGATTCTTATGTAGT
 TGCTTTTTCCGATGAAATGGTACCATGTCCAGTGACTACAGATATGACCTTACAACAGGTTTAAATGGCTATG
 AGTCAGATCCCAGCAGGTGGAAGTATTGCTCTCTTCCAATGATCTGGGCTCAGAAGACAAACACACCTGCT
 GATGTCTTCATTGTATTCACTGATAATGAGACCTTGTCTGGAGGTGTCCATCCTGCTATTGCTCTGAGGGAGT
 ATCGAAAGAAAATGGATATTCCAGCTAAATTGATTGTTTGTGGAATGACATCAAATGGTTTCACCATTGCAGA
 CCCAGATGATAGAGCATGTTGGATATGTGCGGCTTTGATACTGGAGCTCTGGATGTAATTCGAAATTTTACA
 TTAGATATGATT

Protein Sequence (Seq ID No. 23):

>sp|P10155|RO60_HUMAN 60 kDa SS-A/Ro ribonucleoprotein OS=Homo sapiens OX=9606 GN=RO60 PE=1
 SV=2

MEESVNQMQLNEKQIANSDQGYVWQVTDNRLHRFLCFGSEGGTYIYKEQKLGLENAEALIRLIEDGRGCEVIQ
 EIKSFSQEGRTTKQEPMLFALAICSQCSDISTKQAAFKAVSEVCRIPHLFTFIQFKDLKESMKCGMWGRALRKAI
 ADWYNEKGGMALALAVTKYQQRNGWSHKDLLRLSHLKPSEGLAIVTKYITKGWKEVHELYKEKALSVEKLLK
 YLEAVEKVKRTRDELEVIHLEEHRLVREHLLTNHLKSKEVWKALLQEMPLALLRNLGKMTANSVLEPGNSEVSL
 VCEKLCNEKLLKKARIHPFHILIALETYKTGHGLRGKLRPDEEILKALDAFYKTFKTEPTGKRFLAVDVSASM
 NQRVLGSILNASTVAAAMCMVTRTEKDSYVAFSDEMVPVTTDMTLQQVLMAMSQIPAGGTDCSLPMIWAQ
 KTNPADVFIVFTDNETFAGGVHPAIALREYRKMDIPAKLIVCGMTSNGFTIADPDDRGMLDMCGFDTGALDVIR
 NFTLDMI

ZHX2

Q9Y6X8

Zinc fingers and homeoboxes protein 2

Nucleotide Sequence (Seq ID No. 3):

> P002188_Q305_Q305p3_ZHX2_22882_Homo sapiens zinc fingers and homeoboxes 2_BC042145.1_Q9Y6X8

ATGGCTAGCAAACGAAAATCTACAACCTCCATGCATGGTTTCGGACATCACAAGTAGTAGAACAAGATGTGCCC
 GAGGAAGTAGACAGGGCCAAAGAGAAAGGAATCGGCACACCACAGCCTGACGTGGCCAAGGACAGTTGGG
 CAGCAGAACTTGAAGAACTCTTCCAAAGAAAACGAAGTGATAGAGGTGAAATCTATGGGGGAAAGCCAGTCCA
 AAAAATCCAAAGTGTTATGAGTGCAAATACTGCCCTACTCCACGCAAAACCTGAACGAGTTACACGGAGC
 ATGTCGACATGCAGCATCCCAACGTGATTCTCAACCCCTCTACGTGTGTGCAGAATGTAACCTCACAACCAA
 AAAGTACGACTCCCTATCCGACCACAACCTCAAGTTCATCCCGGGGAGGCCAATTCAGCTGAAGTTAAT
 TAAACGCAATAATCAAAGTGTCTTGAACAGTCCATCGAAACCACCAACCATGTCGTGTCCATCACCACAGT
 GGCCCTGGAAGTGGTGACAGTGATTCTGGGATCTCGGTGAGTAAAACCCCATCATGAAGCCTGGAAAACC
 AAAAGCGGATGCCAAGAAGGTGCCAAGAAGCCCGAGGAGATCACCCCGAGAACCACGTGGAAGGGACC
 GCGCGCTGGTGACAGACAGCTGAGATCCTCTCGAGACTCGCGGGGTGGAGCTCCTCCAAGACACATT
 AGGACACGTATGCCCTTCTGTACAGCTGCCACCAATATCAACCTTGTCGCCAAGGTCCCTGCTCCCAATAA
 TACTACCAATAACAACCTCTGCCCTGGATACAAATGCCACGATGATCAACTCTTTCAACAAGTTTCTTACCCG
 ACCCAGGCTGAGTTGTCTGGCTGACAGCTGCCTCCAAACACCCAGAGGAGCACATCAGAATCTGGTTTGC
 CACCCAGCGCTTAAAGCATGGCATCAGCTGGTCCCCAGAAGAGGTGGAGGAGGCCCGGAAGAAGATGTTCA
 ACGGCACCATCCAGTCAGTACCCCGACCATCACTGTGCTGCCCGCCAGTTGGCCCCACAAAGGTGACG
 CAGCCCATCTCCAGACGGCTCTACCGTGCCAGATCCTCGGCCAGACTAGCCTGGTGCTGACTCAGGTGAC
 CAGCGGGTCAACAACCGTCTCTTGCTCCCCATCACACTTGCCGTGGCAGGAGTACCAACCATGGCCAGA
 AGAGACCTTGGTGACTCCCCAAGCTGCCCCGAACCCAAGCGTCCACACATCGCTCAGGTGCCAGAGCCC
 CCACCCAAGGTGGCCAACCCCCGCTCACACCAGCCAGTGACCGCAAGAAGACAAAGGAGCAGATAGCAC
 ATCTCAAGGCCAGCTTTCTCCAGAGCCAGTTCCTGACGATGCCGAGGTTTACCGGCTCATCGAGGTGACT
 GGCCTTGCCAGGAGCGAGATCAAGAAGTGGTTCAGTGACCACCGATATCGGTGTCAAAGGGGCATCGTCCA
 CATCACCAGCGAATCCCTTGCCAAAGACCAGTTGGCCATCGCGGCTCCCGACACGGTGCACGTATCATG
 CGTACCCAGACTTTGCCCCCCAGAAGTTCAAAGAGAAAACACAGGGTCAGGTTAAATCTTGGAAGACAGCT
 TTTTGAAAAGTTCTTTTCTTACCCAAGCAGAACTGGATCGGCTAAGGGTGGAGACCAAGCTGAGCAGGAGAG
 AGATCGACTCCTGGTTCTCGGAGAGGCGGAAGCTTCGAGACAGCATGGAACAAGCTGTCTTGATTCCATG
 GGGTCTGGCAAAAAGGCCAAGATGTGGGAGCCCCCAATGGTGCTCTGTCTCGACTCGACCACTCTCCGG
 TGCCAGTTAAACAAGTTCTGCCCAGCCCTTCGCCAGCAATTGCAAAAAGTCAAGAACAGGTTTCATCTCCT
 GAGGAGCACGTTTGCAAGAACCCAGTGGCCTACTCCCCAGGAGTACGACCAGTTAGCGGCCAAGACTGGCC

TGGTCCGAAGTGAAGATTGTGCGTTGGTTCAAGGAGAACAGATGCTTGCTGAAAACGGGAACCGTGAAGTGG
 ATGGAGCAGTACCAGCACCAGCCCATGGCAGATGATCACGGCTACGATGCCGTAGCAAGGAAAGCAACAAA
 ACCCATGGCCGAGAGCCCAAAGAACGGGGGTGATGTGGTCCACAATATTACAAGGACCCCAAAAGCTCT
 GCGAAGAGGACTTGAGAAAGTTGGTGACCAGGGTAAAGTAGGCAGCGAGCCAGCAAAAGACTGTTTGCCA
 GCAAAGCCCTCAGAGGCCACCTCAGACCGGTCAGAGGGCAGCAGCCGGGACGGCCAGGGTAGCGACGAG
 AACGAGGAGTCGAGCGTTGTGGATTACGTGGAGGTGACGGTCGGGGAGGAGGATGCGATCTCAGATAGAT
 CAGATAGCTGGAGTCAGGCTGCGGCAGAAAGGTGTGTCGGAAGTGGCTGAATCAGACTCCGACTGCGTCCCT
 GCAGAGGCTGGCCAGGCC

Protein Sequence (Seq ID No. 24):

>sp|Q9Y6X8|ZHX2_HUMAN Zinc fingers and homeoboxes protein 2 OS=Homo sapiens OX=9606 GN=ZHX2
 PE=1 SV=1

MASKRKSTTPCMVRTSQVVEQDVPEEVDRAKEKGIGTPQPDVAKDSWAAELENSSKENEVIEVKSMGESQSKKL
 QGGYECKYCPYSTQNLNEFTEHVDQMHPNVILNPLYVCAECNFTTKKYDSLSDHNSKFHPGEANFKLKLIRNNQ
 TVLEQSIETTNHVSITSGPGTGDSDSGISVSKTPIMKPGKPKADAKKVPKKPEEITPENHVEGTARLVTDTAELS
 RLGGVELLQDTLGHVMPVSVQLPPNINLVKVPVPLNTTKYNSALDTNATMINSFNKFPYPTQAEISWLTAASKHPE
 EHRIWFATQRLKHGISWSPEEVEEARKKMFNGTIQSVPTITVLPALAPTQVTPILQALPCQILGQTSVLTVQV
 TSGSTTVSCSPITLAVAGVTNHGQKRPLVTPQAAPEPKRPHIAQVPEPPPKVANPPLTPASDRKKTKEQIAHLKAS
 FLQSQFPDDAEVYRLIEVTGLARSEIKKWFSDHRYRCQRGIVHITSESLAKDLAIAASRHGRTHYHAYPDFAPQKF
 KEKTQGVKILEDSFLKSSFTQAELDRLRVETKLSREIDSWSFERRKLKDSMEQAVLDSMSGKKGQDVGAP
 NGALSRLDQLSGAQLTSSLPSPAIAKSQEQVHLRSTFARTQWPTPQEYDQLAAKTGLVRTEIVRWFKENRCL
 LKTGTVKWMEQYQHQPMADDHGYDAVARKATKPMAESPKNGGDVVPQYKDPKKLCEEDLEKLVTRVKVGSE
 PAKDCLPAKPSEATSDRSEGSSRDGQGSDENEESSVDYVEVTVGEEDAISDRSDSWSQAAAEVSELAESDS
 DCVPAEAGQA

PPARD	Q03181	Peroxisome proliferator-activated receptor delta
-------	--------	--

Nucleotide Sequence (Seq ID No. 4):

>P000475_SIG_SIG1-2_PPARD_5467_Homo sapiens peroxisome proliferative activated receptor delta
 transcript variant 2_BC002715.2_AAH02715.1_Q03181_51381.84_0_1086_0_1083

ATGGAGCAGCCACAGGAGGAAGCCCCTGAGGTCCGGGAAGAGGAGGAGAAAGAGGAAGTGGCAGAGGCA
 GAAGGAGCCCCAGAGCTCAATGGGGGACCACAGCATGCACTTCCTTCCAGCAGCTACACAGACCTCTCCC
 GGAGCTCCTCGCCACCCTCACTGCTGGACCAACTGCAGATGGGCTGTGACGGGGCCTCATGCGGCAGCCT
 CAACATGGAGTGCCGGGTGTGCGGGGACAAGGCATCGGGCTTCCACTACGGTGTTTCATGCATGTGAGGGG
 TGCAAGGGCTTCTCCGTCGTACGATCCGCATGAAGCTGGAGTACGAGAAGTGTGAGCGCAGCTGCAAGAT
 TCAGAAGAAGAACCGCAACAAGTGCCAGTACTGCCGCTTCCAGAAGTGCCTGGCACTGGGCATGTACACACA
 ACGCTATCCGTTTTGGTCCGATGCCGGAGGCTGAGAAGAGGAAGCTGGTGGCAGGGCTGACTGCAAAATGA
 GGGGAGCCAGTACAACCCACAGGTGGCCGACCTGCAAGGCCTTCTCCAAGCACATCTACAATGCCTCACTGA
 AAAACTTCAACATGACCAAAAAGAAGGCCCGCAGCATCCTCACCGGCAAAGCCAGCCACACGGCGCCCTTT
 GTGATCCACGACATCGAGACATTGTGGCAGGCAGAGAAGGGGCTGGTGTGGAAGCAGTTGGTGAATGGCC
 TGCCTCCCTACAAGGAGATCAGCGTGCACGTCTTCTACCGCTGCCAGTGCACCACAGTGGAGACCGTGCG
 GGAGCTCACTGAGTTGCCAAGAGCATCCCCAGCTTACGAGCCTCTTCTCAACGACCAGGTTACCCTTC
 TCAAGTATGGCGTGACGAGGCCATCTCGCCATGCTGGCCTCTATCGTCAACAAGGACGGGCTGCTGGTA
 GCCAACGGCAGTGGCTTTGTACCCGTGAGTTCTGCGCAGCCTCCGCAAACCTTCAGTGATATCATTGA
 GCCTAAGTTTGAATTTGCTGTCAAGTTCAACGCCCTGGAACCTTGATGACAGTGACCTGGCCCTATTATTGC
 GGCCATCATTCTGTGTGGAGGTGAG

Protein Sequence (Seq ID No. 25):

>sp|Q03181|PPARD_HUMAN Peroxisome proliferator-activated receptor delta OS=Homo sapiens OX=9606
 GN=PPARD PE=1 SV=1

MEQPQEEAPEVREEEEEKEEVAEAGAPELNNGGPQHALPSSSYTDLSSSSPPSLLDQLQMGCDGASCGSLNM
 ECRVCGDKASGFHYGVHACEGCKGFFRRTIRMKLEYEKERSCKIQKKNRNKCQYCRFQKCLALGMSHNAIRF
 GRMPEAEKRKLVALGTANEGSQYNPQVADLKAFSKHIYNAYLKNFNMTKKKARSILTGTKASHTAPFVIHDIETLW
 QAEKGLVWKQLVNLPPYKEISVHVFYRCQCTTVETVRELTEFAKSIPSFSSFLNDQVTLTKYGVHEAIFAMLASI

VNKDGLLVANGSGFVTREFLRSLRKPFSDIIEPKFEFAVKFNALELDDSDLALFIAAILCGDRPGLMNVPRVEAIQD
TILRALEFHLQANHPDAQYLFPKLLQKMADLRQLVTEHAQMMQRIKKTETETSLHPLLQEIKDMY

SPANXN2 Q5MJ10 Sperm protein associated with the nucleus on the X chromosome N2

Nucleotide Sequence (Seq ID No. 5):

>P003098_Q211_Q211_tube_SPANXN2_494119_0_NM_001009615.1_0_Q5MJ10_0_Insert sequence is gene optimized by GeneArt_0_0_0

ATGGAACAGCCCACCTCTTCCACCAACGGCGAGAAGCGCAAGTCCCCCTGCGAGTCCAACAACAAGAAAAA
CGACGAGATGCAAGAGGCTCCCAACCGTGTGCTGGCTCCCAAGCAGTCCCTGCAAAAGACCAAGACCATC
GAGTACCTGACCATCATCGTGTACTACTACCGCAAGCACACCAAGATCAACTCCAACAGCTCGAGAAGGA
CCAGTCCCGCGAGAAGTCCATCAACCCCGTGCAAGAGGAAGAGGACGAGGGCCTGGACTCCGCTGAGGG
ATCCTCCCAAGAAGATGAGGACCTGGACAGCTCCGAGGGTTCAGCCAAGAGGATGAAGATCTCGACTCCT
CCGAGGGCAGCTCCCAAGAGGACGAGGACTTGGATTCTCCGAGGGATCTAGTCAAGAGGACGAGGATCT
GGACTCTTCCGAAGGCTCATCTCAAGAAGATGAAGATTGGACCCCCCTGAGGGTAGCAGTCAAGAGGATG
AGGACCTCGATTCCAGCGAGGGCTCCTCACAAGAGGGTGGCGAGGAT

Protein Sequence (Seq ID No. 26):

>sp|Q5MJ10|SPXN2_HUMAN Sperm protein associated with the nucleus on the X chromosome N2 OS=Homo sapiens OX=9606 GN=SPANXN2 PE=1 SV=1

MEQPTSSTNGEKRKSPCESNNKKNDEMQEAPNRVLAPKQSLQKTKTIEYLTIIVYYRKHKTINSNQLEKDQSRE
NSINPVQEEDEGLDSAEGSSQEDEDLDSSEGSSQEDEDLDSSEGSSQEDEDLDSSEGSSQEDEDLDSSEGSS
QEDEDLDPPEGSSQEDEDLDSSEGSSQEGGED

HNRNPA2B1 P22626 Heterogeneous nuclear ribonucleoproteins A2/B1

Nucleotide Sequence (Seq ID No. 6):

>P003186_Q311_Q311_tube_HNRNPA2B1_3181_0_NM_002137.3_0_P22626_0

ATGGAACGCGAGAAAGAGCAGTTCCGCAAGCTGTTTCATCGGTGGCCTGTCCTTCGAGACTACCGAGGAATC
CCTGCGCAACTACTACGAGCAGTGGGGCAAGCTGACCGACTGCGTGGTCATGCGTGACCCCGCTTCCAAG
CGTTCCCGTGGTTTCGTTTCGTGACCTTCTCCAGCATGGCTGAGGTGGACGCTGCTATGGCTGCTCGTCC
CCACTCCATCGACGGTCGTGTGGTCGAGCCTAAGCGTGCTGTGGCTCGTGAAGAGTCCGGCAAGCCTGGT
GCTCACGTGACCGTGAAGAAGCTGTTTCGTTGGCGGTATCAAAGAGGACACCGAGGAACACCACCTGAGGG
ACTACTTCGAGGAATACGGCAAGATCGACACCATCGAGATCATCACCGACCGTCAGTCCGGAAAGAAGCGC
GGCTTCGGCTTCGTCACTTTCGACGACACGACCCCGTGGACAAGATCGTGTGCGAGAAGTACCACACCAT
CAACGGTCACAACGCTGAAGTGCGCAAGGCTCTGTCCCGTCAAGAGATGCAAGAGGTGCAGTCTCCCGTT
CCGGTTCGTGGTGGCAACTTCGGATTTCGGCGACTCTCGCGGTGGTGGCGGAACTTCGGTCTGTCGTCG
TTCCAATTCCGTGGTGGTTCGACGGTTACGGCTCCGGAAGAGGTTTCGGCGACGGCTACAACGGCTAC
GGTGGTGGTCTGCGGTGGAAATTCGGTGGTTCCTGTTACGGTGGCGGTTCGCGGTGGATACGGCG
GAGGTGGTCCAGGATACGGCAACCAGGTGGCGGTACGGCGGTGGTACGACAACTACGGTGGCGGCA
ACTACGGTTCGGGAACTACAACGACTTCGGCAATTACAACAGCAGCCCTCCAACACGGCCCCATGAAG
TCTGGCAATTCGGCGGCTCCCGTAACATGGGTGGTCTTACGGTGGTGGAAATTACGGTCCCGGTGGTTC
CGGTGGCTCTGGTGGCTACGGCGGTCTTCCCGTTAC

Protein Sequence (Seq ID No. 27):

>sp|P22626|ROA2_HUMAN Heterogeneous nuclear ribonucleoproteins A2/B1 OS=Homo sapiens OX=9606 GN=HNRNPA2B1 PE=1 SV=2

MEKTLETVPLERKKREKEQFRKLFIGGLSFETTEESLRNYEYQWGKLTDCVVMRDPASKRSRFGFVTFSSMAE
VDAAMAARPHSIDGRVVEPKRAVAREESGKPGAHVTVKKLFVGGIKEDTEHHLRDYFEEYKIDTIEITDRQSG
KKRGFGFVTFDDHDPVDKIVLQKYHTINGHNAEVRKALSRQEMQEVQSSRSRGGNFGFGDSRGGGNGFGPG
PGSNFRGSDGYGSGRFGDGYNGYGGGPGGNGFGGSPGYGGGRGYYGGGPGYGNQGGGYGGGYDNY
GGGNYGSGNYNDFGNYNQQPSNYGPMKSGNFGGSRNMGGPYGGGNYGPGSGSGSGYGGRSRY

TRIB2 Q92519 HUMAN Tribbles homolog 2

Nucleotide Sequence (Seq ID No. 7):

>P001066_KIN2_KIN2p1_TRB2_28951_Homo sapiens tribbles homolog
2_BC002637.2_AA02637.1_Q92519_0_0_1032_0_1029

ATGAACATACACAGGTCTACCCCCATCACAATAGCGAGATATGGGAGATCGCGGAACAAAACCCAGGATTT
CGAAGAGTTGTCGTCTATAAGGTCCGCGGAGCCAGCCAGAGTTTCAGCCCCAACCTCGGCTCCCCGAGC
CCGCCCCGAGACTCCGAACCTTGTGCGATTGCGTTTCTTGATCGGGAAATACTTATTGTTGGAACCTCTGGAG
GGAGACCACGTTTTTCGTGCCGTGCATCTGCACAGCGGAGAGGAGCTGGTGTGCAAGGTGTTTGATATCAG
CTGCTACCAGGAATCCCTGGCACCGTGCTTTTGCCTGTCTGCTCATAGTAACATCAACCAAATCACTGAAAT
TATCCTGGGTGAGACCAAAGCCTATGTGTTCTTTGAGCGAAGCTATGGGGACATGCATTCCTTCGTCCGCAC
CTGCAAGAAGCTGAGAGAGGAGGAGGCAGCCAGACTGTTCTACCAGATTGCCTCGGCAGTGGCCCACTGC
CATGACGGGGGGCTGGTGTGCTGCGGGACCTCAAGCTGCGGAAATTCATCTTTAAGGACGAAGAGAGGACTC
GGGTCAAGCTGGAAGCCTGGAAGACGCCTACATTCTGCGGGGAGATGATGATTCCCTCTCCGACAAGCAT
GGCTGCCCCGGCTTACGTAAGCCCAGAGATCTTGAACACCAAGTGGCAGCTACTCGGGCAAAGCAGCCGACG
TGTGGAGCCTGGGGGTGATGCTGTACACCATGTTGGTGGGGCGGTACCCTTTCCATGACATTGAACCCAGC
TCCCTCTTCAGCAAGATCCGGCGTGCCAGTTCAACATTCCAGAGACTCTGTGCCCCAAGGCCAAGTGCCT
CATCCGAAGCATTCTGCGTCGGGAGCCCTCAGAGCGGCTGACCTCGCAGGAAATTCGGACCATCCTTGGT
TTTCTACAGATTTTAGCGTCTCGAATTCAGCATATGGTGCTAAGGAAGTGTCTGACCAGCTGGTGCCGGACG
TCAACATGGAAGAGAACTTGGACCCTTTCTTTAAC

Protein Sequence (Seq ID No. 28):

>sp|Q92519|TRIB2_HUMAN Tribbles homolog 2 OS=Homo sapiens OX=9606 GN=TRIB2 PE=1 SV=1

MNIHRSTPITIARYGRSRNKTQDFEELSSIRSAEPSQSFSFNLGSPSPPETPNLSHCVSCIGKYLLLEPLEGDHVF
AVHLHSGEELVCKVFDISCYQESLAPCFCLSAHSNINQITEIILGETKAYVFFERSYGDMHSFVRTCKKLREEEAAR
LFYQIASAVAHCHDGLVLRDLKLRKFIFKDEERTRVKLESLEDAYILRGDDDSLSDKHGCPAYVSPILNTSGSY
SGKAADVWSLGVMLYTMVGRYPFHDIPESSLFSKIRRGQFNIPETLSPKAKCLIRSILRREPSERLTSQEILDHPW
FSTDFSVSNSAYGAKEVSDQLVPDVNMEENLDPFFN

CEP55 Q53EZ4 Centrosomal protein of 55 kDa

Nucleotide Sequence (Seq ID No. 8):

>P003121_Q211_Q211_tube_CEP55_55165_0_NM_001127182.1_0_Q53EZ4_0

ATGTCCTCCCGTTCCACCAAGGACCTGATCAAGTCTAAGTGGGGTTCCAAGCCCTCCAAGTCCAAGTCCGA
GACTACCCTCGAGAAGCTGAAGGGCGAGATCGCTCACCTCAAGACCTCCGTGGACGAGATCACCTCCGGC
AAGGGCAAGCTGACCGACAAGGAACGTCACCGTCTGCTCGAGAAGATCCGTGTGCTCGAGGCTGAGAAGG
AAAAGAACGCTTACCAGCTGACTGAGAAGGACAAGGAAATCCAGCGTCTGCGCGACCAGCTGAAGGCTCGT
TACTCCACCACCGCTCTGCTGGAACAGCTGGAAGAAACCACCGCGAGGGCGAGCGTCTGCGAGCAGGTCC
TGAAGGCTCTGTCCGAAGAGAAGGACGTGCTGAAGCAGCAGCTGTCCGCTGCTACCTCCCGTATCGCTGA
GCTGGAATCCAAGACCAACACCTGCGTCTGTCCCAGACCGTGGCTCCCAACTGCTTCAACTCCTCCATCA
ACAACATCCACGAGATGGAATCCAAGTGAAGGACGCTCTCGAGAAGAACCAGCAGTGGCTGGTGTACGAC
CAGCAGCGCGAGGTGTACGTGAAGGGCCTGCTGGCTAAGATCTTCGAGCTGGAAAAGAAGACCGAGACTG
CTGCTCACTCCCTGCCCCAGCAGACCAAGAAGCCCCGAGTCCGAGGGTTACCTGCAAGAGGAAAAGCAGAA
GTGCTACAACGACCTGCTGGCTTCCGCTAAGAAGGACCTGGAAGTCGAGCGTCAGACCATCACCCAGCTGT
CCTTCGAGCTGTCCGAGTTCGCTAGGAAGTACGAAGAGACTCAGAAGGAAGTCCACAACCTGAACCAGCTG
CTGTACTCCCAGCGTCGTGCTGACGTGCAGCACCTCGAGGACGACCGTCACAAGACTGAGAAGATCCAGA
AGCTGCGCGAAGAGAACGATATCGCTCGTGGCAAGCTCGAGGAAGAGAAGAAGCGTTCCGAGGAAGTCTGCT
GTCCCAGGTGCAGTTCCTGTACACCTCCCTGCTCAAGCAGCAAGAGGAACAGACCCGTTGTGGCTCTGTTGG
AGCAGCAGATGCAGGCTTGACCCCTGGACTTCGAGAACGAGAAGCTGGACCGTCAGCACGTCCAGCACCA
GCTGCACGTGATCCTGAAGGAACTGCGCAAGGCTCGTAACCAGATCACCCAGTTGGAGTCCCTGAAGCAG
CTGCACGAGTTCGCTATCACCGAGCCCCCTGGTCACCTTCCAAGGCGAGACTGAGAACCAGCGAGAAGGTGG
CCGCTTCCCCCAAGTCCCCCACCCTGCTCTGAACGAGTCCCTGGTCGAGTGCCCCAAGTGAACATCCA
GTACCCCGCTACCGAGCACCGTGACCTGCTGGTGCACGTCGAGTACTGCTCCAAG

Protein Sequence (Seq ID No. 29):

>sp|Q53EZ4|CEP55_HUMAN Centrosomal protein of 55 kDa OS=Homo sapiens OX=9606 GN=CEP55 PE=1 SV=3

MSSRSTKDLISKWGSKPSNSKSETTLEKLKGEIAHLKTSVDEITSGKGKLTDKERHRLLEKIRVLEAEKEKNAYQ
LTEKDKEIQLRDQLKARYSTTTTLEQLEETTREGERRQVLKALSEEKDVLLKQLSAATSRIALESKTNTLRLS
QTVAPNCFNSSINNIHEMEIQLKDALEKNQWLVDQQRREVYVKGLLAKIFELEKKTETAHSLPQQTKKPESEG
YLQEEKQKCYNDLLASAKKDLEVERQTITQLSFELSEFRRYEETQKEVHNLNQLLYSQRRADVQHLEDDRHKHT
EKIQKLREENDIARGKLEEEKRSEELLQVQFLYTSLLKQQEEQTRVALLEQQMQACTLDFENEKLDROHVQH
QLHVILKELRKARNQITQLESLLQLHEFAITEPLVTFQGETENREKVAASPKSPTAALNESLVECPKCNQYPATEH
RDLVHVEYCSK

SH3GL1 Q99961 Endophilin-A2

Nucleotide Sequence (Seq ID No. 9):

>P000121_CAN_CAN1-1_SH3GL1_6455_Homo sapiens SH3-domain GRB2-like
1_BC001270.1_AAH01270.1_Q99961_0_0_1107_0_1104

ATGTCGGTGGCGGGGCTGAAGAAGCAGTTCTACAAGGCGAGCCAGCTGGTCAGTGAGAAGGTCGGAGGG
GCCGAGGGGACCAAGCTGGATGATGACTTCAAAGAGATGGAGAAGAAGGTGGATGTCACCAGCAAGGCGG
TGACAGAAGTGCTGGCCAGGACCATCGAGTACCTGCAGCCCCAACCCAGCCTCGCGGGCTAAGCTGACCAT
GCTCAACACGGTGTCCAAGATCCGGGGCCAGGTGAAGAACCCCGGCTACCCGCAGTCGGAGGGGCTTCTG
GGCGAGTGATGATCCGCCACGGGAAGGAGCTGGGCGGCGAGTCCAACCTTGGTGACGCATTGCTGGATG
CCGGCGAGTCCATGAAGCGCCTGGCAGAGGTGAAGGACTCCCTGGACATCGAGGTCAAGCAGAACTTCAT
TGACCCCTCCAGAACCTGTGCGAGAAAGACCTGAAGGAGATCCAGCACCACTGAAGAACTGGAGGGC
CGCCGCTGGACTTTGACTACAAGAAGAAGCGGCGAGGGCAAGATCCCCGATGAGGAGCTACGCCAGGCGC
TGGAGAAGTTCGAGGAGTCCAAGGAGGTGGCAGAAACCAGCATGCACAACCTCCTGGAGACTGACATCGA
GCAGGTGAGTCAGCTCTCGGCCCTGGTGGATGCACAGCTGGACTACCACCGGCAGGCCGTGCAGATCCTG
GACGAGCTGGCGGAGAAGCTCAAGCGCAGGATGCGGGAAGCTTCTCACGCCCTAAGCGGGAGTATAAGC
CGAAGCCCCGGGAGCCCTTTGACCTTGGAGAGCCTGAGCAGTCCAACGGGGGCTTCCCCTGCACCACAGC
CCCCAAGATCGCAGCTTCATCGTCTTTCCGATCTTCCGACAAGCCCATCCGGACCCCTAGCCGGAGCATGC
CGCCCTGGACCAGCCGAGCTGCAAGGCGCTGTACGACTTCGAGCCCGAGAACGACGGGGAGCTGGGCT
TCCATGAGGGCGACGTCATCACGCTGACCAACCAGATCGATGAGAACTGGTACGAGGGCATGCTGGACGG
CCAGTCGGGCTTCTTCCCGCTCAGCTACGTGGAGGTGCTTGTGCCCCTGCCGCGAG

Protein Sequence (Seq ID No. 30):

>sp|Q99961|SH3G1_HUMAN Endophilin-A2 OS=Homo sapiens OX=9606 GN=SH3GL1 PE=1 SV=1

MSVAGLKKQFYKASQLVSEKVGGAEGTKLDDDFKEMEKKVDVTSKAVTEVLARTIEYLQPNPASRAKLTMLNTV
SKIRGQVKNPGYPQSEGLLGECMIRHGKELGGESNFGDALLDAGESMKRLAEVKDSLIDIEVKQNFIDPLQNLCEK
DLKEIQHHLKKLEGRRLLDFDYKKRQKGKIPDEELRQALEKFEESKEVAETSMHNLLETIDIEQVSQLSALVDAQLDY
HRQAVQILDELAELKRRMREASSRPKREYKPKPREPFDLGEPEQSNGGFPCTTAPKIAASSSFRSSDKPIRTPS
RSMPLDQPSCKALYDFEPENDGELGFHEGDVITLTNQIDENWYEGMLDGQSGFFPLSYVEVLVPLPQ

FN3K Q9H479 Fructosamine-3-kinase

Nucleotide Sequence (Seq ID No. 10):

>P002359_Q106_Q106p1_FN3K_64122_Homo sapiens fructosamine 3
kinase_BC042680.1_AAH42680.1_Q9H479_0_0_930_0_927

ATGGAGCAGCTGCTGCGCGCCGAGCTGCGCACCGCGACCCTGCGGGCCTTCGGCGGGCCCCGGCGCCGG
CTGCATCAGCGAGGGCCGAGCCTACGACACGGACGCAGGCCAGTGTCGTCAAAGTCAACCGCAGGACG
CAGGCCCGGCAGATGTTTGAGGGGGAGGTGGCCAGCCTGGAGGCCCTCCGAGACAGGGCCTGGTGCG
GGTGCCGAGGCCCATGAAGGTCATCGACCTGCCGGGAGGTGGGGCCGCCTTTGTGATGGAGCATTGAAG
ATGAAGAGCTTGAGCAGTCAAGCATCAAACTTGGAGAGCAGATGGCAGATTTGCATCTTTACAACCAGAAG
CTCAGGGAGAAGTTGAAGGAGGAGGAGAACACAGTGGGCCGAAGAGGTGAGGGTGCTGAGCCTCAGTATG
TGGACAAGTTCGGCTTCCACACGGTGACGTGCTGCGGCTTCATCCCGCAGGTGAATGAGTGGCAGGATGA

CTGGCCGACCTTTTTCGCCCGGCACCGGCTCCAGGCGCAGCTGGACCTCATTGAGAAGGACTATGCTGAC
 CGAGAGGCACGAGAACTCTGGTCCCGGTACAGGTGAAGATCCCGGATCTGTTTTGTGGCCTAGAGATTGT
 CCCC CGCTTGCTCCACGGGGATCTCTGGTCGGGAAACGTGGCTGAGGACGACGTGGGGGCCATTATTTAC
 GACCCGGCTTCCTTCTATGGCCATTCCGAGTTTGAACCTGGCAATCGCCTTGATGTTTGGGGGGTCCCCAG
 ATCCTTCTTACCCGCTACCACCGGAAGATCCCCAAGGCTCCGGGCTTCGACCAGCGGCTGCTGCTCTACC
 AGCTGTTTAACTACCTGAACCACTGGAACCACTTCGGGCGGGAGTACAGGAGCCCTTCGTTGGGCACCATG
 CGAAGGCTGCTCAAG

Protein Sequence (Seq ID No. 31):

>sp|Q9H479|FN3K_HUMAN Fructosamine-3-kinase OS=Homo sapiens OX=9606 GN=FN3K PE=1 SV=1

MEQLLRAELRTATLRAFGGPGAGCISEGRAYDTDAGPVFVKVNRRTQARQMFEGEVASLEALRSTGLVRVPRP
 MKVIDLPGGGAFAFVMEHLKMKSLSSQASKLGEQMA DLHLYNQKLREKLKEEENTVGRRGEGAEPQYVDKFGFH
 TVTCCGFIPQVNEWQDDWPTFFARHRLQAQLDLIEKDYADREELWSRLQVKIPDLFCGLEIVPALLHGD LWS
 GNVAEDDVGPPIYDPASFYGHSEFELAIALMFGGFP RSFFTAYHRKIPKAPGFDQRLLLYQLFNYLNHNHFGRE
 YRSPSLGTMRRLLK

PANK3	Q9H999	Pantothenate kinase 3
-------	--------	-----------------------

Nucleotide Sequence (Seq ID No. 11):

>P002239_Q106_Q106p2_PANK3_79646_Homo sapiens pantothenate kinase
 3_BC013705.1_AAH13705.1_Q9H999_0_0_1113_0_1110

ATGAAGATCAAAGATGCCAAGAAACCCTCTTTCCCATGGTTTGGCATGGACATTGGGGGAACTCTAGTAAAG
 CTCTCGTACTTTGAACCTATTGATATCACAGCAGAGGAAGAGCAAGAAGAAGTTGAGAGTTTAAAAAGTATTC
 GGAAATATTTGACTTCTAACGTGGCATATGGATCCACCGGCATTCCGGATGTACACCTTGAACCTGAAAGATT
 TAACACTTTTTGGCCGAAGAGGGAACCTTGCACTTTATCAGGTTTCCAACCCAGGACCTGCCTACTTTTATCCA
 AATGGGAAGAGATAAAAACTTCTCAACATTGCAGACGGTGCTATGTGCTACAGGAGGTGGTGCTTACAAGTT
 TGA AAAAGATTTTCGCACAATTGGAAACCTCCACCTGCACAACTGGATGAACTTGACTGCCTTGTAAGGG
 CTTGCTGTATATAGACTCTGTCA GTTTCAATGGACAAGCCGAGTGCTATTATTTTGCTAATGCCTCAGAACCT
 GAGCGATGCCAAAAGATGCCTTTTAACCTGGATGATCCCTATCCACTGCTTGTAAGTGAACATTGGCTCAGGA
 GTCAGTATTTTAGCAGTCCATTCCAAAGACA ACTATAAACGAGTGACTGGGACAAGCCTTGAGGGGGGTAC
 CTTTCTGGGTTTATGCAGTTTATTGACTGGCTGTGAAAGTTTTGAAGAGGCTCTTGAAATGGCATCCAAAGGT
 GATAGCACACAAGCTGACAAGCTGGTCCGTGATTTATGGAGGAGATTATGAAAGATTTGGTTTGCCAGGT
 TGGGCTGTAGCATCTAGTTTGGGAATATGATTTATAAGGAGAAGCGAGAATCTGTTAGTAAAGAAGATCTG
 GCAAGAGCTACTTTAGTTACTATACCAATAACATTGGTTCTGTGGCAGCAATGTGTGCTGTTAATGAGAAAA
 TAAACAGAGTTGCTTTGTTGGAACTTTTACGTGTCAATACCCTCTCAATGAACTTTTGGCATATGCACT
 GGATTACTGGTCAAAAGGTCAACTAAAAGCATTGTTTCTAGAACATGAGGGTTACTTTGGAGCAGTTGGTGC
 ACTTCTTGGGCTGCCAAATTT CAGC

Protein Sequence (Seq ID No. 32):

>sp|Q9H999|PANK3_HUMAN Pantothenate kinase 3 OS=Homo sapiens OX=9606 GN=PANK3 PE=1 SV=1

MKIKDAKKPSFPWFGMDIGGTLVKLSYFEPIDITAE EEEQEEVESLKSIRKYLTSNVAYGSTGIRDVHLELKD LTLFG
 RRGNLHFIRFPTQDLPTFIQMGRDKNFSTLQTVLCATGGGAYKFEKDFRTIGNLHLHLKDELDCLVKG LLYIDSVS
 FNGQAECYYFANASEPERCQKMPFNLD DPYPLL VVNIGSGVSILAVHSDKDNYKRVGTGSLGGGTFLGLCSLLTGC
 ESFEEALEMASKGDSTQADKLVRDIYGGDYERFGLPGWAVASSFGNMIYKEKRESVSKEDLARATLVITNNIGS
 VARMCVNEKINRVVFGNFLRVNTLSMKLLAYALDYWSKGQLKALFLEHEGYFGAVGALLGLPNFS

HPCAL1	P37235	Hippocalcin-like protein 1
--------	--------	----------------------------

Nucleotide Sequence (Seq ID No. 12):

>P003172_Q311_Q311_tube_HPCAL1_3241_0_NM_002149.2_0_P37235_0

ATGGGCAAGCAGAACTCCAAGCTGCGTCCCGAGGTGCTGCAGGACCTGCGCGAGAACACCGAGTTCACCG
 ACCACGAGCTGCAAGAGTGGTACAAGGGTTTCTGAAGGACTGCCCCACCGGTACCTGACCGTGGACGA
 GTTCAAGAAGATCTACGCTA ACTTCTCCCTACGGCGACGCTTCCAAGTTCGCTGAGCACGTGTTCCGTAC

CTTCGACACCAACGGCGACGGCACCATCGACTTCCGCGAGTTCATCATCGCTCTGTCCGTGACCTCCCGTG
GCAAGCTCGAGCAAAAGCTGAAGTGGGCTTTCTCGATGTACGACCTGGACGGCAACGGTTACATCTCCCGT
TCCGAGATGCTCGAGATCGTGCAGGCTATCTACAAGATGGTGTCTCCGTGATGAAGATGCCCCGAGGACGA
GTCCACCCCCGAGAAGCGTACCGACAAGATCTTCCGTGATGGACACCAACAACGACGGAAAGCTGTCCC
TGGAAGAGTTCATCCGTGGTGCTAAGTCCGACCCCTCCATCGTGCCTGTCTGTCAGTGCACCCATCCTCC
GCTTCCAGTTC

Protein Sequence (Seq ID No. 33):

>sp|P37235|HPCL1_HUMAN Hippocalcin-like protein 1 OS=Homo sapiens OX=9606 GN=HPCL1 PE=1 SV=3

MGKQNSKLRPEVLQDLRENTFTHLQEWYKGLKDCPTGHLTVDEFKKIYANFFPYGDASKFAEHVFRFTD
NGDGTIDFREIIALSVTSRGKLEQKLKWFMSYDLGNGYISRSEMLEIVQAIYKVMVSSVMKMPEDSTPEKRTD
KIFRQMDTNNDGKLSLEEFIRGAKSDPSIVRLLQCDPSSASQF

THRA P10827 Thyroid hormone receptor alpha

Nucleotide Sequence (Seq ID No. 13):

>P000757_TRN_TRNp2_THRA_7067_Homo sapiens Homo sapiens thyroid hormone receptor alpha
(erythroblastic leukemia viral (v-erb-a) onc_BC000261.1_AAH00261.1_P10827_0_0_1473_0_1470

ATGGAACAGAAGCCAAGCAAGGTGGAGTGTGGGTGAGACCCAGAGGAGAAGAGTCCAGGTACCAGATG
GAAAGCGAAAAAGAAAGAACGGCCAATGTTCCCTGAAAACCAGCATGTGAGGGTATATCCCTAGTTACCTG
GACAAAGACGAGCAGTGTGTCGTGTGGGGACAAGGCAACTGGTTATCACTACCGCTGTATCACTTGTGA
GGGCTGCAAGGGCTTCTTTCGCCGCACAATCCAGAAGAAGCTCCATCCACCTATTCTGCAATATGACA
GCTGCTGTGTCATTGACAAGATCACCCGCAATCAGTGCCAGCTGTGCCGCTTCAAGAAGTGCATCGCCGTG
GGCATGGCCATGGACTTGGTTCTAGATGACTCGAAGCGGGTGGCCAAGCGTAAGCTGATTGAGCAGAACC
GGGAGCGGGCGGGAAGGAGGAGATGATCCGATCACTGCAGCAGCGACCCAGAGCCCACTCCTGAAGAGT
GGGATCTGATCCACATTGCCACAGAGGCCCATCGCAGCACCAATGCCAGGGCAGCCATTGAAACAGAG
GCGGAAATTCCTGCCCGATGACATTGGCCAGTCACCCATTGTCTCCATGCCGACGGAGACAAGGTGGAC
CTGGAAGCCTTCAGCGAGTTTACCAAGATCATCACCCGGCCATCACCCGTGTGGTGGACTTTGCCAAAAA
ACTGCCCATGTTCTCCGAGCTGCCTTGCGAAGACCAGATCATCCTCCTGAAGGGGTGCTGCATGGAGATCA
TGTCCTGCGGGCGGCTGTCCGCTACGACCCTGAGAGCGACACCCTGACGCTGAGTGGGGAGATGGCTGT
CAAGCGGGAGCAGCTCAAGAATGGCGGCCTGGGCGTAGTCTCCGACGCCATCTTTGAAGTGGGCAAGTCA
CTCTCTGCCTTAACCTGGATGACACGGAAGTGGCTCTGCTGCAGGCTGTGCTGCTAATGTCAACAGACCG
CTCGGGCCTGCTGTGTGTGGACAAGATCGAGAAGAGTCAGGAGGCGTACCTGCTGGCGTTCCGAGCACTAC
GTCAACCACCGCAAACACAACATTCCGCACTTCTGGCCCAAGCTGCTGATGAAGGAGAGAGAAGTGCAGAG
TTCGATTCTGTACAAGGGGGCAGCGGCAGAAGGCCGGCGGGCGGGTCACTGGGCGTCCACCCGGAAGG
ACAGCAGCTTCTCGGAATGCATGTTGTTACGGGTCCGAGGTCCGGCAGCTTGAAGCAGCAGCTTGGTGAA
GCGGGAAGTCTCAAGGGCGCGTCTTTCAGCACAGAGCCCGAAGAGCCCGCAGCAGCGTCTCCTGGAG
CTGCTCCACCGAAGCGGAATTCTCCATGCCGAGCGGTCTGTGGGAAGACGACAGCAGTGAAGGCGGACT
CCCCGAGCTCCTCTGAGGAGAACCGGAGGTCTGCGAGGACCTGGCAGGCAATGCAGCCTCTCCC

Protein Sequence (Seq ID No. 34):

>sp|P10827|THA_HUMAN Thyroid hormone receptor alpha OS=Homo sapiens OX=9606 GN=THRA PE=1
SV=1

MEQKPSKVECGSDPEENSARSPDGKRKRKNGQCSLKTSMSGYIPSYLDKDEQCVVCGDKATGYHYRCITCEGC
KGFFRRTIQKNLHPTYSCYDSCCVIDKITRNQCQLCRFKKCIAGVMAMDLVDDSKRVAKRKLIEQNRERRRKE
EMIRSLQQRPEPTPEEWDLIHATEAHRSTNAQGSWWKQRRKFLPDDIGQSPIVSMPDGDVLEAFSEFTKIITP
AITRVVDFAKKLPMFSELPCEQIILLKGCCMEIMSLRAAVRYDPESDTLTLGEMAVKREQLKNGGLGVVSDAIF
ELGKLSAFNLDDTEVALLQAVLLMSTDRSGLLCVDKIEKSQEAYLLAFEHYVNRKHNIPHFVWPKLLMKEREVQ
SSILYKGAAGAEGRPGSLGVHPEGQQLGMHVQGPQVRQLEQQLGEAGSLQGPVLQHQSPKSPQQRLELL
HRSGILHARAVCGEDDSSEADSPSSSEEEPEVCEDLAGNAASP

AIFM1 O95831 Apoptosis-inducing factor 1, mitochondrial

Nucleotide Sequence (Seq ID No. 14):

>P003305_Q311_Q311_tube_AIFM1_9131_Apoptosis-inducing factor, mitochondrion-associated, 1 [Homo sapiens]_NM_001130846.2_0_0_0_0_0_0

ATGGAAGGTCCGCCGAGGGTGTCAAGGTCATGCCCAACGCTATCGTGCAGTCCGTGGGCGTGTCTCT
CCGGCAAGCTGCTGATCAAGCTGAAGGACGGTCGCAAGGTGGAAACCGACCATCGTGGCTGCTGTGGG
CCTCGAGCCCAACGTCGAGCTGGCTAAGACCGGTGGCCTCGAGATCGACTCCGACTTCGGTGGTTTCCGT
GTGAACGCTGAGCTGCAGGCTCGTTCCAACATCTGGGTGGCCGGCGACGCTGCTTCTTACGACATCAA
GCTGGGTGCTGCTGCTGTCGAGCACCACGACACGCTGTGGTGTCCGGTCTGCTGGCTGGCGAGAACATG
ACCGGTGCTGCTAAGCCCTACTGGCACCAGTCCATGTTCTGGTCCGACCTGGGTCCCGACGTGGGTACG
AGGCTATCGGCCTGGTGGACTCCTCCCTGCCACCGTGGGAGTGTTTCGCTAAGGCTACCGCTCAGGACAA
CCCCAAGTCCGCTACCGAGCAGTCCGGCACCAGTATCCGTTCCGAGTCCGAGACTGAGTCCGAGGCTTCC
GAGATCACCATCCCCCCTCCACCCCGCTGTGCCTCAAGCTCCTGTGCAGGGCGAGGACTACGGCAAGG
GTGTCATCTTCTACCTGCGTGACAAGGTGGTCTGTTGATCTGCTGTGGAACATCTTCAACCGTATGCCTA
TCGCCCAGAGATCATCAAGGACGGCGAGCAGCAGGACCTGAACGAGGTGGCCAAGCTGTTCAACAT
CCACGAGGAC

Protein Sequence (Seq ID No. 35):

>sp|O95831|AIFM1_HUMAN Apoptosis-inducing factor 1, mitochondrial OS=Homo sapiens OX=9606
GN=AIFM1 PE=1 SV=1

MFRCGGLAAGALKQKLVLVPLVTVCSRPRQRNRLPGNLFQRWHVPLELQMTRQMASSGASGGKIDNSVLVLIV
GLSTVGAGAYAYKTMKEDEKRYNERISGLGLTPEQKQKKAALSASEGEEVPQDKAPSHVPFLLLIGGGTAFAAAA
RSIRARDPGARVLIVSEDPPELPMRPPLSKELWFSDDPNVTKTLRFKQWNGKERSIYFQPPSFYVSAQDLPHIEN
GGVAVLTGKKVVQLDVRDNMVKLNDGSQITYEKLIATGGTPRSLAIDRAGAEVKSRTTLFRKIGDFRSLKISR
EVKSITIIGGFLGSELACALGRKARALGTEVIQLFPEKGNMGKILPEYLSNWTMEKVRREGVKVMPNAIVQSVGV
SSGKLLIKLDGRKVETDHIVAAVGLPNVELAKTGGLIDSDFGGFRVNAELQARSNIWVAGDAACFYDIKLGRR
RVEHHDHAVVSGRLAGENMTGAAKPYWHQSMFWSLDLGPVGYEAGLVDSLPTVGVFAKATAQDNPKSATE
QSGTGIRSESETESEASEITIPPSTPAVPQAPVQGEDYGKGVIFYLRDKVVVGIVLWNIFNRMPIARKIHKDGEQHE
DLNEVAKLFNIHED

ODC1 P11926 Ornithine decarboxylase

Nucleotide Sequence (Seq ID No. 15):

>P000568_SIG_SIG1-3_ODC1_4953_Homo sapiens ornithine decarboxylase
1_BC025296.1_AAH25296.1_P11926_62117_0_1386_0_1383

ATGAACAACTTTGGTAATGAAGAGTTTGACTGCCACTTCTCGATGAAGGTTTTACTGCCAAGGACATTCTG
GACCAGAAAATTAATGAAGTTTCTTCTCTGATGATAAGGATGCCTTCTATGTGGCAGACCTGGGAGACATT
CTAAAGAAACATCTGAGGTGGTTAAAAGCTCTCCCTCGTGTACCCCCCTTTATGCAGTCAAATGTAATGATA
GCAAAGCCATCGTGAAGACCCTTGCTGCTACCGGGACAGGATTTGACTGTGCTAGCAAGACTGAAATACAG
TTGGTGCAGAGTCTGGGGGTGCCTCCAGAGAGGATTATCTATGCAAATCCTTGTAACAAGTATCTCAAATT
AAGTATGCTGCTAATAATGGAGTCCAGATGATGACTTTTGATAGTGAAGTTGAGTTGATGAAAGTTGCCAGA
GCACATCCCAAAGCAAAGTTGGTTTTGCGGATTGCCACTGATGATTCAAAGCAGTCTGTGCTCTCAGTGTG
AAATTCGGTGCCACGCTCAGAACCAGCAGGCTCCTTTTGGAAACGGGCGAAAGAGCTAAATATCGATGTTGT
TGGTGTGAGCTTCCATGTAGGAAGCGGCTGTACCGATCTGAGACCTTCGTGCAGGCAATCTCTGATGCCC
GCTGTGTTTTTACATGGGGGCTGAGGTTGGTTTCAGCATGTATCTGCTTGATATTGGCGGTGGCTTTCTCTG
GATCTGAGGATGTGAACTTAAATTTGAAGAGATCACCGGCGTAATCAACCCAGCGTTGGACAAATACTTTC
CGTCAGACTCTGGAGTGAGAATCATAGCTGAGCCCGGCAGATACTATGTTGCATCAGCTTTCACGCTTGCA
GTTAATATCATTGCCAAGAAAATTGTATTAAGGAACAGACGGGCTCTGATGACGAAGATGAGTCGAGTGAG
CAGACCTTTATGTATTATGTGAATGATGGCGTCTATGGATCATTAAATTGCATACTCTATGACCACGCACATG
TAAAGCCCCCTTCTGCAAAAGAGACCTAAACCAGATGAGAAGTATTATTCATCCAGCATATGGGGACCAACAT
GTGATGGCCTCGATCGGATTGTTGAGCGCTGTGACCTGCCTGAAATGCATGTGGGTGATTGGATGCTCTTT
GAAAACATGGGCGCTTACACTGTTGCTGCTGCCTCTACGTTCAATGGCTTCCAGAGGCCGACGATCTACTAT
GTGATGTCAGGGCCTGCGTGGCAACTCATGCAGCAATTCCAGAACCCCGACTTCCCACCCGAAGTAGAGGA
ACAGGATGCCAGCACCTGCCTGTGCTTGTGCCTGGGAGAGTGGGATGAAACGCCACAGAGCAGCCTGT
GCTTCGGCTAGTATTAATGTG

Protein Sequence (Seq ID No. 36):

>sp|P11926|DCOR_HUMAN Ornithine decarboxylase OS=Homo sapiens OX=9606 GN=ODC1 PE=1 SV=2

MNFGNEEFDCHFLDEGFTAKDILDQKINEVSSDDKDAFYVADLGDILKKHLRWLALPRVTPFYAVKCNDSKAI
VKTAAATGTGFDCASKTEIQLVQSLGVPPERIIYANPCKQVSQIKYAANNGVQMMTFDSEVELMKVARAHPKAKL
VLRIATDDSKAVCRSLSVKFGATLRTRSLLLERAKELNIDVVGVSVFHVSGCTDPETFVQAIKDARCVFDMGAEVG
FSMYLLDIGGGFPGSESVKLKFEETGVINPALDKYFSDSGVRIIAEPGRYYVASAFTLAVNIIAKKIVLKEQTGSD
DEDESSEQTFMYVNDGVYGSFNCILYDHAHVKPLLQKRPKPDEKYYSSSIWGPTCDGLDRIVERCDLPEMHVG
DWMLFENMGAYTVAAASTFNGFQRPTIYYVMSGPAWQLMQFQNPDPPEVEEQDASTLPVSCAWESGMKR
HRAACASASINV

RPS6KA4

O75676

Ribosomal protein S6 kinase alpha-4

Nucleotide Sequence (Seq ID No. 16):

>P001321_Q305_Q305p4_RPS6KA4_8986_Homo sapiens RPS6KA4 ribosomal protein S6 kinase, 90kDa,
polypeptide 4_BC047896_AAH47896_O75676_0_0_1575_0_1572

ATGGGGGACGAGGACGACGATGAGAGCTGCGCGTGGAGCTGCGGATCACAGAAGCCAACCTGACCGGG
CACGAGGAGAAGGTGAGCGTGAGAACTTCGAGCTGCTCAAGGTGCTGGGCACGGGAGCCTACGGCAAG
GTGTTCTGCTGCGGAAGGCGGGCGGGCACGACGCGGGGAAGCTGTACGCCATGAAGGTGCTGCGCAAG
GCGGCGCTGGTGACGCGCGCAAGACGCAGGAGCACACGCGCACCGAGCGCTCGGTGCTGGAGCTGGT
GCGCCAGGCGCCCTTCTGGTCACGCTGCACTACGCTTTCAGACGGATGCCAAGCTGCACCTCATCTG
GACTATGTGAGCGGCGGGGAGATGTTACCCACCTCTACCAGCGCCAGTACTTCAAGGAGGTGAGGTGC
GCGTGTATGGGGGTGAGATCGTCTGGCCCTGGAACACCTGCACAAGCTCGGCATCATTTACCGAGACCT
GAACTGGAGAATGTGCTGCTGGACTCCGAGGGCCACATTGTCCTCACGGACTTCGGGCTGAGCAAGGAG
TTCCTGACGGAGGAGAAAGAGCGGACCTTCTCCTTCTGTGGCACCATCGAGTACATGGCCCCGAAATCAT
CCGTAGCAAGACGGGGCATGGCAAGGCTGTGGACTGGTGGAGCCTGGGCATCTTGCTCTTCGAGCTGCTG
ACGGGGGCTCGCCCTTACCCTGGAGGGCGAGAGGAACACGCAGGCTGAGGTGTCTCGACGGATCCTG
AAGTGCTCCCCTCCCTTCCCCCTCGGATCGGGCCCGTGGCGCAGGACCTGCTGCAGCGGCTGCTTTGTA
AGGATCCTAAGAAGCGATTGGGCGCGGGGCCCGAGGGGACACAAGAAGTCCGGAACCATCCCTTCTTCCA
GGGCTCGATTGGGTGGCTCTGGCTGCCAGGAAGATTCCAGCCCCATTCCGGCCCCAATCCGCTCAGAG
CTGGATGTGGGCAACTTTGCGGAGGAATTCACCTCGGCTGGAGCCTGTCTACTCACCCCCCTGGCAGCCCC
CACCTGGGGACCCCCGAATCTTTCAGGGATACTCCTTTGTGGCACCCCTCCATTCTCTTTGACCACAACAACG
CGGCCGAGATCATGTGCAAAATCCGCGAGGGGCGCTTCTCCCTTGACGGGGAGGCTGGCAGGGGTGTATC
CGAGGAAGCCAAGGAGCTGGTCCGAGGGCTCCTGACCGTGGACCCCGCCAAGCGGCTGAAGCTCGAGGG
ACTGCGGGGACAGCTCGTGGCTGCAGGACGGCAGCGCGCGCTCCTCGCCCCGCTCCGGACGCCCGACGT
GCTCGAGTCTCTGGGCCCCGAGTGCCTCGGGTCTCAACGCCACCTTCATGGCATTCAACCGGGGCAAG
CGGGAGGGCTTCTTCTGAAGAGCGTGGAGAACGCACCCCTGGCCAAGCGGCGGAAGCAGAAGCTGCGG
AGCGCCACCGCCTCCCGCGGGGCTCCCCTGCACAGCCAACCCGGGCGGAGCCCCCGTCCGCCCAAA
GGGGCCCCCGCGGAGCCAACGGCCCCCTGCCCCCTCC

Protein Sequence (Seq ID No. 37):

>sp|O75676|KS6A4_HUMAN Ribosomal protein S6 kinase alpha-4 OS=Homo sapiens OX=9606 GN=RPS6KA4
PE=1 SV=1

MGDEDDDESCAVELRITEANLTGHEEKVSVENFELLKVLGTGAYGKVFVLRKAGGHDAGKLYAMKVLKAAALVQ
RAKTQEHTRTERSVLELVRQAPFLVTLHYAFQTDAKLHLILDYVSGGEMFTHLYQRQYFKEAEVRVYGGIEVLAL
EHLHLKLIYRDLKLENVLLDSEGHIVLTDGLSKEFLTEEKERTSFSCGTIEYMAPEIIRSKTGHGKAVDWWSLGIL
LFELLTGASPTLEGERNTQAEVSRRLKCSPPFPFRIGPVAQDLLQRLCKDPKKRLGAGPQGAQEVNHPFFQ
GLDWVALAARKIPAPFRPQIRSELDVGNFAEEFTRLEPVYSPPGSPPGDPRIFQYSFVAPSILFDHNNVMTD
GLEAPGAGDRPGRAAVARSAMMQDSPFFQYELDLREPALGQGSFVCRRCRQRQSGQEFVAKILSRLEAN
TQREVAALRLCQSHPNVNLHEVHHDQLHTYLVLLELLRGELLEHIRKKRHFSESEASQILRSLVSAVSFMHEEA
GVVHRDLKPENILYADDTPGAPVKIIDFGFARLRPQSPGVPMQTPCFTLQYAAPELLAQQGYDESCDLWSLGVIL
YMLLSGQVPFQGASGQGGQSQAAEIMCKIREGRFSLDGEAWQGVSEEAKELVRGLLTVDPKRLKLEGLRGSS
WLQDGSARSSPLRTPDVLESSGPAVRSGLNATFMAFNRGKREGFFLKSVENAPLAKRRKQKLRSATASRRGS
PAPANPGRAPVASKGAPRRANGPLPPS

EEF1D P29692 Elongation factor 1-delta

Nucleotide Sequence (Seq ID No. 17):

>P001467_CAG_CAGp2_EEF1D_1936_Homo sapiens Homo sapiens eukaryotic translation elongation factor 1 delta (guanine nucleotide exchang_BC007847.2_AAH07847.1_P29692_0_0_1944_0_1941

ATGAGGAGCGGGAAGGCCTCCTGCACCCTGGAGACCGTGTGGGAAGACAAGCACAAAGTATGAGGAGGCC
 GAGCGGCGCTTCTACGAACACGAGGCCACACAGGCGGCCCTCCGCCCAGCAGCTGCCAGCCGAGGGG
 CCAGCCATGAATGGGCCCCGCCAGGACGACCCTGAGGACGCTGATGAGGCGGAAGCCCCCTGACGGCGGC
 AGCAGGCGTGATCCCAGGAAGAGCCAGGACAGCAGGAAGCCCCCTGCAGAAAAAGAGGAAGCGCTCCCCC
 AAGAGCGGGCTCGGCCCCCGCGGACCTGGCCCTCCTGGGCCTCTCGGCCGAACGTGTGTGGCTGGACAAG
 TCACTTTTCGACCAGGCAGAGAGCTCCTACCGCCAGAAGCTGGCAGATGTGGCTGCCAGGCAGCCTGGC
 CTCCTGCCTTGGCCCCCTTGGGGTCTCTGCACCCATGGAACACAGGTGGCCTGCCACCACGTGACCTGGGG
 GATCTGGGTCAACAAGTCTCTTCGACCAGGCTGAGCGGGCCTTCGTGGAGTGGTCTCAGGCCCTGTTG
 CTGGCCCCCGAGGGCAGCCGAGGCAGGGGACTCCCAACACAGGCCAGCAGGTGGCCGTCCCCGACCTG
 GCCCACCAGCCCAGCCCACCGGTCAATGGCCAGCCCCCGCTGGGCAGCCTGCAGGCACTGGTTCGGGAG
 GTGTGGCTGGAGAAGCCCCGGTATGATGCAGCCGAGAGGGGCTTCTACGAGGCCCTGTTTGACGGCCATC
 CCCCAGGGAAGGTGCGCCTGCAAGAGCGAGCCGGCCTGGCCGAGGGTGCCCGCGGGGCCGACAGAC
 CGGCGGGGCCGCAACATCTTAGGGAACAAGCGGGCCGGGCTGCGACGGGCCGATGGGGAGGCCCCCTC
 TGCCTTGCCCTACTGTTACTTCTGCAGAAGGATGCAGAGGCCCCCTGGCTCAGCAAGCCTGCCTACGACA
 GCGCCGAGTGCCGCCACCACGCTGCCGAGGCCCTGCGTGTGGCCTGGTGCCTCGAAGCTGCCTCCCTGT
 CTCACCGACCCGGTCTCGGTCTGGCCTGTCCGTGTCCAGCCTGAGACCCAACAGAAAAATGGCTACAAAC
 TTCCTAGCACATGAGAAGATCTGGTTCGACAAGTTCAAATATGACGACGCAGAAAGGAGATTCTACGAGCAG
 ATGAACGGGCTGTGGCAGGTGCCTCCCGCCAGGAGAACGGCGCCAGCGTGATCCTCCGTGACATTGCGA
 GAGCCAGAGAGAACATCCAGAAATCCCTGGCTGGAAGCTCAGGCCCGGGGCTCCAGCGGCACCAGCG
 GAGACCACGGTGAGCTCGTCCGATTGCCAGTCTGGAAGTGGAGAACCAGAGTCTGCGTGGCGTGGT
 ACAGGAGCTGCAGCAGGCCATCTCCAAGCTGGAGGCCCGGCTGAACGTGCTGGAGAAGAGCTCGCCTGG
 CCACCGGGCCACGGCCCCACAGACCCAGCACGTATCTCCCATGCGCCAAGTGGAGCCCCCAGCCAAGAAG
 CCAGCCACACCAGCAGAGGATGACGAGGATGATGACATTGACCTGTTTGGCAGTGACAATGAGGAGGAGG
 ACAAGGAGGCGGCACAGCTGCGGGAGGAGCGGCTACGGCAGTACGCGGAGAAGAAGGCCAAGAAGCCTG
 CACTGGTGGCCAAGTCTCCATCCTGCTGGATGTCAAGCCTTGGGATGATGAGACGGACATGGCCCCAGCT
 GGAGGCCTGTGTGCGCTCTATCCAGCTGGACGGGCTGGTCTGGGGGGCTTCCAAGCTGGTGCCCGTGGG
 CTACGGTATCCGGAAGCTACAGATTCAAGTGTGTGGTGGAGGACGACAAGGTGGGGACAGACTTGCTGGAG
 GAGGAGATCACCAAGTTTGAGGAGCACGTGCAGAGTGTGCATATCGCAGCTTCAACAAGATC

Protein Sequence (Seq ID No. 38):

>sp|P29692|EF1D_HUMAN Elongation factor 1-delta OS=Homo sapiens OX=9606 GN=EEF1D PE=1 SV=5

MATNFLAHEKIWFDFKFKYDDAERRFYEQMNGPVGASRQENGASVILRDIARARENIQKSLAGSSGPGASSGTS
 GDHGELVVRASLEVENQSLRGVVQELQQAISKLEARLNVLEKSSPGHRATAPQTQHVSPMRQVEPPAKKPATP
 AEDDEDIDLFSDNEEDKEAAQLREERLRQYAEKKAKKPALVAKSILLDVKPWDDDETMAQLEACVRSIQL
 DGLVWGASKLVPVGYGIRKLQIQCVVEDDKVGTDLLEEITKFEEHVQSVDIAAFNKI

KLF10 Q13118 Krueppel-like factor 10

Nucleotide Sequence (Seq ID No. 18):

>P000598_TRN_TRNp1_TIEG_7071_Homo sapiens TGFB inducible early growth response_BC011538.1_AAH11538.1_Q53QU8_0_0_1410_0_1407

ATGGAGGAAAGAATGGAAATGATTTCTGAAAGGCCAAAAGAGAGTATGTATTCTGGAACAAAAGTGCAGAG
 AAAAGTGATTTTGAAGCTGTAGAAGCACTTATGTCAATGAGCTGCAGTTGGAAGTCTGATTTTAAAGAAATACG
 TTGAAAACAGACCTGTTACACCAGTATCTGATTTGTGAGAGGAAGAGAATCTGCTTCCGGGAACACCTGATT
 TTCATACAATCCCAGCATTTTGTGTTGACTCCACCTTACAGTCTTCTGACTTTGAACCCTCTCAAGTGTCAAA
 CTGATGGCACCAGCGCCATCTACTGTACACTTCAAGTCACTCTCAGATACTGCCAAACCTCACATTGCCGCA
 CCTTTCAAAGAGGAAGAAAAGAGCCAGTATCTGCCCCAAACTCCCCAAAGCTCAGGCAACAAGTGTGAT
 TCGTCATACAGCTGATGCCAGCTATGTAACCACCAGACCTGCCCAATGAAAGCAGCCAGCATCCTCAACTA
 TCAGAACAATTTCTTTAGAAGAAGAACCCACCTAAATGTTGAGGCTGCAAGAAAGAACATACCATGTGCCGC
 TGTGTCACCAAACAGATCCAAATGTGAGAGAAACACAGTGGCAGATGTTGATGAGAAAGCAAGTGCTGCAC

TTTATGACTTTTCTGTGCCTTCCTCAGAGACGGTCATCTGCAGGTCTCAGCCAGCCCCTGTGTCCCCACAAC
 AGAAGTCAGTGTTGGTCTCTCCACCTGCAGTATCTGCAGGGGAGTGCCACCTATGCCGGTCATCTGCCAG
 ATGGTTCCCTTCCTGCCAACAACCCTGTTGTGACAACAGTCGTTCCAGCACTCCTCCAGCCAGCCACC
 AGCCGTTTGCCCCCTGTTGTGTTTCATGGGCACACAAGTCCCCAAAGGCGCTGTCATGTTTGTGGTACCCC
 AGCCCGTTGTGCAGAGTTCAAAGCCTCCGGTGGTGAGCCGAATGGCACCAGACTCTCTCCCATGCCCCCT
 GCTCCTGGGTTTTCCCTTCAGCAGCAAAAGTCACTCCTCAGATTGATTCATCAAGGATAAGGAGTCACATC
 TGTAGCCACCCAGGATGTGGCAAGACATACTTTAAAGTTCCTCATCTGAAGGCCACACGAGGACGCACAC
 AGGAGAAAAGCCTTTCAGCTGTAGCTGGAAAGGTTGTGAAAGGAGGTTGCCCCGTTCTGATGAACTGTCCA
 GACACAGGCGAACCACACGGGTGAGAAGAAATTTGCGTGCCCCATGTGTGACCGGCGGTTTCATGAGGAG
 TGACCATTTGACCAAGCATGCCCGGCGCCATCTATCAGCCAAGAAGCTACCAAAGTGGCAGATGGAAGTGA
 GCAAGCTAAATGACATTGCTCTACCTCCAACCCCTGCTCCACACAG

Protein Sequence (Seq ID No. 39):

>sp|Q13118|KLF10_HUMAN Krueppel-like factor 10 OS=Homo sapiens OX=9606 GN=KLF10 PE=1 SV=1

MLNFGASLQQTAEERMEMISERPKESMYSWNKTAEKSDFEAVEALMSMSCSWKSDFKKYVENRPVTPVSDLSE
 EENLLPGTPDFHTIPAFCLTPPYSPDFEPSQVSNLMAPASTVHFKSLSDTAKPHIAAPFKEEEKSPVSAPKLPK
 AQATSVIRHTADAQLCNHQTCPMKAASILNYQNNFRRRTHLNVEAARKNIPCAAVSPNRSKERNVADVDEK
 ASAALYDFSVPSETVICRSQPAPVSPQQKSVLVSPPAVSAGGVPPMPVICQMVPLANNPVVTTVPSTPPSQ
 PPAVCPVVMFGTQVPKGAVMFFVPQPVVQSSKPPVSPNGTRLSPAPGFPSPAANKVTPQIDSSRIRSHICS
 HPGCGKTYFKSSHLKAHTRHTGEKPFSCSWKGCERRFARSDELRSRHRHTHTGEKKFACPMCDRRFMRSDHL
 TKHARRHLSAKKLPNWQMEVSKLNDIALPPTAPTQ

EPHA2 P29317 Ephrin type-A receptor 2

Nucleotide Sequence (Seq ID No. 19):

>P003284_Q311_Q311_tube_EPHA2_1969_0_NM_004431.3_0_P29317_0

ATGGAAGTGCAGGCTGCTCGTGCTTGCCTCTGCTGTGGGGTTGCGCTTGGCTGCTGCAGCTGCTGC
 TCAGGGCAAAGAGGTGGTCTGCTGGACTTCGCTGCTGCCGGTGGCGAACTGGGATGGCTGACTACCCCT
 TACGGCAAGGGCTGGGACCTGATGCAGAACATCATGAACGACATGCCATCTACATGTACTCCGTGTGCAA
 CGTGATGTCCGGCGACCAAGGACAAGTGGCTGCGTACCAACTGGGTGTACCGTGGCGAGGCTGAGCGCATC
 TTCATCGAGCTGAAGTTCACCGTGCGCGACTGCAACTCCTTCCCTGGTGGTGCTTCCAGCTGCAAAGAGAC
 TTTCAACCTGTACTACGCTGAGTCCGACCTGGACTACGGCACCAACTTCCAGAAGCGTCTGTTACCAAGAT
 CGACACTATCGCTCCCGACGAGATCACCGTGTCTCCGACTTCGAGGCTCGTCACGTGAAGCTGAACGTGC
 AGGAACGCTCCGTGGGTCCCCTGACCCGCAAGGGATTCTACCTGGCTTTCAGGACATCGGTGCTTGGCT
 GGCCCTGCTGTCCGTGCGTGTGTAACAAGAAGTGCCCCGAGCTGCTCCAGGGCCTGGCTCACTCCCT
 GAGACTATCGCTGGTTCGACGCTCCCTCCCTGGCTACCGTTGCTGGTACTTGGCTGGACACGCTGTGGT
 GCCACCTGGTGGCGAGGAACCTCGTATGCACTGCGCTGTGGACGGCGAGTGGCTGGTGCCTATCGGTCAA
 TGCCTGTGCCAGGCTGGTTACGAGAAGGTCGAGGACGCTTGCCAGGCTTGCTCCCCGGTTTCTTCAAGTT
 CGAGGCTCCGAGTCCCCCTGCCTGGAATGCCCTGAACACACCTGCCTTCCCGAGGGGTGCTACCTCCT
 GCGAGTGCGAAGAGGGCTTCTTCCGTGCTCCCGAGACCCCGCTTCTATGCCTTGACCCCGTCTCCCTCC
 GCTCCCCACTACCTGACTGCTGTGCGCATGGGTGCTAAGGTCGAGCTGCGTTGGACCCCCCTCAGGATT
 CTGGTGGTGGCGAGGACATCGTCTACTCCGTGACCTGCGAGCAGTGTGGCCTGAGTCTGGCGAATGCGG
 TCCCTGCGAGGCTTCTGTGCGCTACTCTGAGCCTCCTACGGCCTGACCCGTACCTCTGTGACCGTGTCCG
 ACCTCGAGCCCCACATGAAGTACACCTTACCGTTCGAGGCCCGTAACGGTGTCTCCGGACTGGTCACTCC
 CGTTCCTTCCGTACCGCTTCCGTGTCCATCAACGAGCCGAGCCCCCAAAGTGGCGCTGGAAGGACGTTT
 TACCACCTCCCTGTCGTGCTTGGTCCATCCCCCACCTCAGCAGTCCCGTGTGTGGAAGTACGAAGTGA
 CCTACCGCAAGAAGGGCGACTCCAACCTTACAACGTGCGTGTACCGAGGGTTTCAGCGTGACCTGGAC
 GACCTGGCTCCCGACACCACTACCTGGTGAAGTGCAGGCTCTGACCCAAGAGGGCCAGGGTGTGGTT
 CCAAGGTGCACGAGTTCAGACCCTGTCCCCGAGGGTCCGGAACTTGGCTGTGATCGGCGGTGTGCG
 TGTGGGTGTGCTGCTGCTGTTGGTGTGGCTGGTGTGCGGCTTCTTATCCACCGTGTGCGAAGAACCAGC
 GTGCTCGTCAGTCCCCTGAGGACGTGTAATCTCCAAGTCCGAGCAGCTGAAGCCCCCTCAAGACCTACGTG
 GACCCCCACACTTACGAGGACCCCAACCAGGCTGTGCTCAAGTTCACTACCGAGATCCACCCCTCCTGCGT
 GACCCGTGAGAAAGTGATCGGTGCTGGCGAGTTCGGCGAGGTGTACAAGGGAATGCTCAAGACTTCCTCC
 GGCAAGAAAGAGGTGCCCGTGTGCTATCAAGACCCTGAAGGCTGGCTACACCGAGAAGCAGCGTGTGGACT
 TCCTGGGAGAGGCTGGTATCATGGGCCAGTCTCCACCACAACATCATCCGTCTGGAAGGTGTATCTCC
 AAGTACAAGCCCATGATGATCATTACCGAGTACATGAAAACGGCGCCCTGGACAAGTTCTGCGCGAGAA

GGATGGCGAGTTCTCCGTGCTGCAGCTCGTGGGAATGCTGCGTGGTATCGCTGCTGGCATGAAGTACCTG
 GCCAACATGAATTACGTGCACAGGGACCTGGCTGCTCGCAACATCCTGGTCAACTCCAACCTCGTGTGCAA
 GGTGTCAGACTTCGGCCTGTCCCGCGTGCTCGAGGACGATCCTGAGGCTACCTACACCACCTCCGGTGGA
 AAGATCCCCATCCGTTGGACCGCTCCCGAGGCTATCTCTTACCGCAAGTTCACCTCCGCTTCCGACGTGTG
 GTCCTTCGGTATCGTGATGTGGGAAGTGATGACCTACGGCGAGCGCCCTACTGGGAGCTGTCTAACCAC
 GAAGTCATGAAGGCTATCAACGACGGTTTCCGTCTGCCACCCCTATGGACTGCCCCCTCCGCTATCTACCA
 GCTGATGATGCAATGCTGGCAGCAAGAGCGTGCTAGGCGTCCCAAGTTCGCTGACATCGTGTCTATCCTCG
 ACAAGCTGATCCGCGCTCCTGACTCCCTGAAAACCTGGCTGACTTCGACCCCGTGTGTCCATCCGCCTG
 CCTTCTACCTCTGGCTCCGAGGGTGTCCCTTTCCGTACTGTGTCCGAGTGGCTCGAGTCCATCAAGATGCA
 GCAGTACACCGAGCACTTCATGGCTGCTGGTTACACCGCTATCGAGAAGGTGGTGCAGATGACCAACGACG
 ACATCAAGCGTATCGGCGTGCGTCTGCCCGTCAACCAGAAGAGGATCGCTTACTCCCTGCTGGGCCTGAA
 GGACCAAGTGAACACCGTGGGTATCCCCATC

Protein Sequence (Seq ID No. 40):

>sp|P29317|EPHA2_HUMAN Ephrin type-A receptor 2 OS=Homo sapiens OX=9606 GN=EPHA2 PE=1 SV=2

MELQAARACFALLWGICALAAAAAQKKEVLLDFAAAGGELGWLTHPYGKGWDLQMNMNDMPIYMYSVCNV
 MSGDQDNWLRNWWYRGEAERIFIELKFTVRDCNSFPGGASSCKETFNLYAESLDYGTNFQKRLFTKIDTIAP
 DEITVSSDFEARHVKLNVEERSVGPLTRKGFYLAQFDIGACVALLSVRVYKKCPPELLQGLAHFPETIAGSDAPSL
 ATVAGTCVDHAVVPPGGEEPMMHCAVDGEWLVPICQCLQAGYEKVEDACQACSPGFFKFEASESPCLECPE
 HTLPSPEGATSCECEEGFFRAPQDPASMPCTRPPSAPHYLAVGMGAKVELRWTPPDQSGGREDIVYSVTCEQ
 CWPESEGCGPCEASVRYSEPPHGLTRTSVTVDLEPHMNYTFTVEARNVSGSLVTSRSFRTASVSINQTEPPKV
 RLEGRSTTSLSVWSIPPPQQSRVWKYEVTYRKKGDSNSYNVRRTEGFSVTLDDLAPDTTYLVQVQALTQEGQ
 GAGSKVHEFQTLSPGSGNLAVIGGVAVGVLLVLAGVGFFIHRRRKNQRARQSPEDVYFSKSEQLKPLKTYV
 DPHTYEDPNQAVLKFTTEIHPSCVTRQKVIGAGEFGEVYKMLKTSSGKKEVPVAIKTLKAGYTEKQRVDFLGEA
 GIMGQFSHHNIIRLEGVISKYKPMIITEYMENGALDKFLREKDGESVLQLVGMRLGIAAGMKYLANMNYVHRDL
 AARNILVNSNLVCKVSDFGLSRVLEDDPEATYTTSGGKIPIRWTAPEAISYRKFTSASDVWSFGIVMWEVMTYGE
 RPYWELSNHEVMKAINDGFRLPMDCPSAIYQLMMQCWQERARRPKFADIVSILDKLIRAPDSLKTADFDPR
 VSIRLPSTSGSEGVFPRTVSEWLESIMQQYTEHFMAAGYTAIEKVVQMTNDDIKRIGVRLPGHQKRIAYSLLGLK
 DQVNTVGIPI

PRKAR1A P10644 cAMP-dependent protein kinase type I-alpha regulatory subunit

Nucleotide Sequence (Seq ID No. 20):

>P000113_CAN_CAN1-1_PRKAR1A_5573_Homo sapiens protein kinase cAMP-dependent regulatory type I
 alpha (tissue specific e_BC036285.1_AAH36285.1_P10644_0_0_1146_0_1143

ATGGAGTCTGGCAGTACCGCCGCCAGTGAGGAGGCACGCAGCCTTCGAGAATGTGAGCTCTACGTCCAGA
 AGCATAACATTCAAGCGCTGCTCAAAGATTCTATTGTGCAGTTGTGCACTGCTCGACCTGAGAGACCCATGG
 CATTCTCAGGGAATACTTTGAGAGGTTGGAGAAGGAGGCAAAACAGATTCAGAATCTGCAGAAAAGCA
 GGCACCTCGTACAGACTCAAGGGAGGATGAGATTTCTCCTCCTCCACCCAACCCAGTGGTTAAAGGTAGGAG
 GCGACGAGGTGCTATCAGCGCTGAGGTCTACACGGAGGAAGATGCGGCATCCTATGTTAGAAAGGTTATAC
 CAAAAGATTACAAGACAATGGCCGCTTTAGCCAAAGCCATTGAAAAGAATGTGCTGTTTTACATCTTGATGA
 TAATGAGAGAAGTGATATTTTTGATGCCATGTTTTCGGTCTCCTTTATCGCAGGAGAGACTGTGATTGAGCAA
 GGTGATGAAGGGGATAACTTCTATGTGATTGATCAAGGAGAGACGGATGTCTATGTTAACAATGAATGGGCA
 ACCAGTGTTGGGGAAGGAGGGAGCTTTGGAGAATTGCTTTGATTTATGGAACACCGAGAGCAGCCACTGT
 CAAAGCAAAGACAAATGTGAAATTGTGGGGCATCGACCGAGACAGCTATAGAAGAATCCTCATGGGAAGCA
 CACTGAGAAAAGCGGAAGATGTATGAGGAATTCCTTAGTAAAGTCTCTATTTTAGAGTCTCTGGACAAGTGGG
 AACGTCTTACGGTAGCTGATGCATTGGAACAGTGCACTTTGAAGATGGGCAGAAGATTGTGGTGCAGGGA
 GAACCAGGGGATGAGTTCTTCATTATTTAGAGGGGTGAGCTGCTGTGCTACAACGTGGTCAGAAAATGAA
 GAGTTTGTGAAAGTGGGAAGATTGGGGCCTTCTGATTATTTGGTGAATTGCACTACTGATGAATCGTCCT
 CGTGCTGCCACAGTTGTTGCTCGTGCCCTTGAAGTGCGTTAAGCTGGACCGACCTAGATTTGAACGTGT
 TCTTGGCCCATGCTCAGACATCCTCAAACGAAACATCCAGCAGTACAACAGTTTTGTGTCACTGTCTGTC

Protein Sequence (Seq ID No. 41):

>sp|P10644|KAP0_HUMAN cAMP-dependent protein kinase type I-alpha regulatory subunit OS=Homo sapiens
 OX=9606 GN=PRKAR1A PE=1 SV=1

MESGSTAASEEARSLRECELYVQKHNIQALLKDSIVQLCTARPERPMAFLREYFERLEKEEAKQIQNLQKAGTRT
 DSREDEISPPPNPVVKGRRRRGAI SAEVYTEEDAASYVRKVIPKDYKTMAALAKAIEKNVLFSLDDNERSDIFD
 AMFSVSFIAGETVIQQGDEGDNFYVIDQGETDVYVNNNEWATSVGEGGSFGELALIYGTTPRAATVKAKTNVKLWGI
 DRDSYRRILMGSTLRKRKMYEEFLSKVSILESLDKWERLTVADALEPVQFEDGGQKIVVQGEPPGDEFFIILEGSAAV
 LQRRSENEEFVEVGRLGPSDYFGEIALLMNRPRRAATVVARGPLKCVKLDPRPRFERVLGPCSDILKRNIQQYNFSV
 SLSV

EAPP Q56P03 E2F-associated phosphoprotein

Nucleotide Sequence (Seq ID No. 21):

>P001616_Q106_Q106p2_EAPP_55837_Homo sapiens chromosome 14 open reading frame
 11_BC001245.1_AA01245.1_xx_0_0_858_0_855

ATGAACCGGCTTCCGGATGACTACGACCCCTACGCGGTTGAAGAGCCTAGCGACGAGGAGCCGGCTTTGA
 GCAGCTCTGAGGATGAAGTGGATGTGCTTTTACATGGAATCCTGACCAAAAACGAAAACATCAGAGAAT
 GTCTTACCGGAGAAAAGTGAATCATCTAGTGAAGATGAATTTGAAAAGGAGATGGAAGCTGAATTAATTCTA
 CCATGAAAACAATGGAGGACAAGTTATCCTCTCTGGGAACTGGATCTTCCTCAGGAAATGAAAAGTTGCAA
 CAGTCCGACAAGGTACTACGATGATATATATTTTGATTCTGATTCCGAGGATGAAGACAGAGCAGTACAGG
 TGACCAAGAAAAAAGAAGAAACAACAAGATTCCAACAAATGACGAATTACTGTATGATCCTGAAAAAGA
 TAACAGAGATCAGGCCTGGGTTGATGCACAGAGAAGGGGTTACCATGGTTTGGGACCACAGAGATCACGTG
 AACACAGCCTGTTCCAAATAGTGATGCTGTCTTGAATTGTCCTGCCTGCATGACCACACTTTGCCTTGATTG
 CCAAAGGCATGAATCATACAAAACCTCAATATAGAGCAATGTTTGAATGAATTGTTCTATTAACAAAGAGGAG
 GTTCTAAGATATAAAGCCTCAGAGAACAGGAAGAAAAGGCGGTCCATAAGAAGATGAGGTCTAACCAGGA
 AGATGCTGCTGAGAAGGCAGAGACAGATGTGGAAGAAATCTATCACCCAGTCATGTGCACTGAATGTTCCA
 CTGAAGTGGCAGTCTACGACAAGGATGAAGTCTTTCATTTTTTCAATGTTTTAGCAAGCCATTCC

Protein Sequence (Seq ID No. 42):

>sp|Q56P03|EAPP_HUMAN E2F-associated phosphoprotein OS=Homo sapiens OX=9606 GN=EAPP PE=1
 SV=4

MNRLPDDYDPYAVEEPSDEEPALSSSEDEVLLHGTPDQKRKLIRECLTGESESSSEDEFKEMEAELNSTMK
 TMEDKLSSLGTGSSSGNGKVATAPTRYDDIYFDSSEDEDRAVQVTKKKKKKQHKIPTNDELLYDPEKDNRDQ
 AWVDAQRRGYHGLGPQRSRQQPVPNSDAVLNCPACMTTLCLDCQRHESYKTQYRAMFVMNCSINKEEVLRY
 KASENRKKRRVHKKMRSNREDAAEAETDVEEIYHPVMCTECSTEVAVYDKDEVFHFNNVLASHS

Claims

1. A method for predicting a response to adalimumab from a sample extracted from a rheumatoid arthritis patient prior to treating the patient with adalimumab, said
5 response being classified as a good response corresponding to anti-drug antibody negative or a poor response corresponding to anti-drug antibody positive, comprising the steps of:
 - (i) testing the sample for the presence of autoantibody biomarkers; and
 - (ii) determining whether the patient will develop a good response or a bad
10 response to treatment with adalimumab, based on the detection of said autoantibody biomarkers;characterised in that said autoantibody biomarkers comprise autoantibodies to antigens SSB, TROVE2 and ZHX2, wherein ZHX2 is associated with the good response, and SSB and TROVE2 are associated with the poor response.
15
2. The method according to claim 1 wherein the autoantibody biomarkers further comprise autoantibodies to one or more antigens from the group comprising of PPARD, SPANXN2, HNRNPA2B, TRIB2, CEP55, SH3GL1, FN3K, PANK3, HPCAL1, THRA, AIFM1, ODC1, RPS6KA4, EEF1D, KLF10, EPHA2,
20 PRKAR1A and EAPP.
3. The method according to claim 1 or 2 wherein the antigens are biotinylated proteins.

4. The method according to claim 3 wherein each biotinylated protein is formed from a Biotin Carboxyl Carrier Protein folding marker which is fused in-frame with a protein.
- 5 5. The method according to claim 3 or 4 wherein the biotinylated proteins are bound to a streptavidin-coated substrate.
6. The method according to claim 5 wherein the substrate comprises a hydrogel-forming polymer base layer.
- 10 7. The method according to any preceding claim wherein the antigens are exposed to a sample extracted from a person, such that autoantibody biomarkers from the sample may bind to the antigens.
- 15 8. The method according to claim 7 wherein the antigens are subsequently exposed to a fluorescently-tagged secondary antibody to allow the amount of any autoantibodies from the sample bound to the antigens to be determined.
- 20 9. The method according to claim 8 wherein the patient's response to treatment with adalimumab corresponds to the relative or absolute amount of autoantibodies from the sample specifically binding to the antigens.

10. The method according to any preceding claim wherein the sample comprises any or any combination of exosomes, blood, serum, plasma, urine, saliva, amniotic fluid, cerebrospinal fluid, breast milk, semen or bile.
- 5 11. The method according to any preceding claim wherein the steps are performed *in vitro*.
12. The method according to any preceding claim wherein the method comprises detecting upregulation/downregulation of one or more biomarkers.
- 10 13. A method for manufacturing a kit for predicting a response to adalimumab from a sample extracted from a rheumatoid arthritis patient prior to treating the patient with adalimumab, comprising the steps of:
- for each antigen in a panel, cloning a biotin carboxyl carrier protein folding
- 15 marker in-frame with a gene encoding the antigen and expressing the resulting biotinylated antigen;
- binding the biotinylated antigens to addressable locations on one or more streptavidin-coated substrates, thereby forming an antigen array;
- such that the amount of autoantibodies from the sample binding to the
- 20 antigens on the panel can be determined by exposing the substrate to the sample and measuring the response;
- characterised in that the antigens comprise SSB, TROVE2 and ZHX2.

14. The method according to claim 13 wherein the antigens further comprise one or more of PPARD, SPANXN2, HNRNPA2B, TRIB2, CEP55, SH3GL1, FN3K, PANK3, HPCAL1, THRA, AIFM1, ODC1, RPS6KA4, EEF1D, KLF10, EPHA2, PRKAR1A and EAPP.
- 5
15. A composition comprising a panel of antigens for predicting an immunogenic and/or therapeutic response to adalimumab in a rheumatoid arthritis patient who has not previously been treated with adalimumab, characterised in that the antigens comprise SSB, TROVE2 and ZHX2.
- 10
16. A composition according to claim 15 wherein the antigens further comprise one or more of PPARD, SPANXN2, HNRNPA2B, TRIB2, CEP55, SH3GL1, FN3K, PANK3, HPCAL1, THRA, AIFM1, ODC1, RPS6KA4, EEF1D, KLF10, EPHA2, PRKAR1A and EAPP.
- 15
17. A composition according to claim 15 or 16 wherein the antigens are biotinylated proteins.
18. A composition according to any of claims 15-17 wherein the amount of one or more autoantibody biomarkers binding *in vitro* to the antigens in a sample from a patient can be measured to predict the response.
- 20
19. A composition comprising a panel of autoantibody biomarkers for predicting an immunogenic and/or therapeutic response to adalimumab in a rheumatoid arthritis

patient who has not previously been treated with adalimumab, wherein the level of the autoantibody biomarkers are measured in a sample collected from the patient;

characterised in that the autoantibody biomarkers are specific to antigens comprising SSB, TROVE2 and ZHX2.