

IDENTIFICATION OF HEALTH STATUS IN THE ELDERLY USING IMMUNOLOGICAL BIOMARKERS.

Field of Invention

- 5 The invention relates to the detection of immunological biomarkers, particularly autoantibodies, to determine the health status and/or aging trajectory in the elderly.

Background

10 Despite technological advances in the area of proteomics research, there are only a handful of biomarkers that have entered the clinic, and 90% of the biomarkers are protein biomarkers [1]. 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). Some of the main reasons for failure of biomarkers [2] to make it into clinical practice are:

- 15 1) Low sensitivity and specificity
 2) Low prognostic/predictive value
 3) Not important for clinical decision making
 4) Original claims fail validation (false discoveries)

20 The management of care of elderly individuals depends less on age than on the effect of their comorbidity history (past and present) on their current health status [3]. These comorbidities impose a certain stress on the immune system which has been challenged over the years to deal with infections, cancer or chronic inflammatory diseases [4].

25 An aim of the invention is therefore to provide an improved panel of autoantibody biomarkers for assessing the health status of elderly individuals.

Summary of Invention

30 In one aspect of the invention, there is provided a method for determining the health of an individual from a sample extracted from that individual, comprising the steps of:

- (i) testing the sample for the presence of biomarkers specific for health;
- (ii) determining whether the subject is healthy, is of intermediate health, or is unhealthy, based on the detection of said biomarkers;

characterised in that the biomarkers are autoantibodies to antigens comprising AURKA, FEN1, GLRX3, PHLDA1, PPM1A, FKBP3, CD96 and MAPK13.

In one embodiment the individual is elderly, typically at least 60 years old.

5

Advantageously the autoantibody biomarkers can be used in the characterization (or diagnosis) of the health status of an elderly individual (Healthy, Intermediate and Unhealthy) by measuring the distribution of plasma-antibody levels. Furthermore a subset of these autoantibody biomarkers, particularly those associated with Healthy and Intermediate, may have a protective role against non-communicable disease.

10

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.

15

In one embodiment the antigens may include one or more from the group comprising of UBE2I, AAK1, YARS, ASPSCR1, CASP10, FHOD2, TCL1A and MAP4.

It should be noted that not all antigens generate an autoantibody response and it is not possible to predict *a priori* which antigens will do so in a given cohort – of more than 1500 antigens tested, only autoantibodies against the 16 antigens described above are suitable as biomarkers to identify health and aging status.

20

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.

25

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

30

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 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.

- 5 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.
- 10 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.

15

- 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 antibody is anti-human IgG, but it will be appreciated that other secondary
- 20 antibodies could be used, such as anti-IgM, anti-IgG1, anti-IgG2, anti-IgG3, anti-IgG4 or anti-IgA.

25

In one embodiment the healthiness of the individual corresponds to the relative or absolute amount of autoantibodies from the sample specifically binding to the antigens.

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

30

In one embodiment the method comprises detecting upregulation/downregulation of one or more biomarkers.

In a further aspect of the invention, there is provided a method for manufacturing a kit for determining the health of an elderly individual from a sample extracted from that individual, 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 said 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 antigens on the panel can be determined by exposing the substrate to the sample and measuring the response;

characterised in that the antigens comprise AURKA, FEN1, GLRX3, PHLDA1, PPM1A, FKBP3, CD96 and MAPK13.

In one embodiment the antigens may include one or more from the group comprising of UBE2I, AAK1, YARS, ASPSCR1, CASP10, FHOD2, TCL1A and MAP4.

In a further aspect of the invention there is provided a method for determining the health of an elderly individual by exposing a composition comprising a panel of antigens as herein described to a sample extracted from that individual, and determining the level of autoantibodies from the sample binding to the antigens.

In a yet further aspect of the invention there is provided a method for determining the health of an elderly individual by exposing a composition comprising a panel of antigens as herein described to a sample extracted from that individual *in vitro*, and determining the level of autoantibodies from the sample binding to the antigens.

In further aspect of the invention, there is provided a composition comprising a panel of antigens for determining the health of an elderly individual, characterised in that the antigens comprise AURKA, FEN1, GLRX3, PHLDA1, PPM1A, FKBP3, CD96 and MAPK13.

In one embodiment the antigens may include one or more from the group comprising of UBE2I, AAK1, YARS, ASPSCR1, CASP10, FHOD2, TCL1A and MAP4.

In one embodiment the antigens are biotinylated proteins

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 health status of the patient.

5

In yet further aspect of the invention, there is provided a composition comprising a panel of autoantibody biomarkers for determining the health status of an elderly patient; wherein the level of one or more autoantibody biomarkers are measured in a sample from the patient;

10

characterised in that the one or more autoantibody biomarkers are selected from autoantibodies specific for one or more of the following antigens: AURKA, FEN1, GLRX3, PHLDA1, PPM1A, FKBP3, CD96 and MAPK13.

Brief Description of Drawings

15

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.

20

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

Figure 2 illustrates the pPRO9 plasmid used as a vector.

25

Figure 3 illustrates proteins associated with cell-cycle and cell-death as (A) a chart; (B) linked pathways.

30

Figure 4 illustrates a clustering analysis: (A) Representation of clusters defined within the elderlies by the 16 antigens by tSNE clustering analysis [5]; (B) Expression density of antibodies for each target protein; (C) Autoantibodies specific to each of the health status groups;.

Figure 5 illustrates the cohort selection: (A) Schematic describing the workflow used to select and categorize elderly individuals in the study; (B) Distribution of elderly and young individuals according to age and gender; Statistical analysis performed with Kruskal-Wallis test with Dunn's correction; (C) Range of clinical variables used for the categorization of elderly individuals; (D) Characteristic of elderly individuals selected in the study for the 6 determining clinical parameters. .

Detailed Description

10 **Materials and Methods**

Gene synthesis and cloning. The pPRO9 plasmid (see Figure 2 below) was constructed by standard techniques and consists of a c-myc tag and BCCP protein domain, preceded by a multi-cloning site. A synthetic gene insert was assembled from synthetic
15 oligonucleotides and/or PCR products. The fragment was cloned into pPRO9 using SpeI and NcoI cloning sites. The plasmid DNA was purified from transformed bacteria and concentration determined by UV spectroscopy. The final construct was verified by sequencing. The sequence congruence within the used restriction sites was 100%. 5µg of the plasmid preparation was lyophilized for storage.

20 The recombinant baculoviruses are generated via co-transfection of a bacmid carrying the strong viral polyhedrin promoter together with a transfer vector carrying the coding sequences of protein of interest, into the Sf9 cell line which is a clonal isolate derived from the parental *Spodoptera frugiperda* cell line IPLB-Sf-21-AE. Homologous
25 recombination initiated by the viral system causes 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 culture medium, and viral amplification were done to generate a higher titre of P1
30 viruses.

Protein Expression. Expressions were 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. 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 protein itself, recombinant protein lysate can be pelleted either from the cultured cell or the cultured medium. Positive recombinant proteins were then analyzed via SDS-PAGE and Western blot against Streptavidin-HRP antibody. In total, 1557 human antigens were cloned and expressed using this methodology.

Array fabrication. HS (hydrogel-streptavidin) slides were purchased from Schott and used to print the biotinylated proteins. 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 between 7.0 and 7.5 were used. The slides were dried by centrifugation ($200 \times g$ 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, each resultant 'Immunome array' comprised 1557 antigens, each printed in quadruplicate.

20

Experimental Procedure. Each critical experimental step of running the Immunome array required a second trained person to thoroughly check, precisely record and cross-check all steps in the protocol, in order to reduce operator bias. Samples were picked, randomised and assigned to assay racks accordingly. These samples were then stored at -20°C until the experimental setup was complete.

25

1. Study cohort

The study cohort was divided into 2 age groups: young control individuals (YC) and the elderly individuals. The YC group ($n=60$) composed of male ($n=34$) and female ($n=26$) individuals of Chinese ethnicity from 18 to 27 years of age. They are clinically healthy with no reported comorbidities nor active medical treatments. The selection of elderly individuals was performed within elderly individuals of Chinese ethnicity of 60 years of

30

age and beyond. This initial selection increases the analytical power and outcome of this study by removing an ethnicity bias.

Further selection of elderly individuals into health classes (Healthy, Intermediate and Unhealthy) was based on the combination of 6 clinical parameters (Figure 5A):

- Comorb5: Variable reflecting the total number of comorbidities excluding eye problems
- NADL: Total number of disabilities affecting Activities of Daily Living of the elderly individual [7]
- Wo_sfl: Parameter measuring the general quality of life.
- Frailty: A clinical syndrome where the elderly individual is progressively highly vulnerable to internal and external stressors. It is a multidimensional variable taking into account the physical strength and cognitive abilities [8]
- MMSEtot: Total score of the Mini-Mental State Examination. This is indicative of the cognitive capabilities of the elderly individual [9].
- GDStot: Total score of the Geriatric Depression Scale. This is a self-reported assessment used to identify the depression in the elderly [10].

The characterization of the health status of the elderly individuals takes into accounts the 6 parameters previously described and resulted in the selection of the following groups (Figure 5A):

- 115 Healthy elderlies,
 - 111 elderlies with Intermediate health status,
 - 114 Unhealthy elderlies.
- There are no significant variations of age between the health groups although a gender difference can be observed as more females are present in each health group (Figure 5B, Figure 5D). In Figure 5D, bold numbers determine the grade of the individuals for the specific category and score. Numbers between brackets correspond to number of individuals with specific traits. ND: Not determined. Ave: Average age of all elderly individuals for each health groups [6].

Overall the repartition of the individuals showed that unhealthy elderly individuals present an accumulation of comorbidities, an increased frailty status and cognitive

decline associated with higher depressive status and an increased quality of life (Figure 5C, Figure 5D).

2. Serum/Plasma Dilution

- 5 Samples were then 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 at full speed and spun down for 3 minutes at 13,000 g using a microcentrifuge. 22.5 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, 10% v/v PBS (20°C) and vortexed to mix three times. The tube
10 was tilted during aspiration to ensure that the sera was sampled from below the lipid layer at the top but does not touch the bottom of the tube in case of presence of any sediment. This Serum/Plasma dilution process was carried out in a class II Biological Safety Cabinet. Batch records were marked accordingly to ensure that the correct samples were added to the correct tubes.

15

3. Biomarker Assay

- The array was removed from the storage buffer using forceps, placed in the slide box and rack containing 200 mL of cold SAB (4°C) and shaken on shaker at 50 rpm, for 5 minutes. When the slides have completed washing, the slide was placed, array side up,
20 in a slide hybridization chamber with individual sera which had been diluted earlier. All slides were scanned using the barcode scanner into the relevant batch record and incubated in a refrigerated shaker at 50 rpm for 2 hours at 20°C.

4. Array Washing After Serum Binding

- 25 The protein array slide was then 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.

30 5. Incubation with Cy3-anti IgG

Binding of autoantibodies to the arrayed antigens on replica Immunome arrays was detected by incubation with Cy3-rabbit anti-human IgG. Arrays were immersed in

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

6. Washing After Incubation with Cy3-anti 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 at 240g at room temperature. Slides were then stored at room temperature until scanning (preferably the same day). Hybridization signals were measured with a microarray laser
- 10 scanner (Agilent Scanner) at 10µm resolution. Fluorescence intensities were detected according to the manufacturer's instructions, whereby each spot is plotted using Agilent Feature Extraction software.

- Spot segmentation** Semi-automatic QC process was carried out in order to produce a
- 15 viable result. The output from the microarray scanner is a raw .tiff format image file. Extraction and quantification of each spot on the array were performed using the GenePix Pro 7 software (Molecular Devices). A GAL (GenePix Array List) file for the array was generated to aid with image analysis. GenePix Pro 7 allows for automatic spot gridding and alignment of each spot on the array for data extraction. Following data
 - 20 extraction, a GenePix Results (.GPR) file was generated for each slide which contains numerical information for each spot; Protein ID, protein name, foreground intensities, background intensities etc.

Bioinformatics analysis.

25

1. Image Analysis: Raw Data Extraction

- 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.
- 30 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 the array. This includes the mean and median of the pixel intensities within a spot along with its local background. A GAL (GenePix Array List) file for the

array is generated to aid with 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 spot; Protein ID, protein name, foreground intensities, background intensities etc. In the
 5 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

For each slide, proteins and control probes are spotted in quadruplicate - 4 arrays on
 10 each slide. The following steps were performed to verify the quality of the protein array data before proceeding with data analysis:

Step 1:

Calculate net intensities for each spot by subtracting background signal intensities from
 15 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:

20 Remove replica spots with $RFU \leq 0$.

Step3:

No saturated pixels should be visible within the spots across array which may exceed scanner's reading capacity (maximum RFU for our scanner is 65536 RFU). Therefore,
 25 spot/s that show saturation in $> 20\%$ of the pixels were removed if it occurs in ≤ 2 replica/s. If saturated spots occur in 3 or more replicas of that protein or probe, these proteins/probes will be flagged as "SAT" and excluded from the downstream analyses.

Step 4:

30 Zero net intensities if only 1 replica spot remaining.

Step 5:

Calculating percentage of coefficient of variant (CV%) of to determine the variations between the replica spots on each slide.

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

- 5 Flag a set of replica spots with only 2 or less replica/s remaining and CV% > 20% as “High CV”. The mean RFU of these replica spots (i.e. proteins) will be excluded from the downstream analysis.

- 10 For proteins/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.

15 Step 6:

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

Step 7:

- 20 Composite normalisation of data using both quantile-based and total intensity-based modules. This method assumes that different samples share a common underlying distribution of their control probes while considering 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.

25

- The quantile module adopts the algorithm described by Bolstad et al., 2003 [11]. 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, 30 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

- 5 *(i = spot number and j = 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 $\langle X_i \rangle$

10 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 ,

$$\text{Scaling factor (k)} = \frac{\sum \langle X_i \rangle}{\sum X_k} \quad \text{Equation 2}$$

15

Data Analysis

The fluorescence signals from the 1557 autoantibody measurements were logarithmically transformed to ensure normality prior to any parametric analysis. One way ANOVA was carried on each of the 1557 autoantibody measurements against i) between all groups (healthy elderlies, elderlies with intermediate health status, unhealthy elderlies and the young controls) and ii) between the elderlies (healthy, intermediate and unhealthy) to identify autoantibodies which were significantly different in at least one of the groups compared to the rest (Table 1). An initial P-value threshold of 0.05 was used to indicate significance. Autoantibody biomarkers towards 16 antigens were identified in this manner: YARS [12], UBE2I [13], TCL1A [14], PPM1A [15], PHLDA1 [16], GLRX3 [17], FHOD2 [18], FEN1 [19], CASP10 [20], MAPK13 [21], MAP4 [22], FKBP3 [23], CD96 [24], AURKA [25], ASPSCR1 [26] and AAK [27], shown in figure 3. Amongst these 16 antigens, AURKA, FEN1, GLRX3, PHLDA1, PPM1A, FKBP3, CD96 and MAPK13 were found to have P-values of < 0.02 in both analyses (between all groups and between elderlies).

Table 1

Biomarker	P-Values	
	between all groups	between elderlies
AURKA	0.000912	1.41E-03
FEN1	0.00772	0.00592
GLRX3	0.00289	0.00633
PHLDA1	0.00476	0.00799
PPM1A	7.61E-05	0.0108
FKBP3	0.00504	0.0114
CD96	0.00593	0.0116
MAPK13	0.0191	0.0165
MAP4	0.0376	0.0172
TCL1A	0.000814	0.0211
FHOD2	1.05E-04	0.0212
CASP10	5.91E-05	0.023
ASPSCR1	0.0317	0.0281
YARS	0.0761	0.045
AAK1	0.0126	0.0453
UBE2I	0.0036	0.0496

Pathway enrichment analysis showed that 4 of the 16 (PHLDA1, AURKA, FEN1 and UBE2I) are involved in Cell Cycle and DNA repair pathways which are altered in the aging process.

Given that each of the 16 individual autoantibodies are weak predictors of the health status on their own, dimension reduction using tSNE was carried out to identify the collective capabilities of the 16 autoantibodies to differentiate the health groups. As seen in the figure 4A, dimension reduction using tSNE show that the 2 tSNE dimensions were able to differentiate the 3 health groups. The specificity of the 16 autoantibodies is shown in figure 4B.

To identify autoantibodies specific to each of the health status, a series of t-tests with Welch correction was used to test each of the health status against the rest for all 16 identified autoantibodies. For each of the autoantibodies, the best t-test result amongst the three health statuses were selected as the autoantibody of choice for that health status. This identified PHLDA1 and CD96 as being specific for the healthy group,

AURKA, FEN1, CASP10 and AAK1 as being specific for the intermediate group and the rest as being specific for the unhealthy group (figure 4C). The healthy and intermediate autoantibodies may have a protective role against non-communicable disease. The mean RFU is shown for each of the health status groups together with the
5 P-values of the t-tests demonstrating the significance of the autoantibody in discriminating the health status group of interest against the rest

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
10 above and in EP1470229. The structure of the *E. coli* BCCP domain is illustrated in Figure 1, 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.

15 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 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
20 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 support via the BCCP tag.

The surface chemistry of the support is designed carefully and may use a three-
25 dimensional thin film polymer base layer (polyethylene glycol; PEG), which retains protein spot morphologies and ensures consistent spot sizes across the array. The PEG layer inhibits non-specific binding, therefore reducing the high background observed using other platforms. The solid support used to immobilize the selected biomarkers is thus designed to resist non-specific macromolecule adsorption and give excellent
30 signal-to-noise ratios and low limits of detection (i.e. improved sensitivity) by minimising non-specific background binding. In addition, the PEG layer also preserves the folded structure and functionality of arrayed proteins and protein complexes post-immobilisation. This is critical for the accurate diagnosis because human serum

antibodies are known in general to bind non-specifically to exposed hydrophobic surfaces on unfolded proteins, thus giving rise to false positives in serological assays on arrays of unfolded proteins, moreover, human autoantibodies typically bind to discontinuous epitopes, so serological assays on arrays of unfolded proteins or mis-
5 folded proteins will also give rise to false negatives in autoantibody binding assays.

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 and is ideal for applications such as protein arrays, SPR and bead-
10 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
15 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 success rate of BCCP folding marker mediated expression of even the most
20 complex proteins is in excess of 98%. The technology can therefore be applied in a highly parallelised pipeline resulting in high-throughput, highly consistent production of functionally validated proteins.

The addition of BCCP permits the monitoring of fusion protein folding by measuring
25 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
30 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.

Figure 3 illustrates discriminating proteins in the elderly, characterized by processes of cell-cycle and cell-death. (A) p-values related to the 16 protein-targets discriminating the various elderly health statuses. Only readouts of serum/plasma antibody to the YARS protein do not also discriminate between the elderly and YC individuals ($p > 0.05$) (B) 5 protein readouts were tightly associated with regards to pathways linked to cell-cycle (PHLDA1, AURKA, FEN1), DNA repair (UBE2I, FEN1) and translation (YARS) by enrichment pathway analysis (String).

References

- [1] Mackay EM, Bathe OF. Identifying Clinically Relevant Proteins for Targeted Analysis in the Development of a Multiplexed Proteomic Biomarker Assay. *Methods Mol Biol.* 2018;1788:123-129.
- [2] Yadav S, Kashaninejad N, Masud MK, Yamauchi Y, Nguyen NT, Shiddiky MJA. Autoantibodies as diagnostic and prognostic cancer biomarker: Detection techniques and approaches. *Biosens Bioelectron.* 2019 May 13;139:111315.
- [3] Makovski TT, Schmitz S, Zeegers MP, Stranges S, van den Akker M. Multimorbidity and quality of life: systematic literature review and meta-analysis. *Ageing Res Rev.* 2019 Apr 29. pii: S1568-1637(19)30006-6.
- [4] Fali T, Vallet H, Sauce D. Impact of stress on aged immune system compartments: Overview from fundamental to clinical data. *Exp Gerontol.* 2018 May;105:19-26. doi: 10.1016/j.exger.2018.02.007.
- [5] Van Der Maaten L., Hinton G. Visualizing data using t-SNE. *Journal of Machine Learning Research.* 2008;9(1):2579–2605
- [6] Valenzuela JF, Monterola C, Tong VJC, Ng TP, Larbi A. Health and disease phenotyping in old age using a cluster network analysis. *Sci Rep.* 2017 Nov 15;7(1):15608.
- [7] Elizabeth A. Phelan Barbara Williams Brenda W. J. H. Penninx James P. LoGerfo Suzanne G. Leveille. Activities of Daily Living Function and Disability in Older Adults in a Randomized Trial of the Health Enhancement Program *The Journals of Gerontology: Series A*, Volume 59, Issue 8, August 2004, Pages M838–M843.

- [8] Fried LP, Ferrucci L, Darer J, Williamson JD, Anderson G. Untangling the concepts of disability, frailty, and comorbidity: implications for improved targeting and care. *J Gerontol A Biol Sci Med Sci*. 2004 Mar;59(3):255-63.
- [9] Crum RM, Anthony JC, Bassett SS, Folstein MF. Population-based norms for the Mini-Mental State Examination by age and educational level. *JAMA*. 1993 May 12;269(18):2386-91.
- [10] Leshner EL, Berryhill JS. Validation of the Geriatric Depression Scale--Short Form among inpatients. *J Clin Psychol*. 1994 Mar;50(2):256-60.
- [11] 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.
- [12] Tracewska-Siemiatkowska A, Haer-Wigman L, Bosch DGM, Nickerson D, Bamshad MJ; University of Washington Center for Mendelian Genomics, van de Vorst M, Rendtorff ND, Möller C, Kjellström U, Andréasson S, Cremers FPM, Tranebjærg L. An Expanded Multi-Organ Disease Phenotype Associated with Mutations in YARS. *Genes (Basel)*. 2017 Dec 11;8(12). pii: E381
- [13] Watanabe TK, Fujiwara T, Kawai A, Shimizu F, Takami S, Hirano H, Okuno S, Ozaki K, Takeda S, Shimada Y, Nagata M, Takaichi A, Takahashi E, Nakamura Y, Shin S. Cloning, expression, and mapping of UBE2I, a novel gene encoding a human homologue of yeast ubiquitin-conjugating enzymes which are critical for regulating the cell cycle. *Cytogenet Cell Genet*. 1996;72(1):86-9.
- [14] Ho MF, Lummertz da Rocha E, Zhang C, Ingle JN, Goss PE, Shepherd LE, Kubo M, Wang L, Li H, Weinshilboum RM. TCL1A, a Novel Transcription Factor and a Coregulator of Nuclear Factor κ B p65: Single Nucleotide Polymorphism and Estrogen Dependence. *J Pharmacol Exp Ther*. 2018 Jun;365(3):700-710.
- [15] Lin X, Duan X, Liang YY, Su Y, Wrighton KH, Long J, Hu M, Davis CM, Wang J, Brunicardi FC, Shi Y, Chen YG, Meng A, Feng XH. PPM1A Functions as a Smad Phosphatase to Terminate TGF β Signaling. *Cell*. 2016 Sep 8;166(6):1597.
- [16] Chen Y, Takikawa M, Tsutsumi S, Yamaguchi Y, Okabe A, Shimada M, Kawase T, Sada A, Ezawa I, Takano Y, Nagata K, Suzuki Y, Semba K, Aburatani H, Ohki R. PHLDA1, another PHLDA family protein that inhibits Akt. *Cancer Sci*. 2018 Nov;109(11):3532-3542.

- [17] Li B1, Chen M1, Lu M1, Xin-Xiang J1, Meng-Xiong P1, Jun-Wu M1. Glutaredoxin 3 promotes migration and invasion via the Notch signalling pathway in oral squamous cell carcinoma. *Free Radic Res.* 2018 Apr;52(4):390-401.
- [18] Otomo T, Tomchick DR, Otomo C, Panchal SC, Machius M, Rosen MK.
5 Structural basis of actin filament nucleation and processive capping by a formin homology 2 domain. *Nature.* 2005 Feb 3;433(7025):488-94.
- [19] Kathera C, Zhang J, Janardhan A, Sun H, Ali W, Zhou X, He L, Guo Z. Interacting partners of FEN1 and its role in the development of anticancer therapeutics. *Oncotarget.* 2017 Apr 18;8(16):27593-27602.
- 10 [20] Horn S, Hughes MA, Schilling R, Sticht C, Tenev T, Ploesser M, Meier P, Sprick MR, MacFarlane M, Leverkus M. Caspase-10 Negatively Regulates Caspase-8-Mediated Cell Death, Switching the Response to CD95L in Favor of NF- κ B Activation and Cell Survival. *Cell Rep.* 2017 Apr 25;19(4):785-797.
- [21] Tan FL, Ooi A, Huang D, Wong JC, Qian CN, Chao C, Ooi L, Tan YM, Chung A,
15 Cheow PC, Zhang Z, Petillo D, Yang XJ, Teh BT. p38delta/MAPK13 as a diagnostic marker for cholangiocarcinoma and its involvement in cell motility and invasion. *Int J Cancer.* 2010 May 15;126(10):2353-61
- [22] Kremer BE, Haystead T, Macara IG. Mammalian septins regulate microtubule stability through interaction with the microtubule-binding protein MAP4. *Mol Biol Cell.*
20 2005 Oct;16(10):4648-59
- [23] Zhu W, Li Z, Xiong L, Yu X, Chen X, Lin Q. FKBP3 Promotes Proliferation of Non-Small Cell Lung Cancer Cells through Regulating Sp1/HDAC2/p27. *Theranostics.* 2017 Jul 22;7(12):3078-3089
- [24] Georgiev H, Ravens I, Papadogianni G, Bernhardt G. Coming of Age: CD96
25 Emerges as Modulator of Immune Responses. *Front Immunol.* 2018 May 17;9:1072.
- [25] Donnella HJ, Webber JT, Levin RS, Camarda R, Momcilovic O, Bayani N, Shah KN, Korkola JE, Shokat KM, Goga A, Gordan JD, Bandyopadhyay S. Kinome rewiring reveals AURKA limits PI3K-pathway inhibitor efficacy in breast cancer. *Nat Chem Biol.* 2018 Aug;14(8):768-777
- 30 [26] Bogan JS1, Hendon N, McKee AE, Tsao TS, Lodish HF. *Nature.* 2003 Oct 16;425(6959):727-33. Functional cloning of TUG as a regulator of GLUT4 glucose transporter trafficking.

[27] Gupta-Rossi N, Ortica S, Meas-Yedid V, Heuss S, Moretti J, Olivo-Marin JC, Israël A. The adaptor-associated kinase 1, AAK1, is a positive regulator of the Notch pathway. *J Biol Chem*. 2011 May 27;286(21):18720-30.

Table 2

Protein Name	UniprotID	Description
AAK1	Q2M2I8	HUMAN AP2-associated protein kinase 1
<i>Nucleotide Sequence (Seq ID No. 1):</i> >P001067_KIN2_KIN2p1_AAK1_22848_Homo sapiens AP2 associated kinase 1_BC002695.2_AAH02695.1_Q2M2I8_0_0_1425_0_1422 ATGAAGAAGTTTTCTGACTCCCGGCGAGAGCAGGGCGGCTCTGGCCTGGGCTCCGGCTCCAGCGGAGGAG GGGGCGACACCTCGGGCCTGGGCAGTGGCTACATCGGAAGAGTCTTCGGCATCGGGCGACAGCAGGTCAC AGTGGACGAGGTGTTGGCGGAAGGTGGATTTGCTATTGTATTCTGGTGAGGACAAGCAATGGGATGAAAT GTGCCTTGAACGCATGTTTGTCAACAATGAGCATGATCTCCAGGTGTGCAAGAGAGAAATCCAGATAATGA GGGATCTTTCAGGGCACAAGAATATTGTGGGTTACATTGATTCTAGTATCAACAACGTGAGTAGCGGTGATGT ATGGGAAGTGCTCATTCTGATGGACTTTTGTAGAGGTGGCCAGGTGTTAAACCTGATGAACCGCGCCTGCA AACAGGCTTTACAGAGAATGAAGTGCTCCAGATATTTTGTGATACCTGTGAAGCTGTTGCCCGCCTGCATCA GTGCAAACTCCTATTATCCACCGGGACCTGAAGGTTGAAAACATCCTCTTGCATGACCGAGGCCACTATGT CCTGTGTGACTTTGGAAGCGCCACCAACAAATTCCAGAATCCACAACTGAGGGAGTCAATGCAGTAGAAGA TGAGATTAAGAAATACACAACGCTGTCCTATCGAGCACCAGAAATGGTCAACCTGTACAGTGGCAAAATCATC ACTACGAAGGCAGACATTTGGGCTCTTGGATGTTTGTGTATAAATTATGCTACTTCACTTTGCCATTTGGGG AAAGTCAGGTGGCAATTTGTATGGAACCTTCAAAATCCTGATAATTCTCGATATTCTCAAGACATGCACTG CCTAATTAGGTATATGTTGGAACAGACCTGACAAAAGGCCGGATATTTACCAGGTGTCTACTTCTCATT AAGCTACTCAAGAAAGAGTGCCCAATCCAAATGTACAGAATCTCCATTCTGCAAAGCTTCTGAACCGAG TGAAAGCCAGTGAGGCAGCTGCAAAAAGACCCAGCCAAAGGCCAGACTGACAGATCCCATTCACCCACACA GAGACTTCAATTGCACCCCGCCAGAGGCCTAAAGCTGGGCAGACTCAGCCGAACCCAGGAATCCTTCCCAT CCAGCCAGCGCTGACACCCCGGAAGAGGGCCACTGTTACAGCCCCACCTCAGGCTGCAGGATCCAGCAAT CAGCCTGGCCTTTTAGCCAGTGTTCCCAACCAAAACCCCAAGCCCCACCCAGCCAGCCTCTGCCGCAAC TCAGGCCAAGCAGCCACAGGCTCCTCCCACTCCACAGCAGACGCTTCTACTCAGGCCCAGGGTCTGCCCG CTCAGGCCCAGGCCACACCCAGCACCAGCAGCATACAATAAACTTAGTATGAACTT <i>Protein Sequence (Seq ID No. 17):</i> >sp Q2M2I8 AAK1_HUMAN AP2-associated protein kinase 1 OS=Homo sapiens OX=9606 GN=AAK1 PE=1 SV=3 MKKFFDSRREQGSSGLSGSSGGGGSTSGLGSYGIRVFGIGRQQVTVDEVLAEAGGFVFLVRTSNGMKCAL KRMFVNNEHDLQVCKREIQIMRDLSGHKNIVGYIDSSINNVSVDVWEVLILMDFCRGGQVNVLMNQRLQTGFTE NEVLQIFCDTCEAVARLHQCKTPPIHRDLKVENILLHDRGHYVLCDFGSATNKFQNPQTEGVNAVEDEIKKYTTLSY RAPQMVNLYSGKIITKADIWALGCLLYKCYFTLPFGESQVACDGNFTIPDNSRYSQDMHCLIRYMLEPDPDKRP DIYQVSYFSFKLLKKECPIPNVQNSPIAKLPEPVKASEAAAKTKQPKARLTDPIPTTETSIAPRQRPKAGQTQPNP GILPIQPALTPRKRATVQPPQAAGSSNQPGLLASVPQPKPQAPPSQPLPQTQAKQPAAPPTPQQTPSTQAQGL PAQAQATPQHQQQLFLKQQQQQQQPPPAQQQPAGTFYQQQQAQTQQFQAVHPATQKPAIAQFPVVSQGGSQ QQLMQNFYQQQQQQQQQQQQQLATALHQQQLMTQQAALQKQPTMAAGQQPQPQAAAPQPAPAEPAIQ APVRQQPKVQTTTPPAVQGGQKVGSLTPSPSKTQRAGHRRILSDVTHSAVFGVPASKSTQLLQAAAAEASLNKS KSATTTSPGSPRTSQQNVNPNPSEGSTWNPFDNDFSKLTAEELLNKDFAKLGEKGHPEKLGGSAESLIPGFQST QGDAFATTSFSAGTAEKRKGGQTVDSGLPLLSVSDPFIPLQVPDAPEKLEGLKSPDTSLLLPDLLPMTDPFGSTS DAVIEKADVAVESLIPGLEPPVPQRLPSQTESVTSNRTDSL TGEDSLLDCSLLSNPTDLEEFAPTAPVHKAA EDSNLISGFDVPEGSDKVAEDEFDPIPVLTITKNPQGGHSRNSGSSSESLPNLARSLLLVDQLIDL		
ASPCR1	Q9BZE9	HUMAN Tether containing UBX domain for GLUT4
<i>Nucleotide Sequence (Seq ID No. 2):</i> >P000270_CAN_CAN1-2_ASPSCR1_79058_Homo sapiens alveolar soft part sarcoma chromosome region candidate 1_BC018722.1_AAH18722.1_Q9BZE9_0_0_1662_0_1659 ATGGCGGCCCCGGCAGGCGGCGGAGGCTCCGCGGTGTGCGGTGCTGGCCCCGAACGGCCGGCGCCACAC GGTGAAGGTGACGCCGAGCACCGTGCTGCTTCAGGTTCTGGAGGACACGTGCCGGCGGCAGGACTTCAAC CCCTGTGAATATGATCTGAAGTTTCAGAGGAGCGTGCTCGACCTTTCTCCAGTGGAGATTGGCCAACTG		

CCCAACAAATGCCAAGCTGGAGATGGTGCCCGCTTCCCGGAGCCGTGAGGGGCCTGAGAACATGGTTTCGA
 TCGCTTTGCAGCTGGACGATGGCTCGAGGTTGCAGGACTCTTTCTGTTTCAGGCCAGACCCTCTGGGAGCTT
 CTCAGCCATTTTCCACAGATCAGGGAGTGCCTGCAGCACCCCGCGGGGCCACCCAGTCTGCGTGTACAC
 GAGGGATGAGGTGACGGGTGAAGCTGCCCTGCGGGGCACGACGCTGCAGTCGCTGGGCCTGACCGGGGG
 CAGCGCCACCATCAGGTTTGTATGAAGTGCTACGACCCCGTGGGCAAGACCCAGGAAGCCTGGGCTCGT
 CAGCGTCGGCTGGCCAGGCAGCCGCCAGCGCTCCACTTCCCTTGAATCTGGGGAGCTCAGCCGCGGCGA
 CTTGAGCCGTCCGGAGGACGCGGACACCTCAGGGCCCTGCTGCGAGCACACTCAGGAGAAGCAGAGCACA
 AGGGCACCCGAGCTGCCCCCTTTGTTCTTTCTCGGGTGGGGGACAGAGACAGGGGGGCCCTCCTGGGC
 CCACGAGGCCTCTGACATCATCTTCAGCTAAGTTGCCGAAGTCCCTCTCCAGCCCTGGAGGCCCTCCAAG
 CCAAAGAAGTCCAAGTCGGGCCAGGATCCCCAGCAGGAGCAGGAGCAGGAGCGGGAGCGGGATCCCCAG
 CAGGAGCAGGAGCGGGAGCGGCCCGTGGACCGGGAGCCCGTGGACCGGGAGCCGGTGGTGTGCCACCC
 CGACCTGGAGGAGCGGCTGCAGGCCTGGCCAGCGGAGCTGCCTGATGAGTTCTTTGAGCTGACGGTGGAC
 GACGTGAGAAGACGCTTGGCCCAGCTCAAGAGTGAGCGGAAGCGCCTGGAAGAAGCCCCCTTGGTGACCA
 AGGCCTTCAGGGAGGCGCAGATAAAGGAGAAGCTGGAGCGCTACCCAAAGGTGGCTCTGAGGGTCTGTT
 CCCCCAGCCGTACGTCCTACAGGGCTTCTTCCGCCCGAGCGAGACAGTGGGGGACTTGCGAGACTTCGTGA
 GGAGCCACCTGGGGAACCCCGAGCTGTATTTTACCTGTTTCATCACCCCTCCAAAAACAGTCTGGACGACC
 ACACGCAGACCCTCTTTCAGGCGAACCTCTTCCCGGCCGCTCTGGTGCCTTGGGAGCCGAGGAGCCGGC
 AGGTGTCTACCTGGAGCCTGGCCTGCTGGAGCATGCCATCTCCCCATCTGCGGCCGACGTGCTGGTGGCC
 AGGTACATGTCCAGGGCCGCCGGTCCCTTCCCCATTGCCAGCCCTGACCCTGCACCTAAGTCTGAGCC
 AGCTGCTGAGGAGGGGGCGCTGGTCCCCCTGAGCCCATCCAGGGACGGCCAGCCCGTGAAGAGGAG
 CCTGGGCAAGGTGCCCAAGTGGCTGAAGTGCCGGCCAGCAAGAGG

Protein Sequence (Seq ID No. 18):

>sp|Q9BZE9|ASPC1_HUMAN Tether containing UBX domain for GLUT4 OS=Homo sapiens OX=9606
 GN=ASPCR1 PE=1 SV=1

MAAPAGGGGSVSVLAPNGRRHTVKVTPSTVLLQVLEDTCRRQDFNPCEYDLKFQRSVLDLSLQWRFANLPNN
 AKLEMVPASRSREGPENMVRIALQLDDGSRLQDSFCSGQTLWELLSHFPQIRECLQHPGGATPVCVYTRDEVTE
 EAALRGTTQLSLGLTGGSATIRFVMKCYDPVGKTPGSLGSSASAGQAAASAPLPLESGELSRGDLRSEDADTS
 GPCCEHTQEKQSTRAPAAAPFVPFSGGGQRLGGPPGPTRPLTSSSAKLPKSLSSPGGSPKPKSKSGQDPQQE
 QEQUERERDPQQEQUERERPVDREPVDREPVVCHPDLEERLQAWPAELPDEFFELTVDDVRRRLAQLKSERKRL
 EAPLVTAKAFREAQIKEKLERYPKVALRVLPDRYVLQGFRRPSETVGDLRDFVRSHLGNPELSFYLFITPPKTVLDD
 HTQTLFQANLFPAALVHLGAEPAAGVYLEPGLLEHAISPSAADVLVARYMSRAAGSPSPLPAPDPAPKSEPAEE
 GALVPPPEPIPGTAQPVKRSGLKVPKWLKLPASKR

AURKA O14965 HUMAN Aurora kinase A

Nucleotide Sequence (Seq ID No. 3):

>P000003_KIN96_KIN_STK6_6790_Homo sapiens serine/threonine kinase 6 transcript variant
 1_BC001280.1_AA01280.1_O14965_56781.92_0_1212_0_1209

ATGGACCGATCTAAAGAAAAGTGCATTTTCAGGACCTGTTAAGGCTACAGCTCCAGTTGGAGGTCCAAAACGT
 GTTCTCGTGACTCAGCAATTTCTTGTGAGAATCCATTACCTGTAAATAGTGGCCAGGCTCAGCGGGTCTTGT
 GTCCTTCAAATTTCTCCAGCGCTTCTTTGCAAGCACAAAAGCTTGTCTCCAGTCACAAGCCGGTTCAGAA
 TCAGAAGCAGAAGCAATTGCAGGCAACCAAGTGTACCTCATCTGTCTCCAGGCCACTGAATAACACCCAAAA
 GAGCAAGCAGCCCTGCCATCGGCACCTGAAAATAATCCTGAGGAGGAAGTGGCATCAAAACAGAAAAATG
 AAGAATCAAAAAAGAGGCAGTGGGCTTTGGAAGACTTTGAAATTGGTCGCCCTCTGGGTAAAGGAAAGTTTG
 GTAATGTTTATTTGGCAAGAGAAAAAGCAAGGTTTATTCTGGCTCTTAAAGTGTATTATTAAGCTCAGCTG
 GAGAAAGCCGGAGTGGAGCATCAGCTCAGAAGAGAAGTAGAAATACAGTCCCACCTTCGGCATCCTAATATT
 CTTAGACTGTATGGTTATTTCCATGATGCTACCAGAGTCTACCTAATTCTGGAATATGCACCACTTGAACAG
 TTTATAGAGAACTTCAGAACTTTCAAAGTTTGATGAGCAGAGAACTGCTACTTATATAACAGAATTGGCAAT
 GCCCTGTCTTACTGTCATTGCAAGAGAGTTATTATAGAGACATTAAGCCAGAGAACTTACTTCTTGGATCAG
 CTGGAGAGCTTAAATTCAGATTTTGGGTGGTCACTACATGCTCCATCTTCCAGGAGGACCACTCTCTGTG
 GCACCCTGGACTACCTGCCCCCTGAAATGATTGAAGGTGGATGCATGATGAGAAGGTGGATCTCTGGAGC
 CTTGGAGTTCTTTGCTATGAATTTTGTGGAAGCCTCCTTTTGGGCAACACATACCAAGAGACCTACA
 AAAGAATATCACGGGTGAATTCACATTCCTGACTTTGTAACAGAGGGAGCCAGGGACCTCATTTCAAGACT
 GTTGAAGCATAATCCAGCCAGAGGCCAATGCTCAGAGAAGTACTTGAACACCCCTGGATCACAGCAAATTC

ATCAAAACCATCAAATTGCCAAAACAAAGAATCAGCTAGCAAACAGTCT

Protein Sequence (Seq ID No. 19):

>sp|O14965|AURKA_HUMAN Aurora kinase A OS=Homo sapiens OX=9606 GN=AURKA PE=1 SV=2

MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQAQKLVSSHKPVQN
QKQKQLQATSVPHVSRPLNNTQKSKQPLPSAPENNPEELASKQKNEESKKRQWALEDFEIGRPLGKGFQNV
YLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNLRLYGYFHDATRVYLILEYAPLGTVYRELQKL
SKFDEQRTATYITELANALSYCHSKRVIHRDIKPENLLGSAAGELKIADFGWSVHAPSSRRRTTLCGTLDYLPPEMIE
GRMHDEKVDLWSLGLVLCYEFVVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLKHNPSQRPMLREV
LEHPWITANSSKPSNCQNKESASKQS

CASP10 Q92851 HUMAN Caspase-10

Nucleotide Sequence (Seq ID No. 4):

>P001817_Q305_Q305p2_CASP10_843_Homo sapiens caspase 10 apoptosis-related cysteine
protease_BC042844.1_AAH42844.1_Q92851_0_0_1569_0_1566

ATGAAATCTCAAGGTCAACATTGGTATTCCAGTTCAGATAAAAACTGTAAAGTGAGCTTTCGTGAGAAGCTTCTG
ATTATTGATTCAAACCTGGGGGTCCAAGATGTGGAGAACCTCAAGTTTCTCTGCATAGGATTGGTCCCCAACAA
GAAGCTGGAGAAGTCCAGCTCAGCCTCAGATGTTTTGAACATCTCTTGGCAGAGGATCTGCTGAGTGAGGAA
GACCTTTTCTTCTGGCAGAACTCCTCTATATCATACGGCAGAAGAAGCTGCTGCAGCACCTCAACTGTACCAA
AGAGGAAGTGGAGCGACTGCTGCCACCCGACAAAGGGTTTCTCTGTTTAGAAACCTGCTCTACGAACTGTCA
GAAGGCATTGACTCAGAGAACTTAAAGGACATGATCTTCTTCTGAAAGACTCGCTTCCCAAACTGAAATGAC
CTCCCTAAGTTTCTGGCATTCTAGAGAAACAAGGTAAAATAGATGAAGATAATCTGACATGCCTGGAGGACCT
CTGCAAAACAGTTGTACCTAACTTTTGAAGAAACATAGAGAAATACAAAAGAGAGAAAGCTATCCAGATAGTGAC
ACCTCCTGTAGACAAGGAAGCCGAGTCTGATCAAGGAGAGGAAGAACTAGTTTCCCAACAGATGTTAAGACAT
TCTTGAAGCCTTACCGCAGGAGTCTGGCAAAATAAGCATGCAGGTAGTAATGGTAACAGAGCCACAAATGGT
GCACCAAGCCTGGTCTCCAGGGGGATGCAAGGAGCATCTGCTAACACTCTAACTCTGAAACCAGCACAAAGA
GGGCAGCTGTGTACAGGATGAATCGGAACCACAGAGGCCTCTGTGTCATTGTCAACAACCACAGCTTACCTC
CCTGAAGGACAGACAAGGAACCCATAAAGATGCTGAGATCCTGAGTCATGTGTTCCAGTGGCTTGGGTTTACA
GTGCATATACACAATAATGTGACGAAAGTGAAATGGAGATGGTCTGCAGAAGCAGAAGTGAATCCAGCCC
ATGCCGACGGGACTGCTTCGTGTTCTGTATTCTGACCCATGGGAGATTTGGAGCTGTCTACTCTTCGGATGAG
GCCCTCATTCCCATTCGGGAGATCATGTCTCACTTCACAGCCCTGCAGTGCCTAGACTGGCTGAAAAACCTAA
ACTCTTTTTCATCCAGGCCTGCCAAGGTGAAGAGATACAGCCTTCCGTATCCATCGAAGCAGATGCTCTGAACC
CTGAGCAGGCACCCACTTCCCTGCAGGACAGTATTCTGCCGAGGCTGACTTCTACTTGGTCTGGCCACTGT
CCCAGGCTATGTATCCTTTCGGCATGTGGAGGAAGGCAGCTGGTATATTCAGTCTCTGTGTAATCATCTGAAGA
AATTGGTCCCAAGACATGAAGACATCTTATCCATCCTCACTGCTGTCAACGATGATGTGAGTCGAAGAGTGGAC
AAACAGGGAACAAAGAAACAGATGCCCCAGCCTGCTTTCACACTAAGGAAAAAACTAGTATTCCCTGTGCCCT
GGATGCACTTTCATTA

Protein Sequence (Seq ID No. 20):

>sp|Q92851|CASP10_HUMAN Caspase-10 OS=Homo sapiens OX=9606 GN=CASP10 PE=1 SV=3

MKSQGGHWYSSSDKNCKVSFREKLLIIDSNLGVQDVENLKFCLGLVPNKKLEKSSASDVFEHLLAEDLLSEDPFF
LAELLYIIRQKKLLQHLNCTKEEVERLLPTRQRVSLFRNLLYELSEGIDSENLDKMIFLLKDSLPTKEMTSLSFLAFLEKQ
GKIDEDNLTLEDLCKTVVPKLLRNIEKYKREKAIQIVTPVDKEAESYQGEELVSQTDVKTFLALPQESWQNKHA
GSNGNRATNGAPSLVSRGMQASANTLNSETSTKRAAVYRMNRNHRGLCVIVNNHSFTSLKDRQGTGTHKDAEILSHV
FQWLGFVTHHNNVTKVEMEMVLQKQKCNPAHADGDCVFICILTHGRFGAVYSSDEALIPREIMSHFTALQCPRLA
KPKLFFIQACQGEIIPSVSIEADALNPEQAPTSQDSIPAEADFLGLATVPGYVSFRHVEEGSWYIQLCNHLKLLV
PRMLKFLEKTMEIRGRKRTVWGAKQISATSLPTAISAQTPRPMMRWSSVS

CD96 P40200 HUMAN T-cell surface protein tactile

Nucleotide Sequence (Seq ID No. 5):

>P002164_Q305_Q305p3_CD96_10225_Homo sapiens CD96
antigen_BC020749.1_AAH20749.1_P40200_0_0_1209_0_1206

ATGGAGAAAAATGGAAATACTGTGCTGTCTATTACATCATCCAGATACATTTTGTCAAGGGAGTTTGGGAAAA
 ACAGTCAACACAGAAGAAAATGTTTATGCTACACTTGGCTCTGATGTCAACCTGACCTGCCAAACACAGACAGT
 AGGCTTCTTCGTGCAGATGCAATGGTCCAAGGTACCAATAAGATAGACCTGATTGCTGTCTATCATCCCCAAT
 ACGGCTTCTACTGTGCCTATGGGAGACCCTGTGAGTCACTTGTGACTTTCACAGAACTCCTGAGAATGGGTCA
 AAATGGACTCTGCACTTAAGGAATATGTCTTGTTCAGTCAGTGAAGGTACGAGTGTATGCTTGTCTGTATCCA
 GAGGGCATTACAGACTAAAATCTACAACCTTCTCATTAGACACACGTTACAGCAGATGAATGGAACAGCAACCA
 TACGATAGAAATAGAGATAAATCAGACTCTGGAAATACCATGCTTTCAAATAGCTCCTCAAAAATTTTCATCTGAG
 TTCACCTATGCATGGTCCGTGGAGGATAATGGAACCTCAGGAAACACTTATCTCCCAAAATCACCTCATCAGCAA
 TTCCACATTACTTAAAGATAGAGTCAAGCTTGGTACAGACTACAGACTCCACCTCTCTCCAGTCCAAATCTTCGA
 TGATGGGCGGAAGTTCTCTTGCCACATTAGAGTCGGTCCCTAACAAAATCTTGAGGAGCTCCACCACAGTCAAGG
 TTTTGTCTAAACCAGAAATCCCTGTGATTGTGGAAAATACTCCACGGATGTCTTGGTAGAGAGAAGATTCACCT
 GCTTACTAAAGATGTATTTCCCAAAGCAAATATCACATGGTTTATAGATGGAAGTTTTCTTCATGATGAAAAAGA
 AGGAATATATATTACTAATGAAGAGAGAAAAAGGCAAAGATGGATTTTTGGAAGTGAAGTCTGTTTTAACAAGGGT
 ACATAGTAATAAACCCAGCCCAATCAGACAACCTTGACCATTTGGTGTATGGCTCTGTCTCCAGTCCCAGGAAATAA
 AGTGTGGAACATCTCATCAGAAAAGATCACTTTTCTCTTAGGTTCTGAAATTTCCCTCAACAGACCCTCCACTGAG
 TGTTACAGAATCTACCCTTGACACCCAACCTTCTCCAGCCAGCAGTGTATCTCCTGCAAGTAAGAATGTTTTAC
 ACTGAGCTAT

Protein Sequence (Seq ID No. 21):

>sp|P40200|TACT_HUMAN T-cell surface protein tactile OS=Homo sapiens OX=9606 GN=CD96 PE=1 SV=2

MEKKWKYCAVYIIQIHFKGVWEKTVNTEENVYATLGSDVNLTCQTQTVGFFVQMWSKVNTKIDLIIVYHPQYGF
 YCAYGRPCESLVFTETPENGSKWTLHLRNMSCSVSGRYECMLVLYPEGIQTKIYNLLIQHTVTADEWNSNHTIEIEIN
 QTLEIPCFQNSSSKISSEFTYAWSVENSSTDWSVLLSKGIKEDNGTQETLISQNHLSNSTLLKDRVKLGTDYRLHLS
 VQIFDDGRKFSCHIRVGPKNILRSSTTVKVFAPKEIPVIVENNSTDVLEVRFTCLLKNVFPKANITWFDGSFLHDEKE
 GIYITNEERKKGDFLELKSVLTRVHSNKPQSDNLTWCMALSPVPGNKVWNISSEKITFLLGSEISSDPPLSVTES
 TLDTPSPASSVSPARYPATSSVTLVDVSALRPNTTPQPSNSSMTTRGFNYPWTSSGTDTKKSVSRIPSETYSSSPS
 GAGSTLHDNVFTSTARAFSEVPTTANGSTKTNHVHITGIVVNKPKDGMSPVIVAALLFCCMILFGLGVRKWCQYQK
 EIMERPPPFKPPPPPIKYTCIQEPNESDLPYHEMETL

FEN1	P39748	HUMAN Flap endonuclease 1
-------------	--------	---------------------------

Nucleotide Sequence (Seq ID No. 6):

>P000413_SIG_SIG1-2_FEN1_2237_Homo sapiens flap structure-specific endonuclease
 1_BC000323.2_AAH00323.1_P39748_53567_0_1143_0_1140

ATGGGAATTCAAGGCCTGGCCAACTAATTGCTGATGTGGCCCCAGTGCCATCCGGGAGAATGACATCAAGA
 GCTACTTTGGCCGTAAGGTGGCCATTGATGCCTCTATGACATTTATCAGTTCTGATTGCTGTTCGCCAGGGT
 GGGGATGTGCTGCAGATGAGGAGGGTGAGACCACGACCCAGTATGGGCATGTTCTACCGCACCATTTCG
 ATGATGGAGATCAAGCCGCTGATGTCTTTGATGGCAAGCCGCCACAGCTCAAGTCAGGCGAGCTGG
 CCAAACGCAGTGAGCGGGGGCTGAGGCAGAGAAGCAGCTGCAGCAGGCTCAGGCTGCTGGGGCCGAGCAG
 GAGGTGAAAAATCACTAAGCGGCTGGTGAAGGTCACTAAGCAGCACAATGATGAGTGCAAACATCTGCTGA
 GCCTCATGGGCATCCCTTATCTTGATGCACCCAGTGAGGCAGAGGCCAGCTGTGCTGCCCTGGTGAAGGCTG
 GCAAAGTCTATGCTGCGGCTACCGAGGACATGGACTGCCTCACCTTCGGCAGCCCTGTGCTAATGCGACACCT
 GACTGCCAGTGAAGCCAAAAAGCTGCCAATCCAGGAATCCACCTGAGCCGGATTCTGCAGGAGCTGGGCCTG
 AACCAGGAACAGTTTGTGGATCTGTGCATCCTGCTAGGCAGTGACTACTGTGAGAGTATCCGGGGTATTGGGC
 CCAAGCGGGCTGTGGACCTCATCCAGAAGCACAAGAGCATCGAGGAGATCGTGGCGGACTTGACCCCAACA
 AGTACCCTGTGCCAGAAAATTGGCTCCACAAGGAGGCTCACCAGCTCTTCTTGAACCTGAGGTGCTGGACCC
 AGAGTCTGTGGAGCTGAAGTGGAGCGAGCCAAATGAAGAAGAGCTGATCAAGTTCATGTGTGGTGAAGAAGCAG
 TTCTCTGAGGAGCGAATCCGCAGTGGGGTCAAGAGGCTGAGTAAGAGCCGCCAAGGCAGCAGCCAGGGCCCG
 CTGGATGATTTCTTCAAGGTGACCGGCTCACTCTCTTCACTAAGCGCAAGGAGCCAGAACCCAGGGATCCA
 CTAAGAAGAAGGCAAAGACTGGGGCAGCAGGGAAGTTTAAAAGGGGAAAA

Protein Sequence (Seq ID No. 22):

>sp|P39748|FEN1_HUMAN Flap endonuclease 1 OS=Homo sapiens OX=9606 GN=FEN1 PE=1 SV=1
 MGIIQLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSIYQFLIAVRQGGDLQNEEGETTSHLMGMFYRTIRMMENGI
 KPVYVFDGKPPQLKSGELAKRSERRAEAEKQLQQAAGAEQEVEKFTKRLVKVTKQHNDCKHLLSLMGIPYLD

PSEAEASCAALVKAGKVYAAATEDMDCLTFGSPVLMRHLTASEAKKLPQEFHLSRILQELGLNQEQFV DLCILLGSD
YCESIRGIGPKRAVDLIQKHKSIEEIVRRDPNKYPVPENWLHKEAHQLFLEPEVLDPESELKWSEPNEEELIKFMCG
EKQFSEERIRSGVKRLSKSRQGSTQGRLLDFFKVTGSLSSAKRKEPEPKGSTKKKAKTGAAGKFKRGK

FKBP3 Q00688 HUMAN Peptidyl-prolyl cis-trans isomerase FKBP3

Nucleotide Sequence (Seq ID No. 7):

>P001211_CAG_CAGp1_FKBP3_2287_Homo sapiens FK506 binding protein 3
25kDa_BC016288.1_AAH16288.1_Q00688_0_0_675_0_672

ATGGCGGCGGCCGTTCCACAGCGGGCGTGGACCGTGGAGCAGCTGCGCAGTGAGCAGCTGCCCAAGAAGGA
CATTATCAAGTTTCTGCAGGAACACGGTTCAGATTCTGTTTCTTGCAGAACATAAATTATTAGGAAACATTAAAAAT
GTGGCCAAGACAGCTAACCAAGGACCACTTGGTTACAGCCTATAACCATCTTTTGAAGCTAAGCGTTTAAAGGGT
ACTGAAAGTATAAGTAAAGTGTCTGAGCAAGTAAAAATGTGAAGCTTAATGAAGATAAACCCAAAGAAACCAAG
TCTGAAGAGACCCCTGGATGAGGGTCCACCAAAATATACTAAATCTGTTCTGAAAAAGGGAGATAAAACCAACTTT
CCCAAAAAGGGAGATGTTGTTCACTGCTGGTATACAGGAACACTACAAGATGGGACTGTTTTTGATACTAATATT
CAAACAAGTGCAAAGAAGAAAAATGCCAAGCCTTTAAGTTTTAAGGTCGGAGTAGGCAAAGTTATCAGAGG
ATGGGATGAAGCTCTCTTGACTATGAGTAAAGGAGAAAAGGCTCGACTGGAGATTGAACCAGAATGGGCTTAC
GGAAAGAAAGGACAGCCTGATGCCAAAATCCACCAATGCAAACTCACTTTTGAAGTGGAATTAGTGGATATT
GAT

Protein Sequence (Seq ID No. 23):

>sp|Q00688|FKBP3_HUMAN Peptidyl-prolyl cis-trans isomerase FKBP3 OS=Homo sapiens OX=9606 GN=FKBP3
PE=1 SV=1

MAAVPQRAWTVEQLRSEQLPKKDIKFLQEHSDFSFLAEHKLGNIKNVAKTANKDHLVTAYNHLFETKRFKGTESI
SKVSEQVKNVKLNEDPKETKSEETLDEGPPKYTKSVLKKGDKTNFPKKGDVVHCWYTGLQDGTVFDNIQTSK
KKKNAKPLSFKVGVGVIRGWDEALLTMSKGEKARLEIEPEWAYGKKGQPDAKIPPNAKLTFEVELVDID

FMNL2 Q96PY5 HUMAN Formin-like protein 2

Nucleotide Sequence (Seq ID No. 8):

>P000661_TRN_TRNp1_FHOD2_114793_Homo sapiens formin homology 2 domain containing
2_BC036492.2_AAH36492.1_Q96PY5_0_0_537_0_534

ATGGAAGCTGACCAAGAGAGAGTACACCATGCATGACCATAACACGCTGCTGAAGGAGTTCATCCTCAACAATGA
GGGGAAGCTGAAGAAGCTGCAGGATGATGCCAAGATCGCACAGGATGCCTTTGATGATGTTGTGAAGTATTTT
GGAGAAAACCCCAAGACAACACCACCCTCTGTCTTCTTCTCTGCTTTGTCCGGTTTGTGAAAGCATATAAGCAA
GCAGAAGAGGAAAATGAGCTGAGGAAAAAGCAGGAACAAGCTCTCATGGAAAACTCCTAGAGCAAGAAGCTC
TGATGGAGCAGCAGGATCCAAAGTCTCCTTCTCATAAATCAAAGAGGCAGCAGCAAGAGTTAATTGCAGAATTA
AGAAGACGACAAGTTAAAGATAACAGACATGTATATGAGGGAAAAGATGGTGCCATTGAAGATATTATCACAGC
CTTAAAGAAGAATAATCACTAAATTTCCAAATGTTCACTCGAGGGTAAGGATTTCTTCTAGCACACCGGTGGT
GGAGGATACACAGAGC

Protein Sequence (Seq ID No. 24):

>sp|Q96PY5|FMNL2_HUMAN Formin-like protein 2 OS=Homo sapiens OX=9606 GN=FMNL2 PE=1 SV=3

MGNAGSMDSQQTDFRAHNVPLKLPMPEPGELEERFAIVLNAMELPPDKARLLRQYDNEKKWELICDQERFQVKNP
PHTYIQKLKGYLDPAVTRKKFRRRVQESTQVLRLEISLRTNHIGWVREFLNEENKGLDVLVEYLSFAQYAVTFDFES
VESTVESSVDKSKPWSRSIEDLHRGNSLSPVGNVSRSRHSALRYNTLPSRRTLKNSRLVSKDDVHVCMCLR
AIMNYQYGFNMVMSHPHAVNEIALSLNNKNPRTKALVELLAAVCLVRGGHEILSAFDNFKEVCGEKQRFKLEMEHF
RNEDNNIDFMVASMQFINIVVHVEDMNFVRHLQYEFTKLGLDEYLDKLKHTESDKLQVQIQAYLDNVFDVGALLEDA
ETKNAALERVEELENISHLSEKLQDTENEAMSKIVELEKQLMQRNKELDVREIYKDANTQVHTLRKMVKEKEEAIQ
RQSTLEKKIHELEKQGTIKIKKGDGDIAPVVASGTLMSGSEVVAGNSVGPTMGAASSGPPPPPPPLPPSSDTPE
TVQNGPVTTPMPPPPPPPPPPPPPPPPPPPLPGAAETVPAPPLAPPLPSAPPLPGTSSPTVVFNGLAAVKKIKP
IKTKFRMPVFVNWALKPNQINGTVFNEIDDERILEDLNVDEFEEIFKTKAQGPAIDLSSSKQKIPQKGSNKVTLLEANRA
KNLAITLRKAGKTADEICKAIHVFDLKLTPVDFVECLMRFLPTENEVKVLRLYERERKPLENLSDEDRFMMQFSKIERL
MQKMTIMAFIGNFAESIQLTPQLHAIIAASVSIKSSQKLKILEIILALGNMNSSKRGAVYGFKLQSLDLLLDTKSTDR

KQTLHYISNVVKEKYHQVSLFYNELHYVEKAAVSLNVLDDVKELQRGMDLTKREYTMHDHNTLLKEFILNNEGKL
KKLQDDAKIAQDAFDDVVKYFGENPKTTPPSVFFPVFVRFVKAYKQAEENELRKKQEALMEKLLEQEALMEQQD
PKSPSHKSKRQQELIAELRRRQVKDNRHVYEGKDGAIEDIITVLKTPFTARTAKRGSRRFFCEPVLTEEYHY

GLRX3 O76003 HUMAN Glutaredoxin-3

Nucleotide Sequence (Seq ID No. 9):

>P000071_KIN96_KIN_TXNL2_10539_Homo sapiens thioredoxin-like clone

MGC:12349_BC005289_AAH05289_O76003_48409.07_0_1008_0_1005

ATGGCGGCGGGGGCGGCTGAGGCAGCTGTAGCGGCCGTGGAGGAGGTCGGCTCAGCCGGGCAGTTTGAGG
AGCTGCTGCGCCTCAAAGCCAAGTCCCTCCTTGTGGTCCATTTCTGGGCACCATGGGCTCCACAGTGTGCACA
GATGAACGAAGTTATGGCAGAGTTAGCTAAAGAACTCCCTCAAGTTTCATTTGTGAAGTTGGAAGCTGAAGGTG
TTCCTGAAGTATCTGAAAAATATGAAATTAGCTCTGTTCCCACTTTTCTGTTTTCAAGAATTCTCAGAAAAATCGA
CCGATTAGATGGTGCACATGCCCCAGAGTTGACCAAAAAAGTTCAGCGACATGCATCTAGTGGCTCCTTCCTAC
CCAGCGCTAATGAACATCTTAAAGAAGATCTCAACCTTCGCTTGAAGAAATTGACTCATGCTGCCCCCTGCATG
CTGTTTATGAAAGGAAGTCTCAAGAACCACGCTGTGGTTTCAGCAAGCAGATGGTGGAAATTCCTCACAAACA
TAATATTCAGTTTAGCAGTTTTGATATCTTCTCAGATGAAGAGGTTTCGACAGGGACTCAAAGCCTATTCCAGTTG
GCCTACCTATCCTCAGCTCTATGTTTCTGGAGAGCTCATAGGAGGACTTGATATAATTAAGGAGCTAGAAGCAT
CTGAAGAAGTAGATACAATTTGTCCCAAAGCTCCCAAATTAGAGGAAAGGCTCAAAGTGTGACAAATAAAGCTT
CTGTGATGCTCTTTATGAAAGGAAACAAACAGGAAGCAAATGTGGATTGAGCAAAACAAATTCGGAAATACTAA
ATAGTACTGGTGTGTAATATGAAACATTCGATATATTGGAGGATGAAGAAGTTCGGCAAGGATTAAGGCTTACT
CAAATTGGCCCAACATACCCTCAGCTGTATGTGAAAGGGGAGCTGGTGGGAGGATTGGATATTGTGAAGGAACT
GAAAGAAAATGGTGAATTGCTGCCTATACTGAGAGGAGAAAAAT

Protein Sequence (Seq ID No. 25):

>sp|O76003|GLRX3_HUMAN Glutaredoxin-3 OS=Homo sapiens OX=9606 GN=GLRX3 PE=1 SV=2

MAAGAAEAAVAVEEVGSAGQFEELLRLKAKSLLVHFVWAPWAPQCAQMNEVMAELAKELPQVSFVKLEAGVPE
VSEKEYEISSVPTFLFFKNSQKIDRLDGAHAPELTKVQRHASSGSFLPSANEHLKEDLNLRLLKLTHAAPCMLFMKGT
PQEPRCGFSKQMVILHKNHNIQFSSFDIFSDEEVRQGLKAYSSWPTYPLQYVSGELIGGLDIIKELEASEELDTICPKA
PKLEERLKVLTNKASVMLFMKGNKQEAFCGFSKQILEILNSTGVEYETFDILEDEEVRQGLKAYSNWPTYPLQYVKGE
LVGGLDIVKELKENGELLPILRGEN

MAP3K13 O43283 HUMAN Mitogen-activated protein kinase kinase kinase 13

Nucleotide Sequence (Seq ID No. 10):

>P001569_Q106_Q106p2_MAP3K13_9175_Homo sapiens MAP3K13 mitogen-activated protein kinase kinase
kinase 13_NM_004721_0_0_0_0_0_0

ATGGCCAACTTCAGGAGCACCTGAGCTGCTCCTCTTCTCCACACTTACCCTTCAGTGAAAGCAAAACCTTCAA
TGGACTACAAGATGAGCTCACAGCTATGGGGAACCACCCTTCTCCCAAGCTGCTCGAGGACCAGCAGGAAAAAG
GGGATGGTACGAACAGAGCTAATCGAGAGCGTGACAGCCCCGTCACCACAACAGTGTGACGAGCGTAAGT
GAGGATTCAGGGACCAGTTTGAGAACAGCGTTCTTCAGCTAAGGGAACACGATGAATCAGAGACGGCGGTGT
CTCAGGGGAACAGCAACACGGTGGACGGAGAGAGCACAAGCGGGAAGTGAAGACATAAAGATTGAGTTGAGCA
GGTCAGGCAGTGGCAGTGGTGGGTTTCTTGAAGGACTATTGGATGCTTAAGGCCTGTATGGAATATCATTGGG
AAGGCATATTCCACTGATTACAAATTGCAGCAGCAAGATACTTGGGAAGTGCCATTTGAGGAGATCTCAGAGCT
GCAGTGGCTGGGTAGTGGAGCCCAAGGAGCGGTCTTCTTGGGCAAGTTCGGGCGGAAGAGGTGGCCATCAA
GAAAGTGAGAGAACAGAATGAGACGGATATCAAGCATTTGAGGAAGTTGAAGCACCTAACATCATCGCATTCA
AGGGTGTGTTGACTCAGGCCCCATGTTATTGTATTATCATGGAATACTGTGCCCATGGACAACCTCTACGAGGTCT
TACGAGCTGGCAGGAAGATCACACCTCGATTGCTAGTAGACTGGTCCACAGGAATTGCAAGTGAATGAATTAT
TTGCACCTCCATAAAATTATTCATCGTGATCTCAAATCACCTAATGTTTTAGTGACCCACACAGATGCGGTAAGAA
TTTCAGATTTTGGTACATCTAAGGAAGTCAAGTACAAAAGTACCAAGATGTCATTTGCTGGCAGCGTCCGATGG
ATGGCGCCAGAGGTGATACGGAATGAACCTGTCTCTGAAAAAGTTGATATATGGTCTTTTGGAGTGGTGGCTTTG
GGAGCTGCTGACAGGAGAGATCCCTTACAAAGATGTAGATTCTTCAGCCATTATCTGGGGTGTGGAAGCAACA
GCCTCCACCTTCCAGTTCCTTCCACTTGCCCTGATGGATTCAAATCCTTATGAAACAGACGTGGCAGAGTAAA
CCTCGAAACCGACCTTCTTTTCGGCAGACACTCATGATTTAGACATTGCCTCTGCAGATGTAATTTGCCACCCC
ACAAGAACTTACTTCAAGTCTCAGGCTGAATGGAGAGAAGAAGTGAACAAAACATTTTGAAGAAGATCAAAAGTGA
AGGAAGTGTATACACCGGTTAGATGAAGAACTGATTGCAAGGCGCAGAGAAGAGCTCAGGCATGCGCTGGAT

ATTCGTGAACACTATGAGCGGAAGCTTGAGCGGGCGAATAATTTATACATGGAATTGAGTGCCATCATGCTGCA
 GCTAGAAATGCGGGAGAAGGAGCTCATTAAGCGTGAGCAAGCAGTGGAAAAGAAGTATCTGGGACCTACAAA
 CGACACCCTGTTCTGTCCTATCATCCATCCCAATGCCATGGAGAACTCATGAAAAGGAAAGGAGTGCCTCACA
 ATCTGGGATGCAGACCAAACGGCCAGACTTGTGAGATCAGAAGGGATCCCCACCACAGAAGTGGCTCCCACT
 GCATCCCCCTTTGTCCGGAAGTCCCAAAATGTCCACTTCTAGCAGCAAGAGCCGATATCGAAGCAAACACGCCA
 CCGCCGAGGGAATAGCAGAGGCAGCCATAGTGACTTTGCCGCAATCTTGAAAAACCAGCCAGCCAGGAAAAT
 TCACCCCATCCCCTTACCTGCACCAAGCTCAATCCCAATACCCTTCTCTTCATCACCATAATTCTCTGCAGCAG
 CAATACCAGCAGCCCCCTCCTGCCATGTCCCAGAGTCACCATCCCAGACTCAATATGCACGGACAGGACATAG
 CAACCTGCGCCAACAACCTGAGGTATTTGGGCCAGCAGCAGCCCTGCGGAGCCCACTCAGCAACCATGCTCA
 GAGACAGCTGCCCCGGCTCGAGCCCTGACCTCATCTCCACAGCCATGGCTGCAGACTGCTGGAGAAGTTCTGA
 GCCTGACAAGGGCCAAGCTGGTCCCTGGGGCTGTTGCCAGGCTGACGCTTATGACCCCTGCCTTCAGTGCAG
 GCCAGAACAGTATGGTCCCTAGACATACCCTCTGCTGAGCCAGTGGGGAGGAGCCCTGACCTTTCCAAGTCA
 CCAGCACATAATCCTCTCTTGGAAAACGCCAGAGTTCTGAGAAAACGGAAGAAAATGAATTCAGCGGCTGTAG
 GTCTGAGTCATCCCTCGGCACCTCTCATCTCGGCACCCCTCCAGCGCTACCTCGAAAAACAAGGCCTCTGCAG
 AAGAGTGGAGATGACTCCTCAGAAGAGGAAGAAGGGGAAGTAGATAGTGAAGTTGAATTTCCACGAAGACAGA
 GGCCCCATCGCTGTATCAGCAGCTGCCAGTCATATTCAACCTTTAGCTCTGAGAAATTTCTGTGTCTGATGGA
 GAAGAGGGAAATACCAAGTACCAGTCAACAGTCTGATGAGTTAGCTGATAAATTGAAGACCGCTTGGCAGA
 GAAGCTAGACGACCTGCTGTCCCAGACGCCAGAGATTCCATTGACATATCCTCACACTCGGATGGGCTCTCT
 GACAAGGAGTGTGCCGTGCCCGTGTGAAGACTCAGATGTCTCTGGGCAAGCTGTGTGTGGAGAACGTGGC
 TATGAGAACCCCATGCAGTTTGAAGAATCGGACTGTGACTCTTCAGATGGGGAGTGTCTGATGCCACAGTTAG
 GACCAATAAACACTACAGCTCTGCTACCTGG

Protein Sequence (Seq ID No. 26):

>sp|O43283|M3K13_HUMAN Mitogen-activated protein kinase kinase kinase 13 OS=Homo sapiens OX=9606
 GN=MAP3K13 PE=1 SV=1

MANFQEHLSCSSSPHLPFSESKTFNGLQDELTAMGNHPSPKLLEDQKEKGMVRTELIESVHSPVTTTTLTSVSEDSR
 DQFENSVLQLREHDESETAVSQGNSNTVDGESTSGTEDIKIQFSRSGSGSGGFLEGLFGCLRVPVWNIIGKAYSTDYK
 LQQQDTWEVPFEEISELQWLGSGAQGAVFLGKFRAEEVAIKKVREQNETDIKHLRKLKHPNIIAFKGVCTQAPCYCII
 MEYCAHGLQYEVLRAGRKITPRLLVDWSTGIASGMNYLHLHKIHRDLKSPNVLVTHTDAVKISDFGTSKELSDKSTK
 MSFAGTVAWMAPEVIRNEPVSEKVDIWSFGVVLWELLTGEIPYKDVDSAIWGVGNSLHLPVPSTCPDGFKILMKQ
 TWQSKPRNRPFRQTLMHLDIASADVLATPQETYFKSQAEEVKKHFEKIKSEGTCIHLRDEELIRRRREELRHAI
 DIREHYERKLERANNLYMELSAIMLQLEMREKELIKREQAQVEKKYPGTYKRHPVRPIIHPNAMEKLMKRKGVPKSG
 MQTKRPDLRLSEGIPTTEVAPTASPLSGSPKMSTSSSKSRYRSKPRHRRGNSRSGSHSDFAILKNQPAQENSPHPT
 YLHQAQSQYPSLHHHNSLQQYQPPAMSQSHHPRLMHMGQDIATCANNLRYFGPAAALRSPLSNHAQRQLPG
 SSPDLISTAMAADCWRSEPDKGQAGPWGCCQADAYDPCLCRPEQYGSGLDIPSAEPVGRSPDLSKSPAHPNPLLE
 NAQSSEKTEENEFSGCRSESLGTSHLGTTPALPRKTRPLQKSGDDSSSEEEGEVDSEVEFPRRQRPHRICISSCQS
 YSTFSSSENFVSVDGEEGNTSDHSNSPDELADKLEDRLAEKLDLLSQTPFIPIDISSHSDGLSDKECAVRRVKTQMSL
 GKLCVEERGYENPMQFEESDCDSSDGECSDATVRTNKHYSATW

MAP4 P27816 HUMAN Microtubule-associated protein 4

Nucleotide Sequence (Seq ID No. 11):

>P000490_SIG_SIG1-3_MAP4_4134_Homo sapiens MAP4 microtubule-associated protein
 4_BC008715.2_AA08715.1_P27816_113843.4_0_2940_0_2937

ATGGCTGACCTCAGTCTTGAGATGCATTAACAGAACCATCTCCAGACATTGAGGGAGAGATAAAGCGGGACTT
 CATTGCCACACTAGAGGCAGAGGCCCTTTGATGATGTTGTGGGAGAACTGTTGAAAAACAGACTATATTCCTC
 TCCTGGATGTTGATGAGAAAACCGGGAACCTCAGAGTCAAAGAAGAAACCGTGCTCAGAACTAGCCAGATTGAA
 GATACTCCATCTTCTAAACCAACACTCCTAGCCAATGGTGGTCATGGAGTAGAAGGGAGCGATACTACAGGGTC
 TCCAACTGAATTCCTTGAAGAGAAAATGGCCTACCAGGAATACCCAAATAGCCAGAAGTGGCCAGAAGATACCA
 ACTTTTGTTCACCTGAGCAAGTGGTCGATCCTATCCAGACTGATCCCTTTAAGATGTACCATGATGATGACC
 TGGCAGATTTGGTCTTTCCCTCCAGTGCGACAGCTGATACTTCAATATTTGCAGGACAAAATGATCCCTTGAAAG
 ACAGTTACGGTATGTCTCCCTGCAACACAGCTGTTGTACCTCAGGGGTGGTCTGTGGAAGCCTTAACTCTCCA
 CACTCAGAGTCCTTTGTTTCCCAGAGGCTGTTGCAGAACCTCCTCAGCCAACGGCAGTTCCCTTAGAGCTAGC
 CAAGGAGATAGAAATGGCATCAGAAGAGAGGCCACCAGCACAAAGCATTGGAATAATGATGGGACTGAAGACT
 ACTGACATGGCACCATCTAAAGAAACAGAGATGGCCCTCGCCAAGGACATGGCACTAGCTACAAAAACCGAGG

TGGCATTGGCTAAAGATATGGAATCACCCACCAAATTAGATGTGACACTGGCCAAGGACATGCAGCCATCCATG
 GAATCAGATATGGCCCTAGTCAAGGACATGGAACACCCACAGAAAAAGAAGTGGCCCTGGTTAAGGATGTCA
 GATGGCCACAGAAACAGATGTATCTTCAGCCAAGAATGTGGTACTGCCACAGAAACAGAGGTAGCCCCAGC
 CAAGGATGTGACACTGTTGAAAGAAACAGAGAGGGGCATCTCCTATAAAAAATGGACTTAGCCCCTTCCAAGGACA
 TGGGACCACCCAAAGAAAAACAAGAAAGAAACAGAGAGGGGCATCTCCTATAAAAAATGGACTTGGCTCCTTCCAAG
 GACATGGGACCACCCAAAGAAAAACAAGATAGTCCCAGCCAAGGATTGGTATTACTCTCAGAAATAGAGGTGGC
 ACAGGCTAATGACATTATATCATCCACAGAAATATCCTCTGCTGAGAAGGTGGCTTTGTCTCAGAAACAGAGG
 TAGCCCTGGCCAGGGACATGACACTGCCCCCGAAACCAACGTGATCTTGACCAAGGATAAAGCACTACCTTT
 AGAAGCAGAGGTGGCCCCAGTCAAGGACATGGCTCAACTCCCAGAAACAGAAATAGCCCCGGCCAAGGATGT
 GGCTCCGTCCACAGTAAAAAAGTGGGCTTGTGAAGGACATGTCTCCACTATCAGAAACAGAAATGGCTCTGG
 GCAAGGATGTGACTCCACCTCCAGAAACAGAAAGTAGTTCTCATCAAGAACGTATGTCTGCCTCCAGAAATGGAG
 GTGGCCCTGACTGAGGATCAGGTCCCAGCCCTCAAACAGAAAGCTCCCACCACCATTTGGTGGGTTGAATAAAA
 AACCCATGAGCCTTGCTTCAGGCTTAGTGCCAGCTGCCCCACCCAAACGCCCTGCCGTGCCTCTGCCAGGCC
 TTCCATCTTACCTTCAAAAGACGTGAAGCCAAAGCCCATTCAGATGCAAAAGGCTCCTGAGAAGCGGGCCTCAC
 CATCCAAGCCAGCTTCTGCCCCAGCCTCCAGATCTGGGTCCAAGAGCACTCAGACTGTTGCAAAAAACACAAC
 AGCTGCTGCTGTTGCCTCAACTGGCCCAAGCAGTAGGAGCCCTCCACGCTCCTGCCCAAGAAGCCCACTGCC
 ATTAAGACTGAGGGAAAAACCTGCAGAAGTCAAGAAGATGACTGCAAAGTCTGTACCAGCTGACTTGAGTCGCCC
 AAAGAGCACCTCCACCAGTTCCATGAAGAAAACCACTCTCAGTGGGACAGCCCCCGCTGCAGGGGTGGTT
 CCCAGCCGAGTCAAGGCCACACCCATGCCCTCCCGGCCCTCCACAACCTCTTTCATAGACAAGAAGCCACCT
 CGGCCAAACCCAGCTCCACCACCCCGGCTCAGCCGCTGGCCACCAATACTTCTGCTCCTGATCTGAAGAA
 TGTCCGCTCAAGGTTGGCTCCACGGAACATCAAGCATCAGCCTGGAGGAGGCCGGGCCAAAGTAGAGAA
 AAAACAGAGGCAGTGTCTAACCCGAAAGCCTGAATCTAATGCAGTCACTAAACAGCCGCCCAATTGCAA
 GTGCACAGAAACAACCTGCGGGGAAAGTCCAGATAGTCTCAAAAAAGTGAGCTACAGCCATATTCAGTCCAAG
 TGTGGTTCCAAGGACAATATTAAGCATGTCCCTGGAGGTGGTAATGTTGAGATTGAGAACAAGAAAGTGGACAT
 CTCTAAGGTCTCCTCAAGTGTGGGTCTAAGGCTAACATCAAGCACAAGCCTGGTGGAGGAGATGTCAAGATT
 GAAAGTCAGAAAGTTGAACCTCAAGGAGAAGGCCAGGCCAAGGTGGGATCCCTCGATAATGTGGGCCACCTAC
 CTGCAGGAGGTGCTGTGAAGACTGAGGGCGGTGGCAGCGAGGCTCCTCTGTGTCCGGGTCCCCCTGCTGGG
 GAGGAGCCGGCCATCTCTGAGGCAGCGCCTGAAGCTGGCGCCCCACTTCAGCCAGTGGCCTCAATGGCCAC
 CCCACCCTGTGAGGGGGTGGTGACCAAGGGAGGCCAGACCTTGACAGCCAGATCCAGGAGACAAGCATC

Protein Sequence (Seq ID No. 27):

>sp|P27816|MAP4_HUMAN Microtubule-associated protein 4 OS=Homo sapiens OX=9606 GN=MAP4 PE=1 SV=3

MADLSLADALTEPSDIEGEIKRDFIATLEAEAFDDVVGETVGKTDYIPLLDVDEKTGNSESKKKPCSETSIQIEDTPSS
 KPTLLANGGHGVEGSDTTGSPTEFLEEKMAYQEYPNSQNWPEDTNFCFQPEQVVDPIQTDPFKMYHDDDLADLVF
 PSSATAADTSIFAGQNDPLKDSYGMSPCNTAVVPQGSVEALNSPHSESFVSPEAVAEPQPTAVPLELAKEIEMASE
 ERPPAQALEIMMGLKTTDMAPSKETEMALAKDMALATKTEVALAKDMESPTKLDVTLAKDMQPSMESDMALVKDME
 LPTEKEVALVKDVRWPTETDVSSAKNVVLPTEVAPAKDVTLLKETERASPIKMDLAPSKDMGPPKENKKETERAS
 PIKMDLAPSKDMGPPKENKIVPAKDLVLLSEIEVAQANDIISSTEISSAEKVALSSETEVALARDMTLPPETNVILTCKDA
 LPLEAEVAPVKDMAQLPETEIAPAKDVAPSTVKEVGLLKDMSPLESETEMALGKDVTPPETEVVLIKNVCLPPEMEVA
 LTEDQVPALKTEAPLAKDGVLTANNVTPAKDVPLSETATPVPIKDMIEAQTQKGISEDSHLESQDVGQSAAPTF
 MISPETVTGTGKKCSLPAEDSVLEKLGERKPCNSQPSLSSETSGIARPEEGRPVVSGTGNIDITPPNKLPPSPEK
 KTKPLATTQPAKTSTSKAKTQPTSLPKQAPATTIGGLNKKPMSLASGLVPAAPPKRPASARPSILPSKDVKPKPIAD
 AKAPEKRASPSKPASAPASRSGSKSTQTVAKTTTAAAVASTGPSSRSPSTLLPKKPTAIKTEGKPAEVKKMTAKSVP
 ADLSRPKSTSTSSMKTTTSLGTAPAAGVVP SRVKATPMPSRPSTTFFIDKKPTSAKPSSTTTPRLSRLATNTSAPDLK
 NVRSKVGSTENIKHQPGGGRKVEKKTEAAATTRKPESSNAVTKTAGPIASAQKQAPAGKVQIVSKKVSYSHIQSKCGS
 KDNKIHVPGGGNVQIQNKVDISKVSSKCGSKANIKHKPGGGDVKIESQKLNFKKAQAKVGSNDNVGHLPAAGGAVK
 TEGGGSEAPLCPGPPAGEEPAISEAAPEAGAPTSASGLNGHPTLSGGGDQREACTLDSIQIETSI

PHLDA1 Q8WV24 HUMAN Pleckstrin homology-like domain family A member 1

Nucleotide Sequence (Seq ID No. 12):

>P002080_Q305_Q305p3_PHLDA1_22822_Homo sapiens pleckstrin homology-like domain family A member
 1_BC018929.2_AAH18929.3_Q8WV24_0_0_780_0_777

ATGCTGGAGAGTAGCGGCTGCAAAGCGCTGAAGGAGGGCGTGCTGGAGAAGCGCAGCGACGGGTTGTTGCA
 GCTCTGGAAGAAAAAGTGTTCATCCTCACCGAGGAAGGGCTGCTGCTTATCCCGCCCAAGCAGCTGCAACAC

>sp|Q8WV24|PHLA1_HUMAN Pleckstrin homology-like domain family A member 1 OS=Homo sapiens OX=9606 GN=PHLA1 PE=1 SV=4

MRRAPAAERLLELGFPPRCGRQEPPFLGVTRGWGRWPIQKRREGARPVPFSESRQEDGRGPAARSSGTLWRIR
 TRLSLCRDPEPPPPPLCLLRVSLLCALRAGGRGSRWGEDGARLLLLPPARAAGNGEAEPSGGPSYAGRMLESSGCK
 ALKEGVLEKRSDGLLQLWKKKCCILTEEGLLLIPPKQLQHQQQQQQQQQQQQQQQPQGGAEPSPQSPGPAVASLE
 PPVKLKLHFSNMKTVDCVERKGKMYFTVYMAEGKEIDFRCPQDQGWNAEITLQMVQYKNRQAILAVKSTRQKQQ
 HLVQQQPPSPQPQPQLQPQPQPQPQPQPQSPQPQPQPQPQPQLHPYHPHPHPHSHPHSHPHHPHPHPH
 PHQIPHHPQPHSQPHGHRLLRSTNSA

>P000364_SIG_SIG1-1_PPM1A_5494_Homo sapiens protein phosphatase 1A (formerly 2C) magnesium-dependent alpha isoform tr BC026691.1 AAH26691.1 P35813 53422.23 0 1149 0 1146

ATGGGAGCATTTTAGACAAGCCAAAGATGGAAGAGATAATGCCAGGGGCAGGGTAATGGGTTGCGATATG
GGCTAAGCAGCATGCAAGGCTGGCGTGTGAAATGGAGGATGCACATACGGCTGTGATCGGTTTGCCAAGTGG
ACTTGAATCGTGGTCATTCTTTGCTGTGTATGATGGGCATGCTGTTCTCAGGTTGCCAAATACTGCTGTGAGC
ATTTGTTAGATCACATCACCAATAACCAGGATTTTAAAGGGTCTGCAGGAGCACCTTCTGTGAAAATGTAAAGA
ATGGAATCAGAACAGGTTTTCTGGAGATTGATGAACACATGAGAGTTATGTCAGAGAAGAAACATGGTGAGAT
AGAAGTGGGTCAACAGCTGTAGGTGTCTTAATTTCTCCCCAACATACTTATTTCATTAAGTGTGGAGACTCAAGA
GGTTTACTTTGTAGGAACAGGAAAGTTCATTTCTTCACACAAGATCACAAACCAAGTAATCCGCTGGAGAAAGAA
CGAATTCAGAATGCAGGTGGCTCTGTAATGATTCAGCGTGTGAATGGCTCTCTGGCTGTATCGAGGGCCCTTG
GGGATTTTGATTACAAATGTGTCCATGGAAGGTCTACTGAGCAGCTTGTCTACCAGAGCCTGAAGTCCAT
GATATTGAAAGATCTGAAGAAGATGATCAGTTCATTATCCTTGCATGTGATGGTATCTGGGATGTTATGGGAAAT
GAAGAGCTCTGTGATTTTGTAAAGATCCAGACTTGAAGTCACTGATGACCTTGAGAAAGTTTGAATGAAGTAGTC
GACACCTGTTTGTATAAGGGAAGTCGAGACAACATGAGTGTGATTTTGATCTGTTTTCCAAATGCACCCAAAGTA
TCGCCAGAAGCAGTGAAGAAGGAGGCAGAGTTGGACAAGTACCTGGAATGCAGAGTAGAAGAAATCATAAAGA
AGCAGGGGGAAGGCGTCCCCGACTTAGTCCATGTGATGCGCACATTAGCGAGTGAGAACATCCCAGCCTCC
CACCAGGGGGTGAATTGGCAAGCAAGAGGAATGTTATTGAAGCCGTTTACAATAGACTGAATCCTTACAAAAAT
GACGACACTGACTCTACATCAACAGATGATATGTGG

>sp|P35813|PPM1A_HUMAN Protein phosphatase 1A OS=Homo sapiens OX=9606 GN=PPM1A PE=1 SV=1

MGAFLDKPKMEKHNAQQGGNGLRYGLSSMQGWRVEMEDAHTAVIGLPSGLESWSFFAVYDGHAGSQVAKYCCF
HLLDHITNNQDFKGSAGAPSVENVKNGIRTGFLEIDEHMRVMSEKKHGADRSGSTAVGLISPQHTYFINGDSRGL
LCRNRKVHFFTQDHKPSNPLEKERIQNAGGSVMIQRVNGLAVSRALGDFDYKCVHGKGPTQLVSPEPEVHDIER
SEEDDQFIILACDGIWDMGNEELCDFVRSRLEVTDLEKVCNEVVDTCLYKGSRDNMSVILICFPNAPKVSPEAVKK
EAELDKYLECRVEEIIKKQGEGVPDLVHVMRTLASENIPSLPPGGELASKRNVIEAVYNRLNPYKNDDTDSTSTDDM
W

TCL1A	P56279	HUMAN T-cell leukemia/lymphoma protein 1A
--------------	--------	---

Nucleotide Sequence (Seq ID No. 14):

>P000179_CAN_CAN1-1_TCL1A_8115_Homo sapiens T-cell leukemia/lymphoma
1A_BC005831.2_AAH05831.1_P56279_0_0_345_0_342

ATGGCCGAGTGCCCCGACACTCGGGGAGGCAGTCACCGACCACCCGGACCGCCTGTGGGCCTGGGAGAAGTT
CGTGATTTGGACGAGAAGCAGCAGCGCCTGGCTGCCCTTAACCATCGAGATAAAGGATAGGTTACAGTTACGG
GTGCTCTTGCGTCGGGAAGACGTCGCTCGGGGAGGCCTATGACCCCCACCCAGATAGGCCCAAGCCTGCTG
CCTATCATGTGGCAGCTCTACCCTGATGGACGATACCGATCCTCAGACTCCAGTTTCTGGCGCTTAGTGATCCA
CATCAAGATTGACGGCGTGGAGGACATGCTTCTCGAGCTGCTGCCAGATGAC

Protein Sequence (Seq ID No. 30):

>sp|P56279|TCL1A_HUMAN T-cell leukemia/lymphoma protein 1A OS=Homo sapiens OX=9606 GN=TCL1A PE=1
SV=1

MAECPVLGEAVTDHPDLWAWKFKVYLDEKQHWLPLTIEIKDRLQLRVLLRREDVVLGRPMTPTQIGPSLLPIMWQ
LYPDGRYRSSDSSFWRVLYHIKIDGVEDMLELLPDD

UBE2I P63279 HUMAN SUMO-conjugating enzyme UBC9

Nucleotide Sequence (Seq ID No. 15):

>P001344_CAG_CAGp1_UBE2I_7329_Homo sapiens ubiquitin-conjugating enzyme E2I (UBC9 homolog yeast)
transcript variant 1_BC000427.2_AAH00427.1_P50550_0_0_477_0_474

ATGTCGGGGATCGCCCTCAGCAGACTCGCCAGGAGAGGAAAGCATGGAGGAAAGACCACCCATTTGGTTTC
GTGGCTGTCCCAACAAAAATCCCGATGGCAGCATGAACCTCATGAAGTGGGAGTGCGCCATTCCAGGAAAGA
AAGGGACTCCGTGGGAAGGAGGCTTGTAACTACGGATGCTTTTCAAAGATGATTATCCATCTTCGCCACCA
AAATGTAAATTCGAACCACCATTTTACCCGAATGTGTACCCTTCGGGGACAGTGTGCCTGTCCATCTTAGAG
GAGGACAAGGACTGGAGGCCAGCCATCACATCAAACAGATCCTATTAGGAATACAGGAACCTCTAAATGAACC
AAATATCCAAGACCCAGCTCAAGCAGAGGCCTACACGATTTACTGCCAAACAGAGTGGAGTACGAGAAAAGG
GTCCGAGCACAAGCCAAGAAGTTTGCGCCCTCA

Protein Sequence (Seq ID No. 31):

>sp|P63279|UBC9_HUMAN SUMO-conjugating enzyme UBC9 OS=Homo sapiens OX=9606 GN=UBE2I PE=1
SV=1

MSGIALSRLAQRKAWRKDHPFGFVAVPTKNPDGTMNLMNWECAIPGKKGTPWEGGLFKLRMLFKDDYPSSPPKC
KFEPPLFHPNVYPSGTVCLSILEDKDWRPATIKQILLGIQELLNEPNIQDPAQAEAYTIYCNQNRVEYEKRVRAQAKKF
APS

YARS P54577 HUMAN Tyrosine--tRNA ligase, cytoplasmic

Nucleotide Sequence (Seq ID No. 16):

>P001370_CAG_CAGp2_YARS_8565_Homo sapiens tyrosyl-tRNA
synthetase_BC004151.2_AAH04151.1_P54577_0_0_1587_0_1584

ATGGGGGACGCTCCCAGCCCTGAAGAGAACTGCACCTTATACCCGGAACCTGCAGGAGGTTCTGGGGGAA
GAGAAGCTGAAGGAGATACTGAAGGAGCGGGAACCTTAAATTTTACTGGGGAACGGCAACCACGGGCAAACCAC
ATGTGGCTTACTTTGTGCCCATGTCAAAGATTGCAGACTTCTTAAAGGCAGGGTGTGAGGTAACAATTCTGTTTG
CGGACCTCCACGCATACCTGGATAACATGAAAGCCCCATGGGAACCTTCTAGAACTCCGAGTCAGTTACTATGAG
AATGTGATCAAAGCAATGCTGGAGAGCATTGGTGTGCCCTTGAGAAGCTCAAGTTCATCAAAGGCACTGATTA
CCAGCTCAGCAAAGAGTACACACTAGATGTGTACAGACTCTCCTCCGTGGTCACACAGCAGCAGATTCCAAGAAG
GCTGGAGCTGAGGTGGTAAAGCAGGTGGAGCACCCTTTGCTGAGTGGCCTCTTATACCCCGGACTGCAGGCTT
TGGATGAAGAGTATTTAAAGTAGATGCCCAATTTGGAGGCATTGATCAGAGAAAGATTTTACCTTTGCAGAGA
AGTACCTCCCTGCACTTGGCTATTCAAAACGGGTCCATCTGATGAATCCTATGGTTCCAGGATTAACAGGCAGC
AAAATGAGCTCTTCAGAAGAGGAGTCCAAGATTGATCTCCTTGATCGGAAGGAGGATGTGAAGAAAAAAGTAA
GAAGGCCTTCTGTGAGCCAGGAAATGTGGAGAACAATGGGGTTCTGTCTTCATCAAGCATGTCTTTTTTCCCC
TTAAGTCCGAGTTTGTGATCCTACGAGATGAGAAATGGGGTGGAAACAAAACCTACACAGCTTACGTGGACCTG

GAAAAGGACTTTGCTGCTGAGGTTGTACATCCTGGAGACCTGAAGAATTCTGTTGAAGTCGCACTGAACAAGTT
 GCTGGATCCAATCCGGGAAAAGTTTAATACCCCTGCCCTGAAAAAACTGGCCAGCGCTGCCTACCCAGATCCC
 TCAAAGCAGAAGCCAATGGCCAAAGGCCCTGCCAAGAATTGAGAACCAGAGGAGGTCATCCCATCCCGGCTGG
 ATATCCGTGTGGGGAAAATCATCACTGTGGAGAAGCACCCAGATGCAGACAGCCTGTATGTAGAGAAGATTGA
 CGTGGGGGAAGCTGAACCACGGAAGTGTGGTGAAGCGGCCTGGTACAGTTCGTGCCCAAGGAGGAAGTGCAGGA
 CAGGCTGGTAGTGGTGTGTGCAACCTGAAACCCAGAGATGAGAGGAGTCGAGTCCCAAGGCATGCTTCTG
 TGTGCTTCTATAGAAGGGATAAACCGCCAGGTTGAACCTCTGGACCCTCCGGCAGGCTCTGCTCCTGGTGAGC
 ACGTGTGTGTGAAGGGCTATGAAAAGGGCCAACCCAGATGAGGAGCTCAAGCCCAAGAAGAAAGTCTTCGAGAA
 GTTGCAGGCTGACTTCAAAATTTCTGAGGAGTGCATCGCACAGTGGAAGCAAACCAACTTCATGACCAAGCTGG
 GCTCCATTTCTGTAAATCGCTGAAAGGGGGGAACATTAGCC

Protein Sequence (Seq ID No. 32):

>sp|P54577|SYYC_HUMAN Tyrosine--tRNA ligase, cytoplasmic OS=Homo sapiens OX=9606 GN=YARS PE=1
 SV=4

MGDAPSPEEKLHLITRNLQEVLGEEKLKEILKERELKIYWGTTGKPHVAYFVPMSKIADFLKAGCEVTILFADLHAYL
 DNMKAPWELLELRVSYENVIKAMLESIGVPLEKLKFIKGTDYQLSKEYTLDVYRLSSVVTQHDSKKAGAEVVKQVEH
 PLLSGLLYPGLQALDEEYLVDAQFGGIDQRKIFTFAEKYLPALGYSKRVHLMNPMVPGLTGSKMSSEEEESKIDLLD
 RKEDVKKKLKFAFCEPGENVENNGVLSFIKHVLFPLKSEFVILRDEKWGGNKTYTAYVDLEKDFAAEVVHPGDLKNSV
 EVALNKLLDPIREKFNTPALKKLASAAYPDPSKQKPMAGPAKNSEPEEVIPSRLDIRVGKIITVEKHPDADSLYVEKID
 VGEAEPRTVVSGLVQFVPKEELQDRLVVVLCNLKPQKMRGVESQGMLLCASIEGINRQVEPLDPPAGSAPGEHVFV
 KGYEKGQPDEELPKKKVFEKLQADFKISEECIAQWKQTNFMTKLGSISCKSLKGGNIS

Claims

1. A method for determining the health of an individual from a sample extracted from that individual, comprising the steps of:
 - 5 (i) testing the sample for the presence of biomarkers specific for health;
 - (ii) determining whether the subject is healthy, is of intermediate health, or is unhealthy, based on the detection of said biomarkers;characterised in that the biomarkers are autoantibodies to antigens comprising AURKA, FEN1, GLRX3, PHLDA1, PPM1A, FKBP3, CD96 and MAPK13.
- 10 2. The method according to claim 1 wherein the antigens further comprise one or more of UBE2I, AAK1, YARS, ASPSCR1, CASP10, FHOD2, TCL1A and MAP4.
- 15 3. The method according to claim 1 or 2 wherein the antigens are exposed to a sample extracted from a person, such that autoantibody biomarkers from the sample may bind to the antigens.
4. The method according to claim 3 wherein the antigens are subsequently exposed
20 to a fluorescently-tagged secondary antibody to allow the amount of any autoantibodies from the sample bound to the antigens to be determined.

5. The method according to claim 4 wherein the health status of the person corresponds to the relative or absolute amount of autoantibodies from the sample specifically binding to the antigens.
- 5 6. 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.
7. The method according to any preceding claim wherein the steps are performed *in vitro*.
10
8. The method according to any preceding claim wherein the method comprises detecting upregulation/downregulation of one or more biomarkers.
- 15 9. The method according to any preceding claim wherein PHLDA1 and CD96 correspond to healthy, AURKA, FEN1, CASP10 and AAK1 correspond to intermediate health, and UBE2I, YARS, ASPSCR1, FHOD2, TCL1A, MAP4, MAPK13, FKBP3, PPM1A and GLRX3 correspond to unhealthy.
- 20 10. The method according to any preceding claim wherein the individual is elderly.
11. The method according to any preceding claim wherein the antigens are biotinylated proteins.

12. The method according to claim 11 wherein each biotinylated protein is formed from a Biotin Carboxyl Carrier Protein folding marker which is fused in-frame with a protein.
- 5 13. The method according to claim 11 or 12 wherein the biotinylated proteins are bound to a streptavidin-coated substrate.
14. The method according to claim 13 wherein the substrate comprises a hydrogel-forming polymer base layer.
- 10 15. A method for manufacturing a kit for determining the health of an elderly individual from a sample extracted from that individual, 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 said antigen and expressing the
- 15 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 antigens on the panel can be determined by exposing the substrate to the sample
- 20 and measuring the response;
- characterised in that the antigens comprise UBE2I, AAK1, YARS, ASPSCR1, CASP10, FHOD2, TCL1A, MAP4, MAPK13, CD96, FKBP3, PPM1A, PHLDA1, GLRX3, FEN1 and AURKA.

16. A composition comprising a panel of antigens for determining the health of an elderly individual, characterised in that the antigens comprise MAPK13, CD96, FKBP3, PPM1A, PHLDA1, GLRX3, FEN1 and AURKA, and optionally comprising one or more of UBE2I, AAK1, YARS, ASPSCR1, CASP10, FHOD2, 5 TCL1A, and MAP4.
17. A composition according to claim 16 wherein the antigens are biotinylated proteins.
- 10 18. A composition according to any of claims 16-17 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 determine the health status of the patient.
- 15 19. A composition comprising a panel of autoantibody biomarkers for determining the health status of an elderly patient:
- wherein the level of one or more autoantibody biomarkers are measured in a sample from the patient;
- characterised in that the one or more autoantibody biomarkers are selected from autoantibodies specific for one or more of the following antigens: MAPK13, 20 CD96, FKBP3, PPM1A, PHLDA1, GLRX3, FEN1 and AURKA;
- and optionally comprising one of more of the following antigens: UBE2I, AAK1, YARS, ASPSCR1, CASP10, FHOD2, TCL1A, and MAP4.