

Serum Antibody Repertoire Profiling Using In Silico Antigen Screen

Xinyue Liu¹, Qiang Hu², Song Liu², Luke J. Tallo¹, Lisa Sadzewicz¹, Cassandra A. Schettine³, Mikhail Nikiforov⁴, Elena N. Klyushnenkova⁵, Yuriy Ionov^{3*}

1 Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, Maryland, United States of America, **2** Department of Biostatistic and Bioinformatics, Roswell Park Cancer Institute, Buffalo, New York, United States of America, **3** Department of Cancer Genetics, Roswell Park Cancer Institute, Buffalo, New York, United States of America, **4** Department of Cell Stress Biology, Roswell Park Cancer Institute, Buffalo, New York, United States of America, **5** Department of Surgery, University of Maryland School of Medicine, Baltimore, Maryland, United States of America

Abstract

Serum antibodies are valuable source of information on the health state of an organism. The profiles of serum antibody reactivity can be generated by using a high throughput sequencing of peptide-coding DNA from combinatorial random peptide phage display libraries selected for binding to serum antibodies. Here we demonstrate that the targets of immune response, which are recognized by serum antibodies directed against sequential epitopes, can be identified using the serum antibody repertoire profiles generated by high throughput sequencing. We developed an algorithm to filter the results of the protein database BLAST search for selected peptides to distinguish real antigens recognized by serum antibodies from irrelevant proteins retrieved randomly. When we used this algorithm to analyze serum antibodies from mice immunized with human protein, we were able to identify the protein used for immunizations among the top candidate antigens. When we analyzed human serum sample from the metastatic melanoma patient, the recombinant protein, corresponding to the top candidate from the list generated using the algorithm, was recognized by antibodies from metastatic melanoma serum on the western blot, thus confirming that the method can identify autoantigens recognized by serum antibodies. We demonstrated also that our unbiased method of looking at the repertoire of serum antibodies reveals quantitative information on the epitope composition of the targets of immune response. A method for deciphering information contained in the serum antibody repertoire profiles may help to identify autoantibodies that can be used for diagnosing and monitoring autoimmune diseases or malignancies.

Citation: Liu X, Hu Q, Liu S, Tallo LJ, Sadzewicz L, et al. (2013) Serum Antibody Repertoire Profiling Using In Silico Antigen Screen. PLoS ONE 8(6): e67181. doi:10.1371/journal.pone.0067181

Editor: Barbara Ensoli, Istituto Superiore di Sanità, Italy

Received: March 18, 2013; **Accepted:** May 15, 2013; **Published:** June 27, 2013

Copyright: © 2013 Liu et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the gift from the Hubbel Family foundation, the IRG -0197-07 grant, the NCI grant P30CA016056, the KCA126849A UMGCC Paul Calabresi Award for Clinical Oncology and UMGCC Pilot and Grant Program to Support Developing Shared Services. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: yuriy.ionov@roswellpark.org

Introduction

The repertoires of serum antibody specificities contain information on the state of health and disease of individual. For example, circulating serum autoantibodies against self-antigens can serve as indicators of autoimmune diseases or of immune response against malignancies [1]. The information contained in the individuals' sera can be investigated using methods for global analysis of serum antibody repertoires.

Random peptide phage display libraries (RPPDL) are widely used for mapping epitopes on defined antigens. [2]. Epitopes recognized by monoclonal as well as by polyclonal antibodies can be identified by the biopanning procedure, an affinity selection for binding to antibodies of phage displayed peptides, followed by sequencing of individual phage DNA [3], [4]. Since the length of a consensus sequence that mimics the core epitope recognized by antibody is frequently in the range from 4 to 6 amino acids [5,6], and since all possible 6-mer amino acid permutations can be represented by 6.4×10^7 sequences, this implies that all possible linear core epitopes of the human proteome can be represented by the commercially available library of random heptapeptides of the

complexity of approximately 109 different sequences. The necessity to sequence individual phage clones until recently limited the application of the RPPDL to identifying epitopes on a defined antigen. With the advance of next generation sequencing (NGS), the phage displayed peptides affinity selected for binding to serum antibodies can be used for generating global profiles of serum antibody specificities [7]. The feasibility of using RPPDL and NGS to analyze antibody specificities of polyclonal sera was demonstrated recently by profiling polyclonal sera derived from HIV infected individuals [8]. The authors demonstrated that a fraction of the most abundant peptides selected for binding to IgG antibodies of HIV positive sera could be aligned by a BLASTP analysis to the HIV protein thus indicating some HIV specificity.

The drawback of using RPPDL for global profiling of serum antibody reactivity is the lack of information on the identities of the real antigens that are mimicked by the antibody-binding peptides. Identifying the targets of antibody immune response using peptide mimotopes is a difficult task, since most of epitopes recognized by antibodies are conformational and discontinuous and only a small fraction of antibodies are raised against linear or continuous epitopes. Furthermore, identifying even linear epitopes

of unknown targets is also complicated since a BLAST search of protein databases retrieves hundreds of proteins that have a sequence match to a short peptide. In this work we demonstrate an algorithm, which we term in Silico Antigen Screen (SAS), for analyzing peptide profiles of serum antibody specificities generated by RPPDL panning and NGS. The algorithm can be used for identifying protein autoantigens if the unknown targets are recognized by antibodies directed against linear epitopes.

Results

Profiling the Immune Response in Mice Immunized With Human Proteins

Peptides selected for binding to serum antibodies in biopanning experiments can mimic linear (continuous) epitopes and conformational (discontinuous) epitopes of protein antigens. They also can mimic non-protein epitopes, such as the carbohydrate structures of glycoproteins or glycolipids, nucleic acids or other chemical structures. We hypothesized that the protein targets of humoral immune response can be identified using peptide profiles of serum antibody repertoires generated by NGS, and that peptides that mimic linear epitopes of proteins will be present among the profile-making peptides. Since for any short peptide the BLAST search of protein databases retrieves about a hundred of proteins that have matches to the peptide sequence, it is practically impossible to determine what kind of epitope the peptide was mimicking. However, the BLAST search for a large number of peptides can retrieve proteins that have multiple matches to different peptides. The probability for a protein to have multiple matches to different peptides due to a chance is proportional to the length of the protein. Therefore, the real protein targets of humoral immune response recognized by antibodies directed at the linear epitopes are likely to have the higher number of matches per protein length than the proteins that have matches to peptides due to a chance.

To test this hypothesis, we used the sera of mice immunized with human proteins, prostate specific antigen (PSA) or prostatic acid phosphatase (PAP) in Complete Freund's adjuvant. All the sera had high (>10,000) titers against the whole PAP or PSA proteins in the ELISA assay (data not shown). Our goal was to test whether it was possible to identify the proteins used for immunization by analyzing peptide profiles of serum antibody repertoires. The peptide profiles for the IgG antibodies from the four anti-PSA sera, four anti-PAP sera and two sera from naive mice were processed using two rounds of panning as described in the "Materials and Methods" section. The peptide-coding DNA fragments from the RPPDL selected for binding to the IgG antibodies were sequenced using Illumina HiSeq 2000 and the DNA sequences were translated into the peptide sequences.

For each serum sample from mice immunized with the PSA antigen, we selected the 500 most abundant peptides that were not shared by the mice immunized with the PAP antigen or the naive mice. Excluding the peptides shared by PAP-immunized mice and by naive mice minimized the input of the response to the components of adjuvant used for immunizations and enriched the peptide list with peptides related to a specific antigen. Similarly, for each serum sample from mice immunized with the PAP antigen we selected 500 the most abundant peptides that were not shared by the PSA immunized mice or by naive mice (Table S1).

The search of the refseq_protein database for the homo sapiens (taxid:9606) using default parameters for the Blastp (protein-protein BLAST) retrieved for each peptide sequence the list of 100 proteins ranked by the decrease in the maximum score or by the increase in the expected threshold value. We tested the following

2-step algorithm for distinguishing the real antigens recognized by serum antibodies from the 'sea' of proteins retrieved from the database due to a chance. In the first step, we selected a limited number of the most abundant peptides and used BLAST homology search against human protein database to identify proteins that contain matches to at least two different peptides. Selecting only the most abundant peptides for this analysis would allow identifying proteins that are recognized by serum antibodies with the highest titer. Such antibodies would be easier to detect for independent confirmation of the immune response using a conventional method such as ELISA.

The number of proteins to be selected for each peptide in the first step can be regulated by varying the threshold parameters of BLAST search such as expected value (E-value) or maximal score. Lowering the E-value or increasing the maximal score allows selecting the lower number of proteins but with higher degree of homology to peptides. We chose to use the maximal score equal 18.5 as a threshold parameter, which corresponded to the match between a peptide and a protein of a stretch of 5 amino acids. For each peptide, the BLAST search retrieved, on average, approximately thirty proteins with the maximal score more than 18.5. All proteins retrieved by the BLAST search, that satisfied the threshold parameter were combined in one list. This protein list was analyzed to select proteins which were present in the list more than once. The selected proteins were ranked by the number of matching peptides per protein length. The proteins with the highest number of matching peptides per protein length were further analyzed in the second step. In the second step, we used the Specialized BLAST tool 'Align two (or more) sequences using BLAST (bl2seq)' to analyze all the 500 peptides in order to identify for each selected protein all the peptides with the homologies. The less stringent threshold parameters of the bl2seq allow identifying also the peptides with lower degree of homology to proteins, which could be missed in the first step of the algorithm. The candidate targets of immune recognition were selected by ranking according to the final score calculated for each selected protein as the sum of scores for all the peptides with the homology to the protein divided by the length of the protein.

When we applied this algorithm to the lists of 500 peptides for the anti-PSA sera, and used a BLAST search against human refseq_protein database (first step), we found that none of the peptides retrieved the PSA protein. This may indicate that the humoral response to the human PSA protein in mice is limited to conformational epitopes. In contrast, the analysis of PAP-specific sera produced multiple hits. Three anti-PAP sera samples (PAP1, PAP2 and PAP3) produced the peptide lists with multiple peptides that retrieved the PAP protein by BLAST search against human refseq_protein database. One anti-PAP serum (PAP4) produced the 500 peptide list that contained only one peptide which retrieved the PAP protein by BLAST search against human refseq_protein database in the first step.

The first 120 most abundant peptides of the PAP1 list contained 2 peptides that retrieved the PAP protein with the maximal score ≥ 18.5 . Besides the 2 isoforms of the PAP proteins NP_001090.2 and NP_001127666.1, there were 194 other proteins that had multiple matches to different peptides. After ranking the proteins according to the initial score calculated as a number of matching peptides to the protein divided by the length of the protein, the PAP isoforms NP_001090.2 and NP_001127666.1 occupied the 56th and 66th positions respectively (Table S2A).

In the second step of the algorithm, we run the bl2seq for the all 500 peptides of the PAP1 list against each of the first top 100 proteins with multiple matches identified in the first step. For this purpose, we wrote software that allows performing bl2seq off line

Table 1. Candidate antigens for mouse sera.

Rank	Proteins selected for PAP1 antiserum	Protein length(aa)	Initial number of matches	Initial score	Sum of overall scores	Final score	Major epitope score
1	NP_065117.1 claudin-2	230	2	0.0086	600.1	2.60913	147.9
2	XP_003119043.1 PREDICTED: hypothetical protein LOC100506191	118	2	0.0169	266.2	2.255932	126.2
3	NP_653235.1 lysozyme-like protein 4 precursor	146	2	0.0136	302.4	2.071233	163
4	NP_001090.2 <i>prostatic acid phosphatase isoform PAP precursor</i>	386	2	0.0051	769.5	1.93523	376
5	NP_001181944.1 T-cell surface glycoprotein CD4 isoform	185	2	0.0108	361.9	1.956216	218.5
6	NP_001098018.1 uncharacterized protein	102	3	0.0294	192.9	1.891176	76.3
7	NP_001127666.1 <i>prostatic acid phosphatase isoform TM-PAP precursor</i>	418	2	0.0047	777.3	1.859569	376
8	NP_057574.2 protein FAM178B isoform B	119	2	0.0168	210.2	1.766387	0
9	NP_001172078.1 putative claudin-24	220	2	0.009	369.4	1.679091	0
10	NP_000139.1 galactoside 2-alpha-L-fucosyltransferase 1	365	2	0.0054	594	1.627397	0
Proteins selected for PAP2 antiserum							
1	NP_001230678.1 modulator of retrovirus infection homolog isoform 2	69	5	0.0724	1260.4	18.26667	1245
2	NP_001011700.2 mitochondrial coiled-coil domain protein 1 precursor	119	3	0.0252	1414.3	11.88487	1350.5
3	NP_001092926.2 neurotrophin W preproprotein	165	2	0.0121	1425	8.636364	1281.6
4	NP_076938.2 modulator of retrovirus infection homolog isoform 1	157	5	0.0318	1293.6	8.23949	1245
5	NP_000396.2 ganglioside GM2 activator isoform 1 precursor	193	2	0.0103	1461.6	7.573057	1236.7
6	NP_065117.1 claudin2	230	2	0.0086	1658.4	7.210435	1154.9
7	NP_997394.1 uncharacterized protein C9orf139	190	3	0.0157	1290.5	6.792105	984
8	NP_114103.2 voltage-dependent calcium channel gamma-6 subunit isoform c	189	3	0.0158	1262.2	6.678307	1157.5
9	NP_942567.1 probable alpha-ketoglutarate-dependent dioxygenase ABH6 isoform 1	161	2	0.0124	1063.8	6.607453	989.5
10	NP_443183.1 deoxynucleotidyl-transferase terminal-interacting protein 1	329	9	0.0273	2045.1	6.216109	1866.1
11	NP_665814.1 voltage-dependent calcium channel gamma-6 subunit isoform b	214	3	0.014	1277.4	5.969159	1157.5
12	NP_001090.2 <i>prostatic acid phosphatase isoform PAP precursor</i>	386	7	0.0181	2245.2	5.81658	1886.9
13	NP_005182.1 T-lymphocyte activation antigen CD80 precursor	288	2	0.0069	1666.3	5.785764	1415.9
14	NP_001010905.1 UPF0762 protein C6orf58 precursor	330	3	0.009	1781.4	5.398182	1276.3
15	NP_001127666.1 <i>prostatic acid phosphatase isoform TM-PAP precursor</i>	418	7	0.0167	2245.2	5.371292	1886.9
Proteins selected for PAP3 antiserum							
1	NP_149101.1 gamma-glutamylaminocyclotransferase	153	5	0.0326	1054.8	6.894118	930.3
2	NP_001165160.1 AMME syndrome candidate gene 1 protein isoform 3	210	4	0.019	1032.3	4.915714	897.8
3	NP_001025046.1 regulator of G-protein signaling 7-binding protein	257	5	0.0194	1254.4	4.880934	1043.1
4	NP_060436.4 vacuolar protein sorting-associated protein 37C	355	2	0.0056	1502.4	4.232113	835.1
5	NP_113633.2 AMMECR1-like protein	310	2	0.0064	1053.7	3.399032	876.9
6	NP_003921.2 src kinase-associated phosphoprotein 2	359	2	0.0055	1207.6	3.363788	949.8

Table 1. Cont.

Rank	Proteins selected for PAP1 antiserum	Protein length(aa)	Initial number of matches	Initial score	Sum of overall scores	Final score	Major epitope score
7	NP_004995.1 NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 186, mitochondrial precursor	186	3	0.0161	618.1	3.323118	243.2
8	NP_001090.2 prostatic acid phosphatase isoform PAP	386	6	0.0155	1159.5	3.003886	906.5
9	NP_001504.2 maleylacetoacetate isomerase isoform 3	161	3	0.0186	481.7	2.991925	374
10	NP_005986.2 T-box transcription factor TBX10	385	2	0.0051	1127.3	2.928052	857.6
11	NP_078870.1 pantothenate kinase	370	3	0.0081	1044.1	2.821892	876.9
12	NP_001127666.1 prostatic acid phosphatase isoform TM-PAP precursor	418	6	0.0143	1169.2	2.797129	906.5
13	NP_665878.2 maleylacetoacetate isomerase isoform 2	174	3	0.0172	478.3	2.748851	374
14	NP_001180318.1 putative E3 ubiquitin-protein ligase UNKL isoform	182	3	0.0164	476.2	2.616484	0
15	NP_061867.1 F-box only protein 42	717	4	0.0055	1764.7	2.461227	1024.7

The table shows the top candidate antigens selected for the antibodies of the PAP1, PAP2 and PAP3 antisera by sorting the data generated by processing the results of BLAST database search for the corresponding peptide sequences.
doi:10.1371/journal.pone.0067181.t001

for the large number of peptides against large number of proteins. These analyses identified all the peptides in the list with matches to selected proteins. The proteins were then ranked according to the final score that was calculated as the sum of the scores for the all peptides with matches against the protein divided by the length of the protein. Such analyses ranked the PAP isoforms NP_001090.2 and NP_001127666.1 at the 4th and the 7th positions, respectively (Table S2A and Table 1).

The Blast2seq analysis for the 500 PAP1 peptides against PAP NP_001090.2 or NP_001127666.1 isoforms identified 48 peptides with matches to the PAP protein. Out of 48 peptides, 20 were related to the sequence NFTLPSWA located at the amino acid positions 220–227 of the PAP protein. Figure 1 shows the location of peptide matches on the sequence of the NP_001090.2 PAP variant. Three other proteins, that were ranked higher than the NP_001090.2 PAP isoform also contained the sequences related to the NFTLPSWA sequence and had multiple matches to related peptides. However, the sum of scores for the peptides related to the NFTLPSWA sequence of the PAP isoforms was the highest among the sums of scores for the related epitopes in other proteins (Table 1).

The list of top 500 candidate peptides for the PAP2 serum was significantly enriched with peptides with homology to the PAP proteins sequence. The BLAST search against human refseq_-protein for the first most abundant 120 peptides of the PAP2 list contained 7 peptides that retrieved the PAP protein isoforms with the maximal score ≥ 18.5 . Out of 500 peptides of the PAP2 list analyzed by the blast2seq, 136 peptides matched the PAP isoforms, of which 110 peptides were related to the NFTLPSWA sequence, which was the same sequence related to the peptides for the PAP1 sample. Besides 2 isoforms of the PAP protein, there were 490 other proteins that had multiple matches to different peptides. After ranking the top 100 proteins identified in the first step of analysis using the blast2seq, the PAP isoforms occupied the 12th and the 15th positions (Table S2B and Table 1). The proteins that were ranked higher than the PAP isoforms all had the sequences related to the NFTLPSWA sequence. The higher final scores of these proteins were caused by the input from peptides that have matches to these sequences. As was the case for the PAP1 serum, the sum of scores for the PAP2 peptides related to the NFTLPSWA sequence of the PAP isoforms was higher than the sum of scores for the peptides that have matches to the related sequences in other proteins.

The BLAST search of the first 120 peptides of the PAP3 500 peptide list also contained PAP isoforms as well as 444 proteins that had multiple matches to different peptides with the threshold maximal score ≥ 18.5 . After ranking the top 100 proteins identified in the first step of analysis using the blast2seq, the PAP isoforms occupied the 8th and the 12th positions (Table S2C and Table 1). Out of 500 peptides of the PAP3 list, the 69 and 70 peptides had the matches to the two respective PAP isoforms, of which 48 peptides were related to the amino acids 336–342 QHEPYPL sequence of the PAP protein. The proteins that were ranked higher than the longest NP_001127666.1 PAP isoform also contained the motifs related to the QHEPYPL sequence. The sum of the scores for the QHEPYPL epitope of the PAP isoforms was ranked as 5th among the similar epitopes in other proteins. Validating the SAS Results of Mouse Sera Profiling by Bioinformatics Analyses.

The BLAST-based analysis of serum antibody repertoire described above can identify the linear epitopes but will miss mimotopes that mimic the non-linear protein or carbohydrate epitopes. To identify all the motifs recognized by serum antibodies that represent linear and non-linear epitopes we analyzed the lists

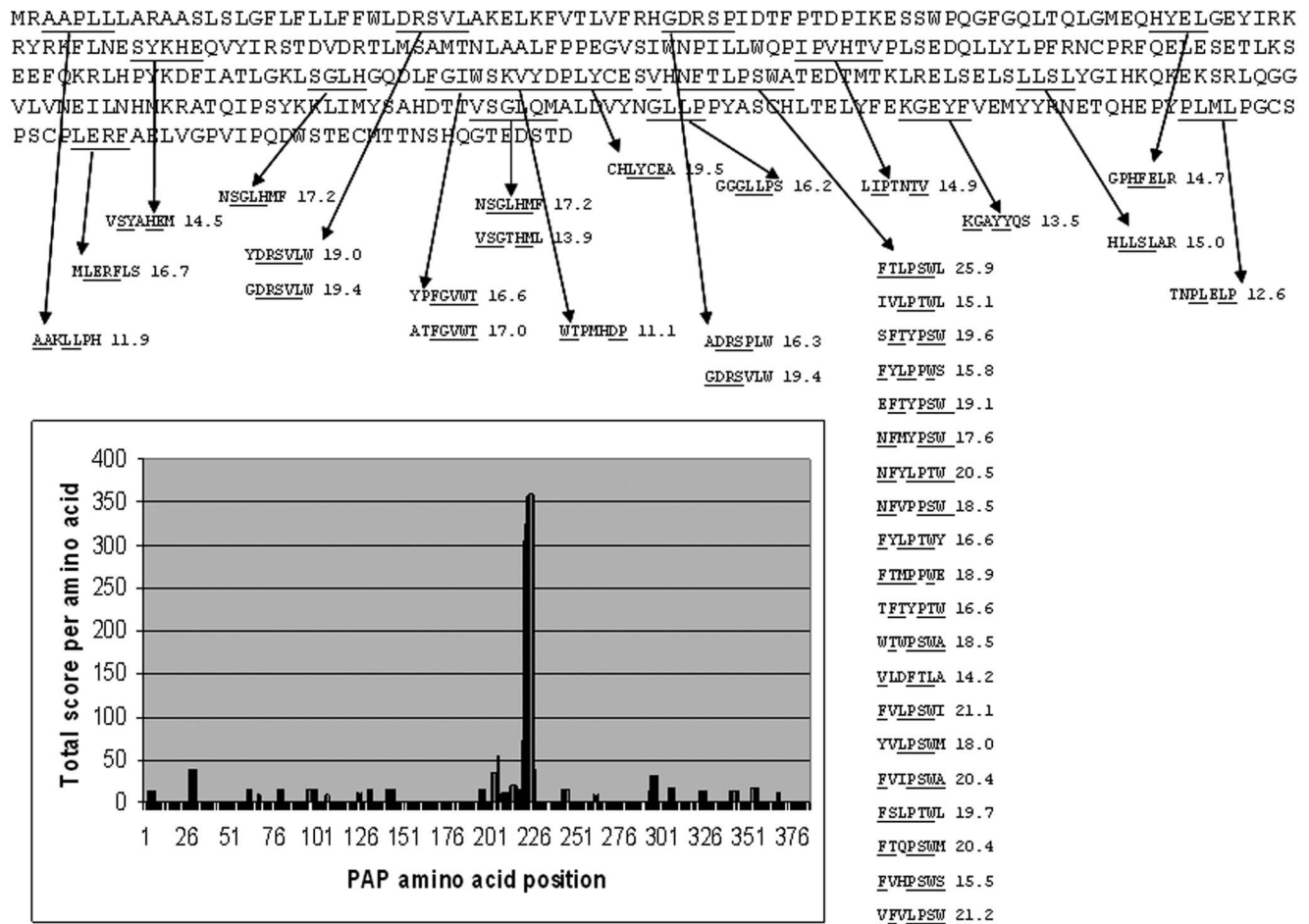


Figure 1. Epitope profile of the PAP protein generated by PAP1 mouse antiserum. (A) The NP_001090.2 PAP variant amino acid sequence and the 40 matching peptides with the highest match scores generated by bl2seq program are shown. The sequences of the protein matching to peptides and the corresponding peptide sequences that match to protein are underlined. (B) The graphic shows the distribution of the total score of the peptide matching over the sequence of the PAP protein.
 doi:10.1371/journal.pone.0067181.g001

of peptides using MEME software available online <http://meme.sdsc.edu/meme/cgi-bin/meme.cgi>. Figure 2 shows the three motifs identified by MEME for the each PAP antiserum. For the PAP1, PAP2 and PAP3 antisera, the most represented motifs were related to the NFTLPSWA and QHEPYPL sequences of the PAP protein. For the PAP4 serum that did not produce significant matches to the PAP protein by BLAST analysis, all three motifs were represented equally.

We also used MEME software to analyze the sequences of proteins that had been selected as the candidate antigens for the PAP1, PAP2 and the PAP3 sera based on their higher final score compared to the PAP isoforms. The MEME analysis identified the same motifs related to the NFTLPSWA and the QHEPYPL sequences of the PAP protein (Figure 3), suggesting that the PAP1, PAP2 and PAP3 sera could cross-react with these proteins.

We also analyzed the PAP protein sequence using available online tool for linear epitope prediction <http://sysbio.unl.edu/SVMTriP/prediction.php>. The software based on the Support Vector Machine algorithm predicted existence of three linear epitopes within the PAP sequence (Table 2). Although the NFTLPSWA sequence was not included in any of the predicted epitopes, the epitope predicted with the highest score included the QHEPYPL sequence recognized by the PAP3 antiserum. Another

predicted epitope contained the match to the peptide NTTNSHG from the PAP3 list, which retrieved the PAP isoforms by the BLAST searching of the peptide sequence against human refseq_protein database.

Validating the SAS Results of Mouse Sera Profiling Using the Anti-peptide ELISA

To prove that the sequences identified by the SAS method represent the real linear epitopes recognized by serum antibodies, we analyzed PAP-specific antisera by ELISA using peptide library consisting of 20-mers that overlap by 10 amino acids and span the mature human PAP amino acid sequence. As shown in Figure 4, PAP1 and PAP2 antisera recognized 20-mer peptide containing the NFTLPSWA sequence, and PAP3 antiserum recognized the 20-mer peptides containing the QHEPYPL sequence. The analysis of PSA-specific antisera by ELISA using the overlapping peptides representing the PSA proteins did not identify any peptide that had a signal significantly higher than that for the background binding (not shown) thus confirming the lack of recognition of linear epitopes of the PSA in the analyzed PSA-specific antisera.

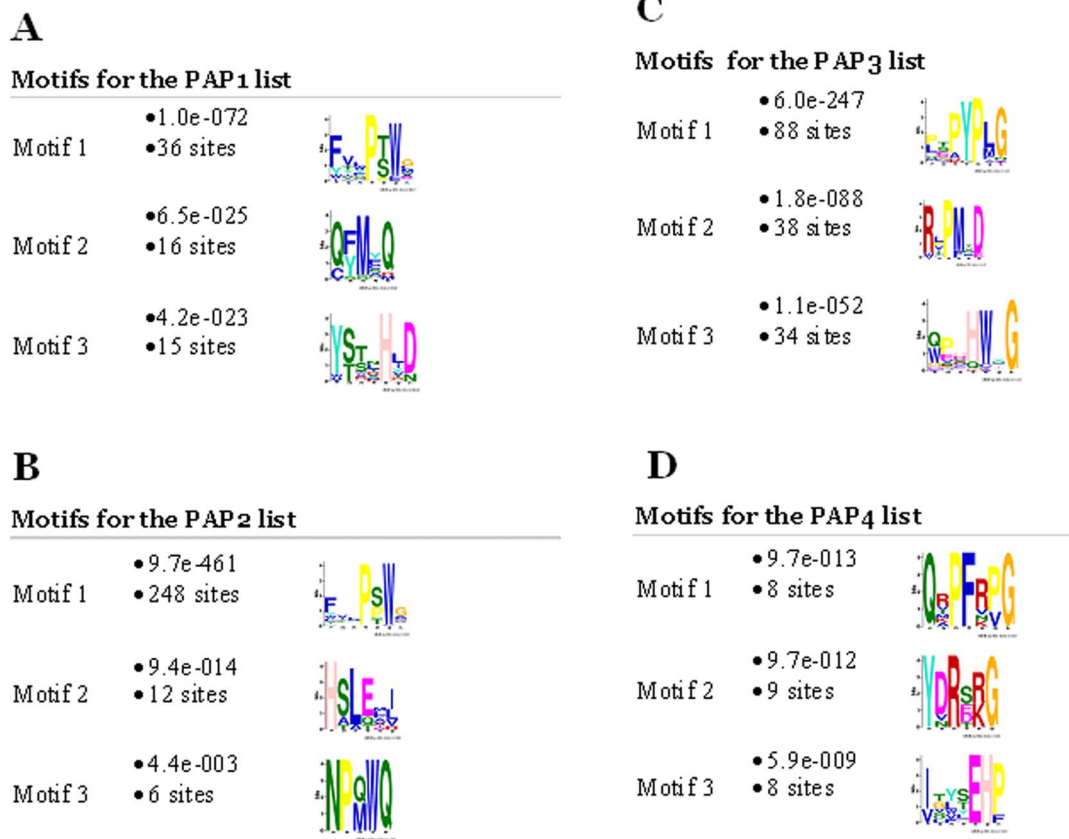


Figure 2. Motifs identified by MEME software for the 500 peptide lists for the PAP1, PAP, PAP3 and PAP4 antisera.
doi:10.1371/journal.pone.0067181.g002

Analyzing Antibody Repertoire of Human Serum

The described analysis of mouse sera using SAS demonstrates that the method can identify the antigen used for immunization, when the immune response involves recognition by serum antibodies of linear epitopes of the antigen. Next we wanted to evaluate the capability of the method to identify autoantigens recognized by serum antibodies produced in the absence of immunization. We analyzed a serum sample from the metastatic melanoma patient, assuming that the serum of a cancer patient can contain autoantibodies against proteins which are overexpressed or aberrantly expressed in tumor cells and had been exposed to the immune system due to tumor cell death. For the serum antibodies of the melanoma patient we identified the 500 most abundant peptides which were not shared with the list of peptides corresponding to the serum sample from a healthy donor. To identify the candidate autoantigens recognized by serum antibodies of the melanoma patients we used the same algorithm as we did for identifying the antigen used for immunization of mice. Table 3 shows the top 10 proteins ranked according to the final score calculated using the bl2seq program for the 500 most abundant peptides. The protein Mast Cell Carboxypeptidase A Precursor (CPA3), which had the highest final score, produced matches to the 52 peptides with the sum of scores equal 917.3. Of these peptides more than half (27) were aligned against a single site with the cumulative score equal 575.4 (Figure 5). This score was comparable to the scores corresponding to the major epitopes recognized by mouse sera on the PAP molecule. To test if the serum antibodies of the melanoma patient will recognize the whole CPA3 protein we analyzed the 293T cells lysate (purchased from

the origene.com) overexpressing the recombinant CPA3 protein using Western blotting. As shown in Figure 5, serum from the melanoma patient but not from the healthy donor reacted with a protein band corresponding to the size of the CPA3 protein in the 293T cells lysate overexpressing the recombinant CPA3 protein. No reactivity was detected in the control 293T cells lysate.

Discussion

We have demonstrated that the lists of peptide sequences generated by the NGS of DNA from RPPDL selected for binding to serum antibodies can be used for identifying protein antigens recognized by serum antibodies, when the humoral immune response is at least partially directed to linear epitopes. The simple BLAST-based algorithm of the SAS strategy generated a list of candidate proteins that contained multiple sequence matches to different peptides from the list. The PAP protein used for immunization of mice was one of the top fifteen candidates for the three out of four immunized mice. It is noteworthy that the proteins that were ranked higher than the PAP protein shared the same motif with the PAP, suggesting that antibodies against the PAP are likely to cross-react with the identified candidate proteins. Moreover, for the two out of three sera that showed a response to linear epitopes of the PAP protein, the score for the PAP epitope was the highest among the similar epitopes in other proteins thus making the PAP protein the top candidate for the target of immune response.

In addition to identifying the candidate proteins recognized by antibodies, the SAS also exposed quantitatively the “fine

Motif for the top proteins derived for the PAP1 list

Name	Start	p-value	Sites	Motif
XP_003119043.1	73	9.51e-08	SLEAFSSRA W T R N P S T A N T K I E	
NP_065117.1	26	3.18e-07	GLLGILVAML L S R K T S S V G A S I V	
NP_001181944.1	113	4.59e-07	QVLLSNIKV L P T S T F P O R A L I V	
NP_001090.2	222	7.35e-07	LYCESVHNFT L S V A T E D T M K I R E	
NP_653235.1	123	7.35e-07	IVKQEGMGA V T S R V Q Q V D I L A F	
NP_001098018.1	43	8.80e-07	ALYRILSLK Y S R V R T P H E D F S G	

Motif for the top proteins derived for the PAP2 list

Name	Start	p-value	Sites	Motif
NP_001090.2	223	2.04e-07	LYCESVHNFT L S V A T E D T M K I R E	
NP_000396.2	159	2.04e-07	KSEFVWPLE L S V A T E D T M K I R E	
NP_001230678.1	13	2.04e-07	TLQSETKIRV L S V A T E D T M K I R E	
NP_001092926.2	91	2.34e-07	EPAAREAPLL L S V A T E D T M K I R E	
NP_065117.1	27	2.96e-07	GLLGILVAML L S R K T S S V G A S I V	
NP_443183.1	299	7.27e-07	IKLEELKSFV L S V A T E D T M K I R E	
NP_005182.1	244	9.05e-07	TKQEHFPDNL L S V A T E D T M K I R E	
NP_114103.2	172	1.58e-06	IGAGCFLLLT L S V A T E D T M K I R E	
NP_942567.1	103	3.38e-06	PHPRGMVFER L S V A T E D T M K I R E	
NP_997394.1	107	3.38e-06	ATDISRSHEL L S V A T E D T M K I R E	
NP_001011700.2	18	4.78e-06	LSRVHFLRL L S V A T E D T M K I R E	
			ISRVHFLRL L S V A T E D T M K I R E	
			ISRVHFLRL L S V A T E D T M K I R E	
			ISRVHFLRL L S V A T E D T M K I R E	

Motif for the top proteins derived for the PAP3 list

Name	Start	p-value	Sites	Motif
NP_001165160.1	31	2.53e-09	GYQERTPRF T N E P Y L F V T M K I R D K	
NP_113633.2	132	8.31e-09	GFGERTPRF T N E P Y L F V T M K I R D K	
NP_149101.1	36	2.29e-08	GSAAFRARGR T N E P Y L F V T M K I R D K	
NP_001090.2	336	5.48e-08	FVEMYRNFT T N E P Y L F V T M K I R D K	
NP_001025046.1	234	2.86e-07	IKNQDGSLL N L E P Y L F V T M K I R D K	
NP_003921.2	7	6.47e-07	MINPSS T N E P Y L F V T M K I R D K	
NP_078870.1	179	1.43e-06	PERQKMPFN T N E P Y L F V T M K I R D K	
NP_005986.2	344	2.55e-06	APSHIGIPRT R A S P Y L F V T M K I R D K	
NP_004995.1	61	2.55e-06	AAKKINMRVE D E P Y L F V T M K I R D K	
NP_001504.2	124	4.10e-06	QVANAERFKV D E P Y L F V T M K I R D K	
			QVANAERFKV D E P Y L F V T M K I R D K	
			QVANAERFKV D E P Y L F V T M K I R D K	
			QVANAERFKV D E P Y L F V T M K I R D K	
			QVANAERFKV D E P Y L F V T M K I R D K	

Figure 3. Motifs identified by MEME software for the top candidate antigens selected for the PAP1 and PAP3 antisera. Motifs for the antigens are related to the motifs identified for the corresponding top 500 the most abundant peptides.
doi:10.1371/journal.pone.0067181.g003

structure” of humoral immune response against the antigen. SAS not only can identify the epitopes recognized by antibodies on the protein sequence but also can show the strength of the response against these epitopes, since the number of peptides containing the epitope sequence should be proportional to the titer of antibodies.

Table 2. Epitopes predicted by the SVMTrIP inside the PAP protein.

Rank	Location	Epitope	Score
1	334–353	ETQHEPYPLMLPGCSPSCPL	2.505
2	136–155	LLWQPIPVHTVPLSEDQLLY	1.152
3	360–379	VGPVIPQDWSTECMT TTNSHQ	1.004

The table shows the hypothetical epitopes predicted by the SVMTrIP software on the PAP protein. In bold are the sequences that produce matches to the peptides recognized by serum antibodies of the Pap3 serum.
doi:10.1371/journal.pone.0067181.t002

The antibody responses against the linear epitopes of the human PAP protein appear to be limited in mice only to few epitopes. Each of the PAP1, PAP2 and PAP3 antisera recognized only a single major linear epitope; for the PAP1 and PAP2 antisera, the same epitope was recognized. This result is not in contradiction with the previously observed fraction of antibodies that recognized linear epitopes. It was shown that less than 2% of the antibodies raised against the cytochrome c could also recognize the synthetic peptides. The authors concluded that linear epitopes were rare in the case of small, globular, and conformationally stable proteins. The inability of the method to identify the PSA protein which was also used for immunization can be explained by the absence of the response against the linear epitopes of the PSA protein, which is shorter than the PAP protein. We think that if the immune response in mice were to be directed against 2 or more major linear epitopes of the PAP, the PAP protein would be ranked first with a big lead over the other proteins, which were ranked higher mainly due to a shorter size than that of the PAP protein. Although the SAS has been designed for identifying proteins with multiple epitopes recognized by serum antibodies, it appears that a

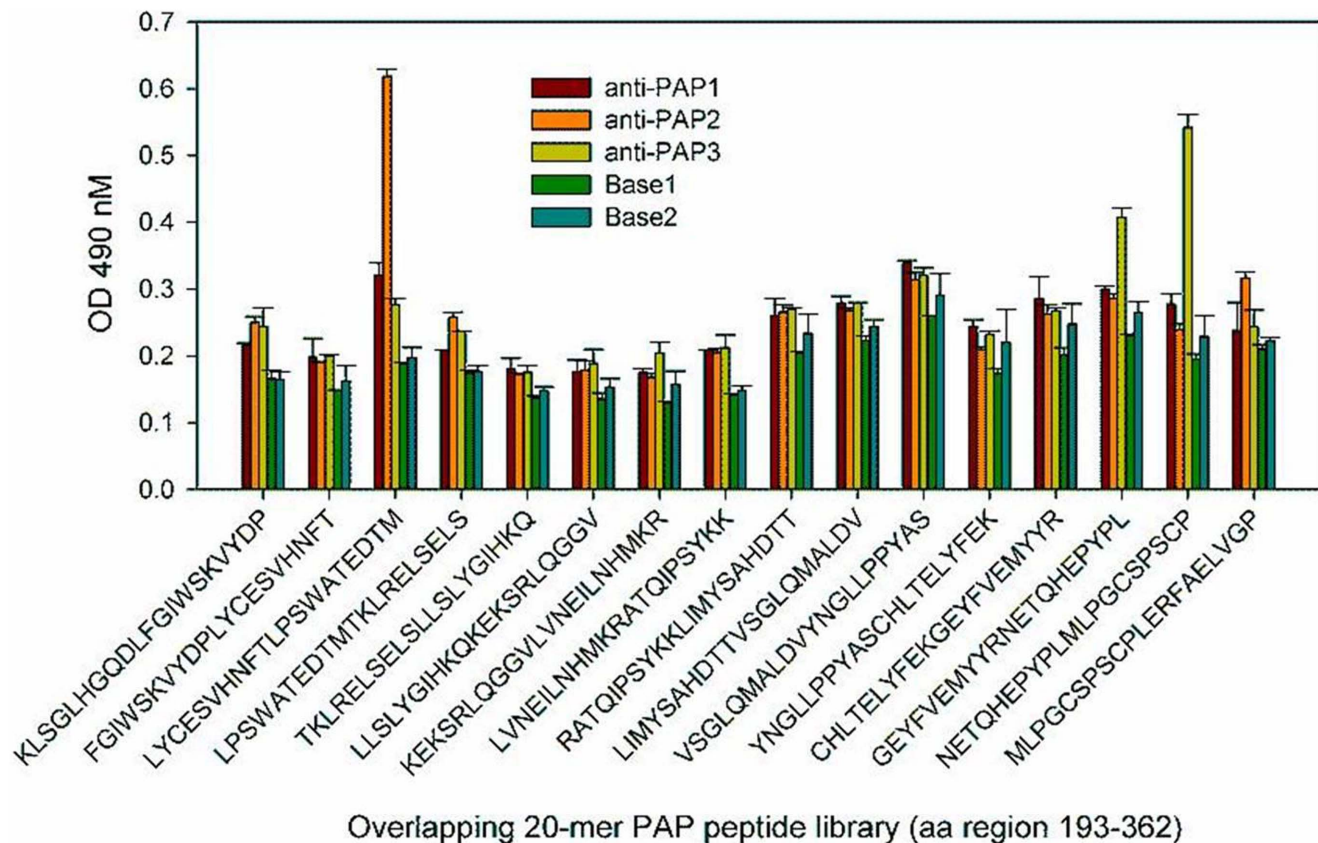


Figure 4. Epitope mapping for human PAP-specific antibodies from four immune sera samples. The reactivity of 4 mouse PAP-specific immune sera (PAP1, PAP2, PAP3 and PAP4) was tested by ELISA using a library of overlapping 20-mer peptides that span the entire amino acid sequence of human PAP. Base1 and base 2 samples derived from naïve mice served as negative controls. The results for the PAP protein region corresponding to amino acids 193–362 are shown. doi:10.1371/journal.pone.0067181.g004

large number of different peptides that have matches to a single epitope may be sufficient for identifying a protein as a candidate antigen. However, when only a single major epitope is detected on the candidate protein, the possibility that the protein is identified due to a chance can not be excluded. Based on the analysis of the PAP3 serum sample we had to choose another protein, and not the PAP protein as the first candidate for the target of the immune response (Table 1). Even when binding of a protein to serum antibodies is confirmed by Western blot, as was the case with the CPA3 protein and the melanoma patient's serum antibodies, this binding does not necessary mean that the immune response was directed against this protein, but could be attributed to the cross reactivity. It is possible that the peptide sequence recognized by serum antibodies of the melanoma patient on the CPA3 protein is the mimic of a conformational or carbohydrate epitope on the other human, viral or bacterial protein, which may represent the unknown real target of the humoral immune response. This limitation of the SAS method is shared with the other methods of serum antibody repertoire profiling which are based on using the cDNA expression libraries, such as SEREX [9].

When we analyzed the reactivity of the melanoma serum sample against the next ranked highest after the CPA3 protein, the serum antibody did not detect the corresponding band on the western blot. The absence of the specific signal for the protein on the western blot does not necessary mean the absence of autoantibodies against this protein. It can be explained by the much higher sensitivity of the SAS over the Western blot. The

detection of the immunoreactivity in the SAS is based on sequencing of the PCR amplified DNA fragment coding for the peptide that binds to antibodies. The method can detect even a single peptide that binds to a single molecule of the antibody. Such sensitivity cannot be achieved by Western blot analysis or ELISA. The extremely high sensitivity of the SAS capable of identifying low affinity or low titer autoantibodies against tumor associated antigens is of importance for development of the diagnostic assays for early detection of cancer.

The efficiency of having multiple epitopes for identifying disease-relevant targets of immune response has been demonstrated previously by searching protein databases for homology against predicted epitopes of two monoclonal antibodies recognizing the same protein [5]. In this work, we demonstrated that combining RPPDL panning with high throughput sequencing allows the generation of global profiles of serum antibody repertoires. We further demonstrated that, using the BLAST algorithm to analyze a large number of peptides against human protein databases, the targets of serum antibodies can be identified without prior knowledge about the proteins that induced immune response. We suggest that the SAS method can be used for identifying novel tumor associated antigens of cancer patients as well as the targets of immune response in autoimmune diseases. The identified by the SAS tumor-associated antigens can be used for developing personalized immunoassays for detecting recurrences and for personalized anticancer immunotherapies. The applicability of the SAS for identifying antigens beyond human proteome needs to be

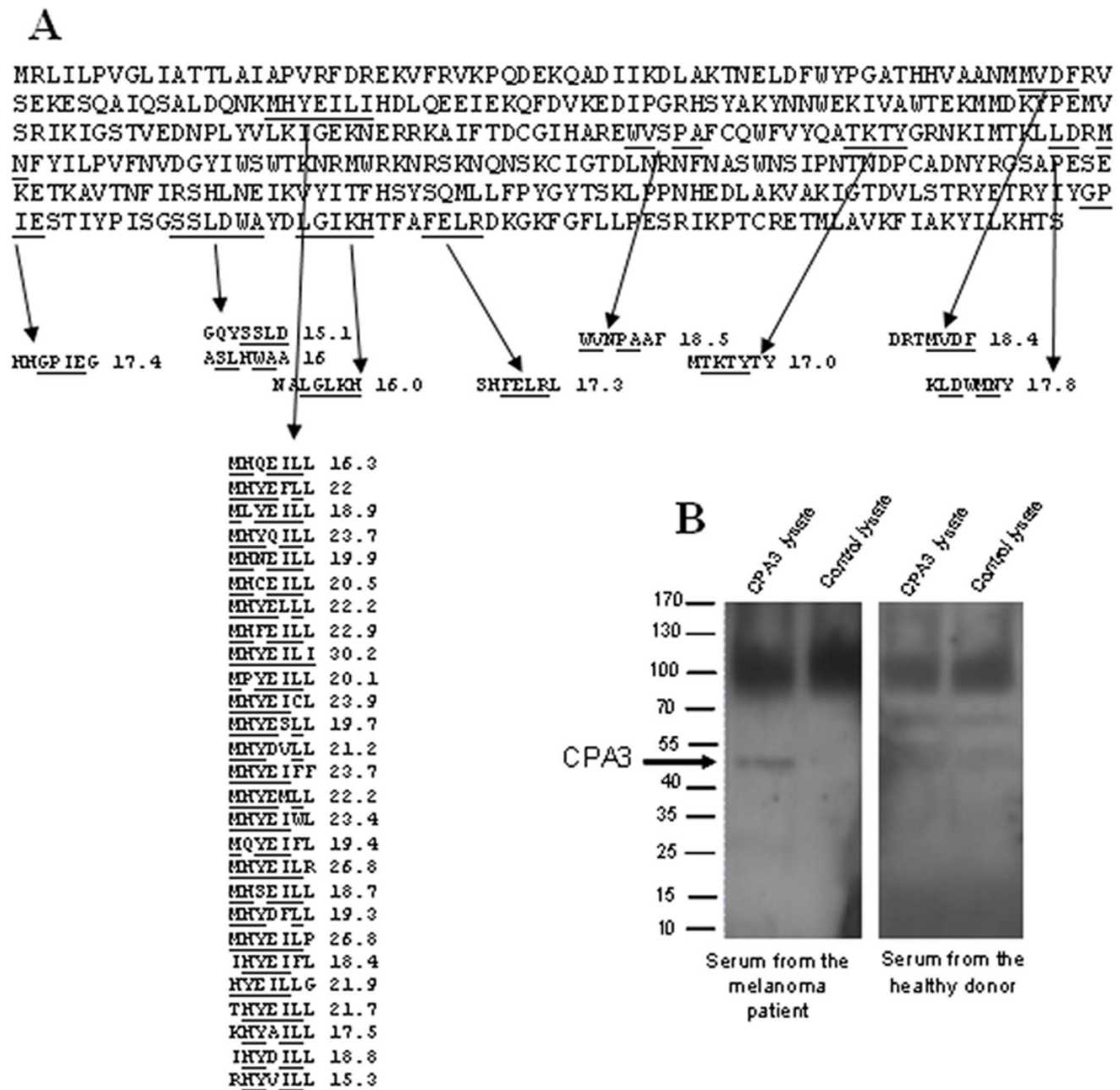


Figure 5. Epitope profile of the CPA3 protein generated by the serum from a melanoma patient. (A). The CPA3 amino acid sequence and the 36 matching peptides with the highest match scores generated by bl2seq program are shown. The sequences of the protein matching to peptides and the corresponding peptide sequences that match to protein are underlined. **(B).** Western blot analysis of the 293T cell lysates overexpressing the recombinant CPA3 protein or transfected with empty vector. The membranes were incubated with serum from either melanoma patient (left panel) or healthy donor (right panel). The reaction was developed as described in the “Materials and Methods” section. doi:10.1371/journal.pone.0067181.g005

tested. It is unlikely that the unknown pathogen can be identified by performing BLAST search for short peptides against the non-redundant protein sequences database that includes more than 30 millions of proteins from more than 20 thousands organisms. However, when the source of a pathogen is known to be associated with viruses or bacteria the BLAST search against bacterial or viral protein database is likely to be capable of identifying proteins containing linear epitopes.

Although the SAS is designed for identifying linear epitopes, a high throughput sequencing of random peptide phage libraries

selected for the large number of serum samples from infected individuals can identify peptide motifs associated with viral or bacterial infections. These peptide motifs mimicking not only linear but conformational and carbohydrate epitopes of viruses and bacteria can be used as biomarkers for diagnosis of infectious diseases. The costs of analyzing large number of serum samples using high throughput sequencing can be essentially decreased by multiplexing the samples using PCR primers containing DNA indexes tagged to serum samples.

Table 3. Top candidate antigens selected by the SAS for the melanoma patient serum.

Rank	Proteins selected for PAP1 antiserum	Protein length(aa)	Sum of overall scores	Final score	Major epitope score
1	NP_001861.2 mast cell carboxypeptidase A precursor	417	917.3	2.19976	517
2	NP_060293.2 dual specificity protein phosphatase 23	150	305.6	2.037333	252.2
3	NP_055150.1 EP300-interacting inhibitor of differentiation 1	187	318.2	1.701604	262.7
4	NP_001535.1 intercellular adhesion molecule 4 isoform 1 precursor	271	454.8	1.678229	none
5	NP_037441.2 zinc transporter 4	429	711	1.657343	280.6
6	NP_071772.1 intercellular adhesion molecule 4 isoform 2 precursor	237	385.5	1.626582	none
7	NP_001093391.1 melanoma-associated antigen B16	324	515.7	1.591667	145.4
8	NP_064576.1 28S ribosomal protein S22, mitochondrial	360	572.9	1.591389	119.7
9	NP_001034221.1 intercellular adhesion molecule 4 isoform 3 precursor	272	432.7	1.590809	none
10	NP_000307.1 parathyroid hormone-related peptide receptor precursor	593	935	1.576728	328.6

The table shows the top candidate antigens selected for the antibodies of the melanoma patient serum by sorting the data generated by processing the results of BLAST database search for the corresponding peptide sequences.

doi:10.1371/journal.pone.0067181.t003

Materials and Methods

Ethics Statement

The archived de-identified human serum samples for this study were accrued from the RPCI Data Bank and BioRepository (DDBR) with the approval of the Office of Research Subject Protection protocol #NHR010810. Written informed consent from all the donors to DDBR was obtained for use the clinical samples in research. Human serum samples were accrued with the approval of the Institutional Review Board committee of the Roswell Park Cancer Institute. The work with mice was approved by the IACUC committee of the University of Maryland School of Medicine.

Immunizations and ELISA

Mice were immunized subcutaneously with proteins in complete Freund's adjuvant as described previously [10,11]. The mouse sera at 500-fold dilution were incubated with the overlapping peptides covalently immobilized to the amine-binding, maleic-anhydride activated plates (Pierce) and the antibody binding to peptides were measured using secondary antibodies labeled with horse radish peroxidase.

Western Blot

Whole cell lysate of 293T cells over-expressing recombinant CPA3 protein or the control lysate of the 293T cells transfected with empty vector were separated on the 10% SDS-PAGE acrylamide gel electrophoresis and transferred to Immobilon PVDF membrane (Millipore). After blocking the membrane with 5% milk in phosphate buffered saline buffer (PBS) the membrane was incubated overnight at 4°C with serum from either melanoma patient or healthy donor. Serum was diluted 100-fold in PBS containing 0.05% Tween 20 and 1% bovine serum albumin. A horseradish peroxidase-conjugated goat anti-human antibody was then added, and secondary antibodies were detected through autoradiography using enhanced chemiluminescence (ECL Plus, General Electric Healthcare, Milwaukee, WI).

Generating Serum Antibody Repertoire Profiles

Twenty µl of mouse or human serum and 10 µl of the Ph.D.7 random peptide library (NEB) were diluted in 200 µl of the Tris Buffered Saline (TBS) buffer containing 0.1% Tween 20 and 1%

BSA and incubated overnight at room temperature. The phages bound to antibodies were isolated as follows. 20 µl of protein G agarose beads (Santa Cruz) were added to the phage-antibody mixture and incubated for 1 hr. To eliminate the unbound phage the mixture with beads was transferred to the well of 96-well MultiScreen-Mesh Filter plate (Millipore) containing 20 µm pore size nylon mesh at the bottom. The unbound phage was removed by applying vacuum to the outside of the nylon mesh using micropipette tip. The beads were washed 4 times by adding to the well 100 µl of TBST buffer and removing the liquid by applying vacuum to the outside of the nylon mesh using micropipette tip. The phage bound to the antibodies was eluted by adding to the beads of 100 µl of 100 mM Tris-glycine buffer pH 2.2 followed by neutralization using 20 µl 1 M Tris buffer pH 9.1. The eluted phages were used for the amplification in bacteria. The amplified phages were incubated with 20 µl of serum overnight at room temperature followed by isolation of the antibodies-bound phages using 20 µl of protein G agarose beads. The phages were then eluted from antibodies/protein G complexes using low pH buffer as described above, and the DNA was isolated using phenol-chloroform extraction and ethanol precipitation. The 21 nt long DNA fragments coding for random peptides were PCR-amplified using primers containing a sequence for annealing to the Illumina flow cell and the sequence complementary to the Illumina sequencing primer. The PCR-amplified DNA library was purified on agarose gel and DNA from all samples were multiplexed by adding 4-base bar code at the beginning of each DNA fragment.

Nextgen Data Processing

The high-throughput sequencing was performed using Illumina Hiseq2000. A total of about 313 million raw reads were generated. The sequences were de-multiplexed to determine its source sample. The 21- base nucleotides representing the 7-amino-acid peptide were extracted between base position 29 and 49. All identical 21-mer DNA sequences were collapsed into single sequence with its coverage (frequency) recorded in the result multi-FASTA file. These DNA sequences were translated to peptide sequences using the first frame. All barcode splitting, read trimming, and sequence collapsing were done using FASTX-Toolkit. Peptide translation, selection of subsets of sequences of shared DNA, abundant DNA passing minimum coverage threshold, were done through customized Perl script. BL2seq was

also automated to handle batch processing of thousands of sequences in a very short time frame. The next generation sequencing data are deposited to the NIH Short Read Archive. The accession number for the sequences in the SRA database is SRP021104.

Supporting Information

Table S1 The table shows 500 the most abundant peptides with corresponding copy numbers selected for the antibodies from each of the four anti-PAP sera. The selected peptides are not shared by the sera from unimmunized mice as well as from mice immunized with the PSA antigen.
(DOC)

Tables S2 S2A, S2B and S2C. The table shows the list of proteins selected by doing protein BLAST of peptide sequences against refseq_protein database for the Homo Sapiens (taxid:9606) with the maximal score ≥ 18.5 threshold parameter. In the column A, the proteins are sorted by the increase of the protein accession number. Highlighted in yellow are the proteins which have been retrieved multiple number of times by BLAST search against different peptides. The column B shows the lengths of proteins in

the number of amino acids. The column C shows the number of the matches to peptides that retrieved proteins at the selected E-value threshold. The column D shows the initial scores calculated as the number of matches in column C divided by protein length in column B. Sorting the data by the descending numbers in column D allowed to select the top candidates for the second step of analysis. The column E shows the sums of scores for all the peptides that produced matches to the protein in Blast2seq analysis of the protein against all the most abundant 500 peptides. The column F shows the final scores calculated as the sums overall scores in column E divided by protein length in column B. Sorting the data by the descending numbers in column E allows to select the candidate antigens containing linear epitopes recognized by serum antibodies. The column G shows the sum of the scores for peptides that match to the single major site on the protein.
(XLS)

Author Contributions

Conceived and designed the experiments: ENK MN YI. Performed the experiments: XL QH SL LJLT LS CAS. Analyzed the data: ENK MN YI. Contributed reagents/materials/analysis tools: XL ENK MN. Wrote the paper: ENK MN YI. Designed the software used in analysis: XL.

References

1. Mintz PJ, Kim J, Do KA, Wang X, Zinner RG, et al. (2003) Fingerprinting the circulating repertoire of antibodies from cancer patients. *Nat Biotechnol* 21: 57–63.
2. Scott JK, Smith GP (1990) Searching for peptide ligands with an epitope library. *Science* 249: 386–390.
3. Folgori A, Tafi R, Meola A, Felici F, Galfre G, et al. (1994) A general strategy to identify mimotopes of pathological antigens using only random peptide libraries and human sera. *EMBO J* 13: 2236–2243.
4. Mennuni C, Santini C, Lazzaro D, Dotta F, Farilla L, et al. (1997) Identification of a novel type 1 diabetes-specific epitope by screening phage libraries with sera from pre-diabetic patients. *J Mol Biol* 268: 599–606.
5. Bastas G, Sompuram SR, Pierce B, Vani K, Bogen SA (2008) Bioinformatic requirements for protein database searching using predicted epitopes from disease-associated antibodies. *Mol Cell Proteomics* 7: 247–256.
6. Ionov Y (2010) A high throughput method for identifying personalized tumor-associated antigens. *Oncotarget* 1: 148–155.
7. Dias-Neto E, Nunes DN, Giordano RJ, Sun J, Botz GH, et al. (2009) Next-generation phage display: integrating and comparing available molecular tools to enable cost-effective high-throughput analysis. *PLoS One* 4: e8338.
8. Ryvkin A, Ashkenazy H, Smelyanski L, Kaplan G, Penn O, et al. (2012) Deep Panning: steps towards probing the IgOme. *PLoS One* 7: e41469.
9. Sahin U, Tureci O, Schmitt H, Cochlovius B, Johannes T, et al. (1995) Human neoplasms elicit multiple specific immune responses in the autologous host. *Proc Natl Acad Sci U S A* 92: 11810–11813.
10. Klyushnenkova EN, Kouiavskaia DV, Kodak JA, Vandenbark AA, Alexander RB (2007) Identification of HLA-DRB1*1501-restricted T-cell epitopes from human prostatic acid phosphatase. *Prostate* 67: 1019–1028.
11. Klyushnenkova EN, Link J, Oberle WT, Kodak J, Rich C, et al. (2005) Identification of HLA-DRB1*1501-restricted T-cell epitopes from prostate-specific antigen. *Clin Cancer Res* 11: 2853–2861.

Supplementary Table S1.

Antiserum	PAP1	PAP2	PAP3	PAP4
Peptide sequences with corresponding copy numbers	>213-8341_1 IVLPTWL >236-7646_1 YSTRHTD >280-6564_1 ADRSPLW >319-5936_1 YSSVHID >355-5487_1 YTSQTPR >406-4879_1 QTSITMY >426-4754_1 HRVLWHQ >436-4720_1 SFTYPSW >439-4684_1 ELCFKCS >501-4296_1 SDRSALW >502-4286_1 VSPGPTG >508-4251_1 SEWLWDH >515-4208_1 QCTLCPR >520-4188_1 YATPWQM >545-3992_1 TLIDPSR >585-3803_1 VLSDPPP >595-3753_1 RIPMQDL >611-3641_1 YVFPTWS >619-3619_1 ERMNHNP >643-3466_1 FYLPPWS >646-3434_1 ASQDITP >649-3422_1 IHTHPWY >679-3345_1 SHGSVHA >705-3237_1 GLGSVKP >714-3209_1 YIWPTWA >716-3205_1 VGFRGL >731-3137_1 STTIALF >745-3081_1 SSPPKRL >746-3080_1	>133-13001_1 AWFYPPW >156-11910_1 YVHPPWW >157-11908_1 LGYPWMQ >166-11423_1 NFRLPWW >178-10841_1 NPQWQTS >217-9446_1 YELPPWM >225-9164_1 FVRPSWI >229-8978_1 AFFGPSW >255-8115_1 WTRPPWE >263-8001_1 TFLPSWA >264-7974_1 YVFPTWS >267-7948_1 APPLFTP >302-7088_1 WTLPSWE >303-7032_1 NFQMPSW >318-6902_1 QWWNHEI >323-6818_1 GCCMHKN >355-6308_1 YINPSWW >363-6226_1 SLVDIWW >364-6210_1 FSLPPWA >371-6099_1 YVWPSWM >376-6041_1 GSLHFLV >380-6015_1 FTSPSWQ >389-5925_1 TYWPFPR >393-5893_1 SPLNIAR >397-5841_1 SYPWLIN >419-5674_1 KDLRDWF >437-5566_1 YVNPWEP >454-5382_1 FWEPWGW >455-5372_1	>166-4304_1 RVIAPSD >204-3575_1 WPVPYPL >259-2906_1 QWDDHW* >326-2221_1 QSPYPLN >366-2046_1 WPAREYF >385-1959_1 TLSGSRF >394-1905_1 TPQPYPL >436-1746_1 YETWEHR >442-1729_1 THSIFVY >471-1651_1 DWRNHWL >481-1612_1 AVSIVKR >504-1553_1 QWDDHWY >510-1534_1 RHEWLLH >543-1431_1 WLTPYPL >581-1355_1 AIPRLEE >608-1296_1 YHDAYPM >634-1249_1 YEGSIPH >646-1236_1 SPSPYPF >679-1179_1 VPSAYPL >689-1168_1 QANNTM >698-1143_1 FWNTEML >702-1138_1 YNPWTHW >708-1131_1 NFRILSI >738-1088_1 THYRVNS >754-1059_1 KQPLHWY >759-1056_1 DLTPYPL >778-1031_1 DQPMHWF >819-979_1 FSWRMMT >823-976_1	>101-10104_1 VFIPYGH >181-5658_1 KVMQLHI >397-2768_1 QMQWAVT >431-2573_1 YRGISVN >461-2449_1 AYLFNPM >502-2225_1 HAPIGNN >563-2049_1 RDLATLQ >574-1993_1 TVNNRMY >585-1953_1 GVWMLPK >643-1824_1 LVSSYNF >645-1813_1 TSIMFRY >672-1742_1 SVSTNSM >681-1724_1 ETRT*TL >683-1714_1 QPGRIVQ >712-1650_1 GIGRVAL >713-1649_1 LVALRVG >753-1558_1 ASGSLCC >758-1552_1 VVPCAGL >763-1539_1 TYTNPGF >767-1533_1 ETGSNTI >776-1520_1 VMDLRIL >791-1500_1 ISLRIEM >826-1450_1 MPFWLAR >840-1424_1 MPYDRSR >875-1363_1 RIHVDLF >878-1361_1 TSIGVFS >915-1301_1 WNSVSRL >933-1283_1 FFWMTRI >975-1234_1

YAPPYQR	GDMHRVM	DHVSWHY	VSPVRQL
>749-3060_1	>456-5363_1	>836-963_1	>987-1223_1
STSELYG	TQVRPPW	QWEDHWQ	MYQINVT
>772-2990_1	>465-5268_1	>885-924_1	>1016-1174_1
YDMSLIG	SFLPPWM	QDPYPMN	SGEQYRI
>818-2853_1	>473-5176_1	>892-921_1	>1021-1170_1
QYMNQHA	WNWPTWE	DHTSWHP	VPMQTRL
>835-2808_1	>488-5067_1	>920-898_1	>1030-1153_1
YDRSVLW	DLYYGRV	SGRIIGS	SHTNPKS
>839-2802_1	>507-4925_1	>931-888_1	>1052-1127_1
HFVWTRS	GLWADNE	SHVSWHT	NILANDR
>847-2777_1	>509-4911_1	>940-883_1	>1067-1113_1
IELRGIY	FNSPSWS	MHRMELR	ISTHYYL
>849-2777_1	>516-4858_1	>964-861_1	>1081-1097_1
SDTTGLP	GFMAPSW	GHEPYPL	WTVVQTL
>853-2770_1	>518-4846_1	>969-858_1	>1096-1083_1
EFTYPSW	LGSAWDV	DVPNWRT	H*EISVL
>875-2721_1	>525-4793_1	>977-845_1	>1181-1021_1
GWMMRFS	IIMPSWE	MLTPYPL	YSYVHTV
>892-2684_1	>526-4787_1	>982-844_1	>1190-1015_1
QDRSILW	GLNAALF	YPNTDMT	HYSATAL
>894-2683_1	>535-4734_1	>984-842_1	>1191-1014_1
ERVMSMM	APQMHLF	AWPRHAA	QPSYWSR
>920-2624_1	>536-4722_1	>986-842_1	>1201-1005_1
LAAKVHR	TDGKSHL	VTITVQR	YVQGSRL
>936-2593_1	>551-4575_1	>993-836_1	>1205-1001_1
MLDPRWE	YVQPSWL	CLAGGRP	WPSFVPK
>944-2574_1	>557-4543_1	>994-836_1	>1210-997_1
NFMYPSW	WEKINHT	NPSAYPM	TFAGYRV
>966-2524_1	>561-4519_1	>998-835_1	>1211-996_1
TPMKISH	AP*MHLF	THVSWHT	SISHRMY
>969-2519_1	>569-4470_1	>1006-830_1	>1223-986_1
AQSNSMH	YLLPPWE	DMPSQWG	HVRANVP
>971-2514_1	>572-4457_1	>1028-809_1	>1227-983_1
ILPVHVS	HYLPSWH	LRQPATG	WNLGGGS
>983-2495_1	>584-4396_1	>1036-805_1	>1231-980_1
VPPGTII	FSFPPWS	DMPLWWT	ASKLSTG
>986-2487_1	>586-4391_1	>1073-780_1	>1238-976_1
QGLGTRG	LVLPSWL	EHPSWYP	YQPVTYR
>991-2474_1	>611-4236_1	>1075-777_1	>1261-960_1
EPKTKHI	EQFYFAM	GGLIIGM	QPDSNVA
>1016-2403_1	>625-4170_1	>1077-774_1	>1272-955_1
HLSYGLN	ILWLLLW	ASNANLR	VIHGNPP
>1017-2399_1	>626-4159_1	>1083-770_1	>1276-954_1
YIFPTWA	NYWPFPPW	WMPMRDR	SF*TDHL
>1047-2361_1	>627-4151_1	>1092-766_1	>1279-952_1
YVFPTWM	VVWDRNP	YPPFTDP	QQHHVHL
>1069-2333_1	>629-4140_1	>1107-755_1	>1300-939_1
QFMSQVY	GYPWISN	GKPCEHC	VSPRAGM
>1074-2324_1	>640-4083_1	>1113-751_1	>1322-928_1
SSGQHAF	ESLNSFA	EFRLPQL	DRNPMSF
>1084-2307_1	>645-4062_1	>1116-749_1	>1334-920_1
FIWPTWS	TMWPFPS	TPQPYPL	DPAT*TV
>1120-2238_1	>649-4044_1	>1117-746_1	>1348-914_1
TGSPQAR	FVRPTWH	FNVGVKP	GSPQVKK
>1124-2235_1	>650-4042_1	>1136-731_1	>1403-878_1
YGVMAFH	HPLQVFF	QCDDHW*	RVQDLPH
>1125-2234_1	>653-4026_1	>1140-728_1	>1413-874_1
RNDLPYT	YFLPPWI	NHTPYPL	SFRLGSH
>1127-2231_1	>656-4007_1	>1146-726_1	>1415-873_1
GQWSRIN	NVLPPWL	LSKFPTS	TVHAAVM

>1155-2183_1	>662-3993_1	>1156-724_1	>1430-865_1
GNNLVRT	FELPSWM	TVAQATS	WGQLGIR
>1156-2182_1	>664-3987_1	>1159-722_1	>1466-848_1
SSAQSSQ	FVWPYWQ	YRPSMNS	HISKGMT
>1168-2165_1	>675-3935_1	>1245-684_1	>1477-843_1
FKFTFAL	IVLPSWW	SWAIRIY	AVQFL*T
>1178-2152_1	>676-3935_1	>1252-678_1	>1493-838_1
SRVVDGY	APPMHII	NWLKVG	GVNIRGI
>1187-2142_1	>683-3923_1	>1256-676_1	>1505-830_1
VPSRESW	EFFFPSF	TAWSVMK	GQRVVIS
>1211-2092_1	>686-3907_1	>1292-660_1	>1515-826_1
QLAMTYV	VYAHTWL	TVEYITN	MVHDMLG
>1231-2061_1	>687-3906_1	>1301-656_1	>1536-817_1
MMNPIRN	FVSPSWK	TFVSMPP	GNGPLHW
>1248-2041_1	>694-3871_1	>1315-648_1	>1537-816_1
GELPGTA	FSMPSWQ	RLPMVEP	VGKMHAN
>1249-2037_1	>696-3868_1	>1330-642_1	>1550-807_1
WIGIDTV	ITLPTWL	NALTGSA	LAAKSGT
>1253-2030_1	>699-3863_1	>1334-640_1	>1552-806_1
LSINVVF	GFKEPSW	HYKPVYP	GTSPDVW
>1265-2018_1	>700-3861_1	>1344-636_1	>1555-802_1
EEWNWLR	TFRAPSW	IPHWVQT	HTLV TAR
>1283-1998_1	>711-3797_1	>1367-626_1	>1563-799_1
LSWTGKS	KAPHLES	RPEPYPL	VSWSSSS
>1317-1959_1	>722-3764_1	>1395-610_1	>1568-798_1
GLFREGW	HFAFPSW	QWEDHC*	ITWSEHP
>1335-1930_1	>730-3721_1	>1397-610_1	>1586-789_1
MAQSSAA	FHRPSWD	GQRVVIS	AYFNLLR
>1348-1918_1	>732-3710_1	>1399-608_1	>1592-786_1
QYMYQLP	FTHPWS	TPQAYPL	ASHAVGV
>1350-1916_1	>742-3676_1	>1401-606_1	>1600-780_1
DFWHHLR	VFVEPSW	THSPYPM	TEWLRFD
>1357-1900_1	>743-3661_1	>1427-593_1	>1615-770_1
ADRSALF	GGKLPLN	AHLSIIF	SLKHMY
>1369-1885_1	>744-3657_1	>1441-587_1	>1625-763_1
VPSFVPL	SQLMPSW	QSSKFLH	MSWKLLQ
>1415-1820_1	>749-3641_1	>1458-579_1	>1627-762_1
QCCLTCA	IWLPPWF	MSMSQWR	LNKLT
>1424-1813_1	>752-3628_1	>1474-574_1	>1654-747_1
DSPLHAT	FLLPPWN	QAALNRM	QTGLSSS
>1439-1803_1	>766-3561_1	>1477-573_1	>1675-737_1
QTYRPSY	GWFLPSW	FRPPMMD	AFHVTRM
>1444-1800_1	>768-3551_1	>1493-567_1	>1677-736_1
ANFWGES	WFINTPW	IGVHHWN	AYPKADS
>1457-1784_1	>771-3539_1	>1495-565_1	>1680-735_1
NFYLPTW	AWWAPSW	WGLHTFS	VGTTVVK
>1496-1734_1	>779-3509_1	>1500-564_1	>1694-730_1
WPISFLI	SFVKPTW	LPGFSVS	TGSLFLT
>1497-1734_1	>784-3499_1	>1503-563_1	>1715-722_1
YVFPTWM	KPVDLTA	KCCATHL	IARFSAN
>1500-1731_1	>785-3490_1	>1505-563_1	>1730-716_1
YSTNHID	FRQPSWY	DQPMHWI	HTPVLSN
>1504-1724_1	>787-3481_1	>1515-559_1	>1737-713_1
WKSPVLM	AQWYPPA	MSHMDSP	VSYGGR
>1513-1716_1	>804-3422_1	>1538-552_1	>1739-711_1
VFGVHAH	WVLPSWM	YILRAGD	SIYSWAV
>1520-1711_1	>807-3405_1	>1540-551_1	>1743-708_1
FNQPTWL	FFDPSWN	CMQAHAM	FRFSCRW
>1522-1710_1	>809-3399_1	>1548-547_1	>1750-703_1
HHSPMIG	GFTLPSW	DHLMHQR	NGLVYRE
>1527-1706_1	>816-3384_1	>1553-544_1	>1752-702_1

SLGKTVA	FDFPPWA	QWEDHWE	RYLDVWA
>1537-1691_1	>818-3373_1	>1573-539_1	>1756-701_1
AWPRHAA	FNEPPWH	MQTRLMA	EMYFTKQ
>1540-1686_1	>831-3340_1	>1607-527_1	>1763-698_1
GIFKYFS	VWMPSWA	HTTKPSL	ISLTEHP
>1542-1684_1	>835-3329_1	>1611-527_1	>1773-694_1
NFVPPSW	HEPTVMS	TVTPYPM	AQMPVNV
>1555-1663_1	>842-3285_1	>1612-527_1	>1778-691_1
MGDIKHV	RMVLPSW	ACLMCLT	YHPHPRWT
>1561-1658_1	>858-3222_1	>1613-526_1	>1790-689_1
TGTDATF	LTDHQRY	HPFNQLQ	FSSLGSK
>1562-1655_1	>863-3215_1	>1620-523_1	>1791-689_1
CYMCQSL	NFILPTW	WTIVGPI	WSTLTTP
>1577-1641_1	>865-3208_1	>1622-522_1	>1795-688_1
QYMFQTS	ANLSQVL	KMEPYPL	SAPDDSF
>1585-1636_1	>872-3191_1	>1630-521_1	>1796-688_1
TQGYATS	GFFLPTW	MHSPYPL	TWIRSAL
>1606-1622_1	>882-3172_1	>1638-518_1	>1797-688_1
NSGLHMF	SDHRHWI	QLDPYPL	MLQTNID
>1619-1603_1	>884-3165_1	>1642-517_1	>1810-682_1
YPHRQVP	RFMVPSW	HLTAVLT	DRHFNYL
>1635-1589_1	>899-3114_1	>1654-513_1	>1816-681_1
YSRLHLD	VQMPASK	MDILAFH	QLSNCCQ
>1637-1586_1	>910-3080_1	>1658-512_1	>1824-678_1
LPVERLN	HSLERLA	KPNIAIL	GSTSPIV
>1648-1578_1	>922-3036_1	>1663-509_1	>1843-671_1
LLAVDAL	FALPSWH	QSYLYGW	LERMPNP
>1649-1578_1	>925-3029_1	>1679-506_1	>1847-667_1
HSPQSIA	GDYIPVK	QLEPYPL	MWRIDFR
>1652-1576_1	>927-3028_1	>1706-498_1	>1849-666_1
SGMTVLI	EDTKTSH	AYPLYPT	SQSTVLQ
>1657-1573_1	>939-2996_1	>1709-497_1	>1857-664_1
WDPHWPK	FVFPSFH	EQPVHWS	HNPNGTHK
>1665-1565_1	>940-2995_1	>1711-497_1	>1871-660_1
VVTIPYP	WTQQWYY	GVQKASN	SGLDLWT
>1667-1563_1	>941-2994_1	>1719-494_1	>1893-654_1
LATRSSH	LVYGDER	MTFTIAA	QTSGSKL
>1675-1556_1	>949-2972_1	>1721-494_1	>1906-649_1
GPHFELR	VPSGKLH	DSFRVWP	QSQIPNS
>1683-1551_1	>955-2945_1	>1738-487_1	>1921-645_1
RVLPSE	VVWDRHL	HLLNGKP	SLPLVYS
>1689-1544_1	>959-2935_1	>1743-486_1	>1934-641_1
ETSSSIA	ILLPSWE	AHDPYPL	HSWIMHT
>1691-1544_1	>963-2920_1	>1747-484_1	>1938-640_1
HLQSSYQ	TFSYPTW	DQPSHWM	WCRGLVW
>1701-1528_1	>966-2909_1	>1748-484_1	>1947-637_1
SIRLHLG	VHNTPYQ	NAHTYGI	TFTSYLT
>1702-1526_1	>967-2902_1	>1750-483_1	>1949-637_1
NGYINAM	QWWNLEI	HGLHQFN	AIMHMMV
>1720-1506_1	>976-2877_1	>1755-480_1	>1957-633_1
FYLPTWY	WHDPSPV	VMRPVTA	ASLDRYN
>1735-1494_1	>981-2861_1	>1756-480_1	>1966-631_1
EFHLGES	DFIIPSW	YPEKSPP	MIKYRVA
>1739-1492_1	>983-2853_1	>1758-479_1	>1967-631_1
MSNRDSE	YPWINQP	LTHATMR	SMLIPHP
>1740-1491_1	>1006-2796_1	>1761-477_1	>1970-630_1
FTLPSWL	RFMDPPW	APHRLTQ	SLYP*IG
>1745-1489_1	>1023-2760_1	>1762-477_1	>1981-625_1
KVTTMKP	FNPPSWG	NMSPYPM	VMQGVLF
>1746-1488_1	>1026-2758_1	>1764-477_1	>1998-619_1
APPLSNL	FINPTWA	GYSLHHR	NHVS*VS

>1762-1481_1 GSAQESH	>1027-2754_1 TWFAPTW	>1772-476_1 ELKIIQP	>2000-619_1 TPHWYVK
>1763-1480_1 HHGAKHF	>1030-2748_1 VILPSWY	>1798-466_1 FRLPMTD	>2006-616_1 TEQYTLN
>1800-1452_1 TEFDVDH	>1034-2744_1 VPGHHLM	>1810-463_1 MFTRGQE	>2010-615_1 QMRPQHT
>1824-1436_1 ITSPWPG	>1041-2737_1 SVPSSSEV	>1816-461_1 RPPMNDL	>2012-614_1 KLYQREP
>1852-1413_1 VAKTYTS	>1043-2732_1 FVMPYWA	>1817-460_1 *IDPYPL	>2017-614_1 GSNSARS
>1864-1407_1 WGQNQSS	>1048-2728_1 FKWPPWN	>1823-458_1 ITSSHVV	>2035-606_1 NATMSAS
>1866-1406_1 LLGNIPP	>1052-2722_1 LIYPSWA	>1825-458_1 LRLPMHD	>2043-604_1 IYPKVGH
>1869-1404_1 TSTTSRG	>1058-2713_1 KSADHYY	>1826-458_1 HFGPRQY	>2052-602_1 STYKSML
>1885-1391_1 YPFGVWT	>1061-2705_1 FSLQSFA	>1840-454_1 GPLMPPY	>2069-597_1 MSRGAFI
>1889-1389_1 NHHAYPT	>1070-2682_1 SFMFPPW	>1847-451_1 VGTIPPL	>2075-596_1 IWFTFQD
>1899-1380_1 ELALWST	>1071-2679_1 SFYEPSW	>1852-450_1 THVSWHI	>2085-593_1 APGGVVR
>1900-1380_1 SWLSIFG	>1072-2671_1 TFTLPTW	>1856-449_1 TSSVSLM	>2093-591_1 ERVLPVR
>1914-1369_1 TDNSARP	>1073-2669_1 HSLEMLW	>1858-449_1 DRSAFSS	>2100-589_1 VPLWYVK
>1924-1363_1 HTAHVVT	>1080-2660_1 SFQSPSW	>1865-448_1 GYCAEDT	>2115-585_1 ATRPLNH
>1927-1361_1 TVCCQTA	>1081-2660_1 SYLRDLW	>1866-448_1 VLGPYPL	>2117-584_1 EQQTIRL
>1946-1351_1 FWFPWE	>1084-2652_1 HTHASGG	>1867-448_1 FRLPMND	>2119-583_1 TWS*AQT
>1948-1350_1 KFIIKAQ	>1085-2650_1 AFRMPSW	>1872-446_1 TCSAKWC	>2132-580_1 QLWYKIN
>1952-1347_1 SVELQFG	>1092-2643_1 LVWDPWP	>1875-445_1 GPTRYPL	>2137-578_1 FPGVQTQ
>1981-1335_1 FYWPTWV	>1101-2625_1 FTQPSWQ	>1876-445_1 HLGPEQP	>2139-578_1 VLFPDIT
>1989-1330_1 MTQDITI	>1119-2575_1 RWVSPPW	>1885-442_1 ERDMRGI	>2145-577_1 TNKTFQL
>1990-1330_1 SPLPWWQ	>1125-2561_1 VLPSWAQ	>1905-438_1 KDWPGR	>2146-577_1 ALFSNGR
>2024-1305_1 WGGTLAR	>1132-2551_1 LIPVALT	>1924-433_1 QVTLSGH	>2153-575_1 EASTKLT
>2025-1304_1 FSTVSQR	>1140-2544_1 YWLPPWS	>1929-432_1 THWTRAD	>2157-574_1 AIRAPWS
>2030-1301_1 RLVAQWP	>1147-2538_1 EQFYPNW	>1937-431_1 NKVIYYP	>2164-573_1 IGLSEHF
>2034-1298_1 WTPMHDP	>1150-2532_1 HVKAVET	>1959-427_1 LKTMPMI	>2179-568_1 SVPYKLL
>2041-1296_1 GVQKASN	>1160-2512_1 YEMPIPL	>1960-427_1 HLQWST*	>2198-563_1 AAPHPAT
>2042-1295_1 DRLHMEP	>1170-2497_1 GASRAPA	>1972-425_1 YGGFPMT	>2226-555_1 YRITIPM
>2048-1290_1 QFFWPAG	>1176-2480_1 NFTRPSW	>1981-424_1 VRSPFPM	>2238-552_1 YVELIFV
>2057-1285_1 WSSGVIF	>1181-2472_1 VWDRSP	>1985-423_1 RPLPYPL	>2261-546_1 TPHWYVR
>2058-1284_1 SLALMMR	>1187-2461_1 MVVSSAN	>1986-423_1 HSLPHER	>2268-545_1 NTSTYLM
>2063-1280_1	>1190-2458_1	>1987-423_1	>2270-544_1

ETHHAHR	FTQPPWR	YLSPYPL	TVLSAAL
>2069-1277_1	>1191-2458_1	>1989-423_1	>2282-542_1
LSHFSIV	FLVPSWH	KATYIVL	TWPEFTK
>2078-1274_1	>1193-2455_1	>1996-421_1	>2286-541_1
FIPRLSY	WFRMPSW	DRNMHIP	YGFTSNR
>2098-1261_1	>1195-2451_1	>1997-421_1	>2299-539_1
SVTQSVI	QMEALMP	ITKPEPN	ELQIVPS
>2116-1250_1	>1197-2450_1	>2000-420_1	>2308-537_1
LMTPIMK	SPPLQMF	GYGKDWK	LT*KAQY
>2129-1242_1	>1200-2439_1	>2008-418_1	>2313-535_1
KFVVKN	EFRSPSW	MSYQIKR	AKVDPED
>2156-1226_1	>1203-2434_1	>2009-418_1	>2314-534_1
DHDRSWF	IFSFPSW	YTLSLSL	YPHPRWN
>2161-1224_1	>1206-2426_1	>2015-417_1	>2316-534_1
TQHSYFL	EVFWARL	NLEPYPL	NFDT*AL
>2164-1222_1	>1207-2423_1	>2017-416_1	>2332-531_1
GNTYGFH	RPVDTPQ	GMQPKPV	VSLTYHI
>2168-1217_1	>1214-2408_1	>2022-414_1	>2340-529_1
IVMPPGP	HSLVLPN	WPGPYPL	GVAHMTT
>2170-1216_1	>1220-2402_1	>2024-414_1	>2347-526_1
VSQSHLQ	QWWAGES	EHPSWAL	NAPWWGS
>2180-1210_1	>1241-2378_1	>2060-406_1	>2352-525_1
FRIPMID	KFQLPSW	ALTPYPL	SLTEKNW
>2183-1208_1	>1243-2376_1	>2070-403_1	>2355-524_1
IFARTTM	TFNPPPW	AHVPFHH	VANTHHG
>2184-1208_1	>1247-2373_1	>2073-402_1	>2373-520_1
QHRALPT	YPWMQFS	DQTGGTN	ECRPCAD
>2190-1206_1	>1252-2362_1	>2079-401_1	>2381-518_1
AVNTFML	FLWPSWT	YARMLLS	NIVSNSS
>2191-1206_1	>1254-2357_1	>2085-401_1	>2387-516_1
KGAYYQS	QFRLPPW	KMTPYPL	DPLLNGV
>2205-1197_1	>1267-2331_1	>2089-399_1	>2390-516_1
VGLSIHI	FNPMWQW	TDWLWWE	YMTGSKP
>2229-1185_1	>1274-2323_1	>2091-399_1	>2393-515_1
DKLYVTL	FHYPWR	EAVPLRS	*TWRGAV
>2231-1183_1	>1275-2322_1	>2093-399_1	>2397-514_1
TVRTGST	FTEPTWA	TGSPSMA	VLYSEHP
>2253-1169_1	>1279-2317_1	>2101-398_1	>2400-513_1
GQDITAQ	VPRWFPP	KQMPYH	KIPLTSK
>2256-1169_1	>1281-2316_1	>2105-398_1	>2401-513_1
ADRSIMW	YYFPSWM	SCPMCYW	SVFSFRL
>2273-1155_1	>1282-2315_1	>2117-396_1	>2423-509_1
THHHPDE	YIVLPPW	QEPYPLN	SSMRQMH
>2276-1151_1	>1283-2314_1	>2119-396_1	>2432-508_1
QQFKSMS	FFSPSWL	KTSADMM	ALVFESY
>2278-1148_1	>1286-2311_1	>2126-394_1	>2441-507_1
LIPTNTV	RADYHAY	GWSEHWS	WPTFIPR
>2281-1147_1	>1292-2289_1	>2128-394_1	>2442-507_1
LGTHLHQ	YLGNYMI	TDSRHVL	LDPHRWM
>2286-1144_1	>1302-2281_1	>2131-393_1	>2452-505_1
VHTGGLA	HLSTKGV	DSWQKMR	TYTGFSV
>2288-1143_1	>1314-2258_1	>2139-392_1	>2456-504_1
IKISYPS	TWWMREL	RPPMFDS	EFAISST
>2291-1143_1	>1316-2254_1	>2141-392_1	>2467-502_1
MPGRGGV	FYSPSWT	HYKASLM	WKSGLT
>2292-1143_1	>1323-2246_1	>2151-390_1	>2472-501_1
ATFGVWT	WPIFWH*	HMTPYPM	GLIVSPM
>2302-1137_1	>1343-2220_1	>2152-390_1	>2487-499_1
NHVQGLP	QLTLPSW	RIPMFDN	NTTTLES
>2304-1134_1	>1356-2206_1	>2163-386_1	>2497-498_1
QFMSQMA	YSGQFNW	MLMPYPL	FQRSTAH

>2309-1131_1 IKLQPSV	>1360-2203_1 NFAFPSW	>2176-384_1 NQAKWVH	>2503-496_1 SHQDNLL
>2321-1124_1 ETTQRPM	>1361-2202_1 HFNEPSW	>2184-383_1 SIPQASF	>2507-495_1 HQRPFVP
>2328-1117_1 GPYHAQP	>1362-2200_1 ISTHYYL	>2191-381_1 WGLHTFR	>2509-494_1 FMILPGN
>2346-1106_1 FTMPPWE	>1364-2197_1 IPLNYNV	>2202-378_1 SLLMPLQ	>2511-494_1 YEGSIPH
>2352-1103_1 FPFKWLA	>1367-2190_1 FISPSWT	>2215-374_1 LSPTYLE	>2516-493_1 WRHVP*T
>2354-1102_1 GASRAPA	>1370-2186_1 NNNPQWQ	>2216-374_1 RVPMNDG	>2527-490_1 LKLLDSY
>2384-1088_1 ADHVSHG	>1372-2183_1 FVAPSWN	>2217-374_1 MTDSRAL	>2528-490_1 QLAYSTR
>2387-1085_1 VTSNFSH	>1378-2172_1 LVWPTWA	>2220-373_1 WGVHQWR	>2529-490_1 GVHYTSA
>2388-1085_1 NSLTSYT	>1384-2166_1 WSWPSWV	>2221-373_1 KFDGPMQ	>2530-490_1 IHMSPLL
>2389-1085_1 MLERFLS	>1385-2165_1 SLIQPSW	>2229-372_1 HLNANSK	>2548-486_1 HVFPCCH
>2397-1080_1 HYFRING	>1388-2162_1 QVLPTWS	>2230-371_1 VYRSVEH	>2549-486_1 TYPERSA
>2401-1078_1 TFTYPTW	>1389-2161_1 TYVYPPW	>2235-371_1 HPFNRTS	>2559-484_1 IPRWYVK
>2406-1074_1 HTHWFHT	>1390-2160_1 TNPMW*M	>2238-370_1 SF*TDHL	>2565-483_1 *TSHKYL
>2414-1070_1 DPKGAYT	>1391-2160_1 HSLEDVY	>2245-369_1 VGMHTWN	>2569-482_1 LARITAE
>2433-1061_1 YSSMHID	>1397-2147_1 LVNPPWA	>2252-369_1 GVAHMTT	>2586-480_1 VFAKANP
>2453-1052_1 IGQLVKL	>1398-2147_1 HMMRTNH	>2257-368_1 DQPMHWL	>2587-480_1 HVDGSLR
>2460-1050_1 DSEHMTS	>1399-2147_1 LDFPPTS	>2262-367_1 QWADHW*	>2610-477_1 YRFVNFQ
>2463-1048_1 WTWPSWA	>1402-2143_1 HVLPPWS	>2264-367_1 YGSAFTM	>2621-475_1 QWRKTES
>2465-1047_1 YTSQTIR	>1404-2141_1 WSEFRTL	>2266-367_1 SPMPYPL	>2624-474_1 MAFTNAM
>2473-1042_1 KFWQGV	>1408-2137_1 HSLEAIH	>2268-366_1 EHFSLHP	>2663-467_1 SFYLP
>2477-1041_1 SKQDIIA	>1409-2136_1 SHYQYPD	>2270-366_1 THSPYPM	>2664-467_1 LTCLHCN
>2490-1031_1 SGMDRHE	>1413-2130_1 HKGFPM	>2273-366_1 MRLPMHD	>2673-466_1 EHPTTQI
>2491-1031_1 VTANTRE	>1417-2124_1 LYLPPWQ	>2280-365_1 NHGPYPL	>2684-464_1 MGAMARV
>2503-1025_1 FSFTLTV	>1418-2123_1 LNLPTWM	>2294-364_1 SLMPYPL	>2697-460_1 AVMDKRN
>2505-1024_1 AKVAITI	>1419-2122_1 FLNPSWQ	>2303-363_1 GQPVHWF	>2699-459_1 DLFARSH
>2509-1022_1 FVRPTWH	>1421-2118_1 WFMPPWD	>2314-361_1 HPVPYPM	>2701-459_1 TLGTPWS
>2511-1021_1 VSGTHML	>1426-2107_1 DYLQNN	>2315-361_1 FPP*KHV	>2706-458_1 SQVAMIR
>2512-1021_1 SPFFWEL	>1427-2107_1 HSKPILT	>2323-360_1 VLTPYPF	>2709-457_1 GSSNLNL
>2517-1019_1 TARSVWI	>1429-2102_1 FIRPTWG	>2329-359_1 VVSASWL	>2719-455_1 EPLSIFS
>2518-1019_1 VSTRHLN	>1431-2098_1 YHWATWE	>2342-358_1 SVSETGG	>2742-452_1 YYPPAHR
>2534-1013_1	>1435-2096_1	>2362-354_1	>2743-452_1

WNAPVPN	HFTMPSW	QDPYPFS	NAFKMVR
>2547-1006_1	>1447-2086_1	>2369-353_1	>2744-452_1
NLDANLV	YPEWGHN	WRLPMSD	YQRPFPN
>2548-1006_1	>1448-2086_1	>2375-352_1	>2748-451_1
KPNFDLH	DQPWHWI	ELGSALL	VSQMRSP
>2551-1004_1	>1450-2080_1	>2380-352_1	>2772-447_1
GAAGDAY	YDMSLIG	TIGRYPV	GQMPPFRP
>2557-1002_1	>1455-2073_1	>2387-351_1	>2777-446_1
TMTRNFM	EFISPPW	KLSPYPL	SPPLGSS
>2575-995_1	>1479-2050_1	>2400-349_1	>2781-446_1
LRLNITW	QVLPSWA	GDQGNP	IVDTTWH
>2581-992_1	>1487-2036_1	>2424-346_1	>2783-445_1
RIPMNDT	HALEMIA	DQPRHWM	TVVMYPG
>2582-991_1	>1488-2036_1	>2426-346_1	>2784-445_1
YTGLQTQ	WSTYRTA	MPDPYPL	SPDPYLR
>2595-985_1	>1492-2032_1	>2437-344_1	>2791-445_1
CIFTCHD	VTNFKTN	VLGWNMN	QVYDRSK
>2596-985_1	>1499-2016_1	>2453-342_1	>2811-441_1
MDRSALW	LILPPWL	LQPHHCY	QGLSRPH
>2608-980_1	>1503-2007_1	>2468-340_1	>2815-440_1
TISKWEE	WSWPSWN	ASIFTTW	SLFPMSL
>2618-973_1	>1504-2007_1	>2471-340_1	>2818-440_1
DRYVSWH	TRLYITD	SWSNSVR	DTPLPPW
>2620-973_1	>1507-2004_1	>2473-339_1	>2826-439_1
EMRTERV	HIGQPRS	SSNLTDR	WMTDRVV
>2624-970_1	>1512-1997_1	>2478-339_1	>2829-437_1
DWNKAGD	VEDYMMR	QSKYPLD	KPPMSKY
>2625-970_1	>1518-1989_1	>2481-338_1	>2833-436_1
CHLYCEA	WQRLPSW	LPSLQVL	SQGMNLR
>2630-969_1	>1521-1984_1	>2485-338_1	>2838-435_1
GEWEYDH	YVRPPWS	LDRQALM	SHPWTSG
>2633-966_1	>1523-1983_1	>2489-338_1	>2841-434_1
WNGMIGM	MVNPTWM	NHMLSAK	WTTTWMY
>2634-966_1	>1527-1973_1	>2510-334_1	>2844-434_1
EQTQARV	LLWDPPF	NQPYHWL	MFSTATA
>2638-963_1	>1528-1973_1	>2533-331_1	>2857-432_1
YSTTHVD	NVSLQMF	YGLHTFH	WMTPRWL
>2644-962_1	>1529-1972_1	>2571-326_1	>2866-430_1
MPAHLNE	QWHEPSW	SADHRML	LPTVRNS
>2646-962_1	>1530-1970_1	>2575-326_1	>2871-429_1
SPFKWLA	LWWSPSW	WRTHTWA	TPKWFEV
>2656-958_1	>1538-1955_1	>2577-326_1	>2879-427_1
HISWRMM	NDAFASM	DQPWHWI	QHYDRSR
>2661-957_1	>1547-1951_1	>2578-326_1	>2880-427_1
QEYHLAS	YIMPSWS	HYSYRSI	HWYDRDR
>2667-955_1	>1549-1949_1	>2579-325_1	>2904-423_1
LLGGAVA	FEWPSWG	NTFAFNH	HHFTLTI
>2670-954_1	>1556-1936_1	>2592-324_1	>2909-423_1
RIPMTDE	FMNPSWK	QYPYPLD	THLLIYM
>2675-953_1	>1562-1930_1	>2594-324_1	>2911-422_1
NICIRCP	FIKPSWV	QWEYHW*	YPWSEQR
>2677-952_1	>1563-1929_1	>2595-324_1	>2927-420_1
QNWVPFN	HHIDSLW	MTPLDAR	ICWACSP
>2680-951_1	>1570-1922_1	>2610-322_1	>2928-420_1
CYMCYQT	WSLHRTG	NYWTQHR	TGSSWAP
>2682-951_1	>1579-1912_1	>2620-321_1	>2930-420_1
SVVLNWN	YNWVQQS	TNATTEL	HLQTSYT
>2684-950_1	>1580-1912_1	>2632-319_1	>2938-419_1
QFMYQPV	FRMPSWT	TRAPMHD	GKLTNVR
>2687-948_1	>1586-1908_1	>2633-319_1	>2943-418_1
YTTLHRD	NTVVLHL	HPFNRSR	VGFEYHP

>2688-947_1 LANMSLL >2703-942_1 FVEPRSH >2704-942_1 SSGELWN >2712-939_1 MTDSHVS >2716-937_1 SYQMDYA >2733-932_1 AAKLLPH >2738-931_1 GRIGLQV >2741-929_1 QYMN*NF >2742-929_1 MPALVKF >2751-925_1 NLTHRGP >2757-922_1 LSLNMHD >2759-921_1 VLDFTLA >2762-920_1 TSYVGIY >2769-917_1 YSVLHLD >2778-914_1 YTSNHL >2781-912_1 KFVVKHT >2785-911_1 HSREFTF >2788-908_1 SHRYDPG >2792-907_1 DDRLNAL >2796-905_1 RCESMQC >2814-899_1 LHLRLSA >2819-897_1 SKPNTVS >2823-896_1 AHLAADM >2824-896_1 GASHHSK >2829-895_1 *TPTQQY >2831-894_1 ESQSLCV >2832-894_1 TNPLELP >2835-892_1 KPQLNYF >2838-891_1 GLLTYGT >2849-887_1 SHTNPKS >2851-887_1	>1590-1905_1 YVLPSWM >1599-1895_1 HSLTIRN >1603-1892_1 TYPWVTQ >1620-1878_1 NWLPTWE >1623-1877_1 FLNPSWE >1627-1875_1 NWMWDAN >1635-1862_1 AFIGPPW >1637-1860_1 EIAMTTL >1640-1855_1 LHHGVPV >1644-1852_1 INLPSWM >1650-1840_1 WTYPPWG >1657-1833_1 HFTHPSC >1661-1828_1 IIGPGWF >1662-1825_1 WHTSIPL >1664-1820_1 TFTNPPW >1665-1819_1 GHYLPLS >1666-1819_1 AQWYPFS >1670-1814_1 FYLPTWY >1675-1809_1 AWDLFRA >1676-1808_1 QWWNREL >1678-1808_1 QTGDSTH >1681-1807_1 LHKGLAV >1684-1804_1 VFQRPSW >1686-1803_1 SWNPKER >1692-1798_1 FTPLWHL >1698-1792_1 VFWDSL >1705-1779_1 SFWPFPS >1706-1779_1 WIIPPWL >1710-1773_1 IVAHVVA >1712-1770_1 SALPKSL >1713-1770_1	>2641-318_1 QWEDHCY >2653-316_1 QLTPYPL >2654-316_1 LHPHHWY >2674-313_1 YTFFPAS >2680-312_1 QWRSHWN >2681-312_1 WPLPYPM >2687-312_1 YEGGSMG >2693-311_1 RVPMVDG >2707-310_1 DPTPYPL >2710-309_1 NAKTSHM >2711-309_1 TDGQLRR >2718-308_1 SLEPYPI >2719-308_1 WMPAMKG >2746-304_1 RIPMVD >2748-304_1 NTALVPF >2754-304_1 KVTNAQM >2763-302_1 GQPLHWA >2764-302_1 NPDPYPI >2772-301_1 NWPYMTW >2781-300_1 NTYKLSP >2785-300_1 DTLDALV >2795-299_1 YHYASMS >2800-298_1 TERHPRL >2811-296_1 KLDPYPL >2817-296_1 SLMPGRP >2823-295_1 SVLLSTK >2831-294_1 SALVALD >2832-294_1 LATPISD >2836-294_1 RLPFIDS >2841-294_1 SGSHKMS >2858-292_1	>2950-418_1 RIPIH*N >2958-417_1 VLYTKNL >2960-417_1 DGGYWHV >2961-416_1 AISLTPR >2962-416_1 SLPT*GS >2969-416_1 RAPVNSV >2973-414_1 QEYNRER >3003-411_1 YPNTDMT >3019-408_1 LDNALDR >3020-408_1 DGKQVSI >3027-407_1 YVRLAVT >3039-405_1 NTQHVT >3040-405_1 YCYWCNP >3046-404_1 SLEDEL >3048-404_1 HEHELKS >3060-401_1 GYSLHHR >3070-400_1 AADNMRS >3074-400_1 SYLLPIA >3079-399_1 IGPASEA >3083-399_1 SWSYYSR >3088-397_1 WMYYPYE >3091-397_1 WGTLGLR >3100-395_1 RPAITTM >3103-395_1 GDKTSPF >3108-395_1 GFGENSG >3115-394_1 ARTVFSY >3125-392_1 REVMHWP >3127-392_1 IQYTEHP >3133-391_1 WTSHHHY >3139-390_1 HAMTEAP >3140-390_1
--	---	--	---

HLKAQFN	WVAPSWQ	*TAYPLT	VTSYFST
>2879-876_1	>1729-1752_1	>2861-291_1	>3141-390_1
FGSRVPP	EYWPVPF	YPGPYPL	MVPRIHP
>2887-872_1	>1738-1746_1	>2868-290_1	>3155-388_1
QNNPYLH	TTPTNTY	YLSPYPF	GMTLATL
>2896-867_1	>1740-1745_1	>2869-290_1	>3156-388_1
VYNSSHM	SYPWMEQ	TSI*TQT	TTKFIK
>2898-866_1	>1741-1743_1	>2886-288_1	>3161-388_1
TVVMYPG	IRMPSWA	KPVSGHC	KWMDHQW
>2900-866_1	>1747-1738_1	>2899-287_1	>3170-386_1
FWYPTFL	FNTPPWE	VGGSLVG	AGAQDSM
>2902-865_1	>1748-1737_1	>2902-286_1	>3182-385_1
WNWPTWE	FTVPPWD	AAMHFNW	LPSPIIT
>2905-865_1	>1753-1733_1	>2909-286_1	>3186-384_1
FSLAGGR	TMQTFSR	AIGASDH	EPTLSKP
>2914-862_1	>1754-1731_1	>2915-285_1	>3187-384_1
IPTWMTK	SWTWPSW	YETWEHR	EQEYYPF
>2919-861_1	>1755-1731_1	>2918-285_1	>3188-384_1
QFMAQKY	AWVSPPW	GPTYPI	FQAPFRP
>2931-856_1	>1758-1728_1	>2927-284_1	>3191-383_1
HPATICR	TDPLSIT	SGKVRPN	TDWLWVE
>2934-856_1	>1767-1718_1	>2928-284_1	>3196-382_1
MLVEATD	HLSVGSL	STHTQPS	RTDYPYV
>2936-855_1	>1771-1715_1	>2929-284_1	>3205-382_1
RTPSELK	FMLPTWN	NGLWHAD	DPWEWSF
>2937-855_1	>1778-1708_1	>2936-284_1	>3211-381_1
HLLSLAR	HSLELIH	TLQPYPL	TQKPFPR
>2941-854_1	>1781-1704_1	>2939-284_1	>3212-381_1
NGTVGYL	DRWNTNV	YGLHSWR	FPPNLYT
>2946-852_1	>1783-1703_1	>2951-282_1	>3222-380_1
LKATVSI	QFLPPPW	ATRPPWI	TDTGANW
>2950-850_1	>1794-1690_1	>2959-281_1	>3234-379_1
GGIKVTL	TADAGVT	WGVHQWS	YMASTHT
>2952-850_1	>1795-1689_1	>2971-280_1	>3235-379_1
HKLHLNI	FFVGPPW	FNAPRPS	FHEHMTA
>2957-849_1	>1797-1686_1	>2977-279_1	>3243-378_1
TVFPWMI	LSNPQWQ	LPMPYPL	SLGRIVT
>2959-847_1	>1805-1677_1	>2999-278_1	>3254-377_1
NDPAGHF	FQLPSWN	HTRMSPF	NIGLLAN
>2967-845_1	>1815-1666_1	>3006-277_1	>3263-376_1
AVHLPYR	NVLPPWA	QSPYPML	TSSMTYL
>2994-834_1	>1818-1665_1	>3010-277_1	>3270-375_1
SIVMART	NPSLQTF	RLPMIDA	TAIWQAS
>3000-831_1	>1824-1660_1	>3015-277_1	>3275-374_1
VGWIGKL	DLHVFSN	TLGLIKL	GLSVRPT
>3001-831_1	>1827-1658_1	>3018-276_1	>3279-374_1
QFMAQVS	FNQPTWL	HFRLHEP	GYWNLAY
>3015-827_1	>1829-1655_1	>3019-276_1	>3281-374_1
YESTPEW	TVSHANG	NGSYTWI	GVQKASN
>3031-822_1	>1838-1644_1	>3025-276_1	>3295-372_1
VLLPTRI	AFRAPTW	RIPMDDP	AVFTRQP
>3033-820_1	>1841-1642_1	>3028-276_1	>3297-371_1
QYFASSD	HSLEWLS	RLPMQDY	SYNGLAS
>3054-813_1	>1842-1641_1	>3032-275_1	>3308-369_1
RPITFHL	DYSLPWW	YPHNLPP	SPMYWGN
>3057-813_1	>1855-1627_1	>3034-275_1	>3312-369_1
YTCSLCG	SFSYPTW	SWSPSGT	GHL*NAL
>3065-810_1	>1857-1623_1	>3035-275_1	>3330-367_1
HAMATTV	APVIGLY	RIPMNDP	NPFTNSA
>3068-809_1	>1867-1617_1	>3044-274_1	>3333-367_1
GWWPLLK	LLREPNY	KPVHPSF	HGPNTST

>3078-806_1 VDGYTNW	>1877-1605_1 FTTPSWH	>3046-274_1 HPHNQTQ	>3339-366_1 MTLSRVP
>3085-804_1 TDSMKNM	>1882-1600_1 QPFESWS	>3051-273_1 TISDWTS	>3352-364_1 TFTATFS
>3088-804_1 WPIPPHS	>1894-1587_1 HALEFIS	>3060-273_1 HWSQLQA	>3353-364_1 GSIITHH
>3102-800_1 VSAAFVS	>1899-1584_1 HFLQPSW	>3065-272_1 GFSGTVW	>3354-364_1 VVTQRHL
>3108-799_1 NVFITHR	>1902-1579_1 SQVEPPW	>3068-272_1 HHIETAR	>3356-363_1 STSELYG
>3109-798_1 FTGRMAI	>1916-1565_1 VDSVHVR	>3082-271_1 WGTRIHA	>3365-362_1 DQRPFKP
>3119-795_1 LVRVTVN	>1919-1565_1 QMNENTY	>3090-270_1 EAARILL	>3367-362_1 ELGAYKL
>3120-794_1 TVLRHDT	>1924-1559_1 QPNVNNY	>3098-269_1 QWEDHWH	>3388-359_1 VALWPHP
>3123-793_1 TTVMAWN	>1927-1555_1 WSMFRTL	>3107-268_1 RPSTEPP	>3389-359_1 IVYLEHP
>3124-792_1 CYMCMTF	>1932-1550_1 HHVQNSY	>3112-268_1 KTLLQND	>3403-358_1 ALNISGL
>3137-789_1 QYMF*MT	>1933-1549_1 GGIMLWW	>3113-268_1 VFSGVPL	>3422-356_1 MPDAVNQ
>3143-786_1 LCYLFCE	>1941-1543_1 FQVGPLT	>3119-268_1 GANRETL	>3425-355_1 GAQLYDN
>3156-782_1 SHTIVQH	>1948-1539_1 FTSPTWM	>3127-267_1 SLSAYPF	>3430-354_1 ELTRFAK
>3157-782_1 TADPTIR	>1949-1539_1 KPLMQSL	>3135-266_1 LIAYDGL	>3449-352_1 YNGKALS
>3161-782_1 YEATLVL	>1952-1538_1 SFTQPTW	>3139-266_1 RVPFTDM	>3458-350_1 LGSLVRG
>3163-781_1 WDAWWRH	>1970-1520_1 LFLPPWS	>3146-265_1 SNLKMGY	>3464-350_1 IVMPQEW
>3165-781_1 WSTTHYN	>1973-1519_1 FPPILHI	>3149-265_1 RLPFTDT	>3466-349_1 QFKTSNT
>3168-780_1 HFRMTSA	>1974-1518_1 NQFMPSW	>3166-264_1 YPAKFSN	>3470-349_1 HQAFFMK
>3169-780_1 HSSFHDM	>1992-1500_1 SFKFPTW	>3172-263_1 RTPMSDV	>3479-348_1 SGSSGHR
>3186-776_1 WWWPPTW	>1993-1499_1 NNGPISK	>3184-262_1 VNREADID	>3483-348_1 TLPSYRN
>3196-773_1 VRATISI	>2013-1482_1 DTPMWHL	>3185-262_1 WSVTLDH	>3489-347_1 YYQAELP
>3199-773_1 NKTMLEL	>2014-1482_1 LPWTDSS	>3187-262_1 ANVDPYS	>3508-345_1 TWAVNTN
>3201-772_1 VI*TITR	>2018-1480_1 RFIMPSW	>3190-261_1 GPLPYPL	>3513-345_1 FKQNLVM
>3204-772_1 LKGSSPL	>2021-1478_1 LTKISTP	>3191-261_1 VGGDYIM	>3517-344_1 FQVPFRV
>3209-771_1 AFVQPSW	>2022-1477_1 YIIPPWL	>3192-261_1 RPPMQDI	>3525-343_1 ASRLTNL
>3212-771_1 SFGPCCS	>2023-1476_1 YVMPSWT	>3207-260_1 WTPMHDP	>3529-342_1 WHQTARE
>3218-770_1 QFMHQTL	>2024-1476_1 QQLLPPW	>3210-260_1 QWEDHL*	>3534-342_1 AGQLKHE
>3220-769_1 SHDNDRQ	>2042-1466_1 QFTAPSW	>3223-259_1 SPFQSRL	>3549-339_1 NILWKSS
>3225-768_1 ARDSSFV	>2043-1464_1 LEAHILL	>3225-259_1 LIPAHSG	>3551-339_1 NPRTWLD
>3233-765_1 VGPGIGK	>2045-1456_1 FMRPSWM	>3226-259_1 TLTPYPL	>3554-339_1 DTKHNNL
>3236-764_1	>2051-1450_1	>3232-258_1	>3567-337_1

FVLPSWI	LYIEPSW	SPGLRSQ	EYSLVAS
>3240-763_1	>2057-1447_1	>3242-258_1	>3573-337_1
DAVQPPI	LWWQPYP	HLAYLMM	STEAWSA
>3242-762_1	>2060-1443_1	>3257-256_1	>3578-336_1
MAAKLVT	KTSVDYD	AHKMAGN	TSIFTRF
>3245-761_1	>2062-1441_1	>3261-256_1	>3587-335_1
TTSWPPL	TNDRYMI	SGVKHDQ	QLFTTYV
>3253-759_1	>2063-1441_1	>3267-256_1	>3596-334_1
QPYNMNW	HHTSRPI	ELSAYPL	GIGGTPN
>3254-759_1	>2065-1439_1	>3284-254_1	>3597-334_1
TIPKEHY	NEQMPLY	TEWKYPL	WPMSVPR
>3256-758_1	>2068-1438_1	>3288-254_1	>3598-333_1
NVWNYSR	YYMPSWS	TPQPYPL	AFWPTFV
>3260-756_1	>2069-1437_1	>3289-254_1	>3610-331_1
LAARIPK	LELPPWL	ESMWIVT	ALLGKSP
>3266-755_1	>2070-1437_1	>3298-253_1	>3622-330_1
GYSNTGK	EYVWSRS	ADP*CQ*	SYPREGY
>3292-748_1	>2082-1430_1	>3309-251_1	>3624-330_1
TFVMPSC	DLHNWGQ	STTSPSK	GSENNHH
>3296-746_1	>2083-1429_1	>3311-251_1	>3629-329_1
ATESTQL	VLASVYR	STVKTEG	VSIHLSL
>3310-743_1	>2084-1429_1	>3314-251_1	>3631-328_1
YVLPSWM	YVFPSWY	FATRITF	VTLKHGG
>3312-741_1	>2085-1429_1	>3316-251_1	>3636-328_1
MASYEMS	GPEHTNP	RRVNH*P	ATPATYY
>3313-741_1	>2100-1418_1	>3328-250_1	>3643-327_1
MLDILHL	EWMPSWA	KLAHACI	SLKLAFR
>3317-738_1	>2111-1412_1	>3330-250_1	>3655-326_1
GAQVLYP	FIEPSWW	QNHDTLK	SQGHMMI
>3326-734_1	>2116-1409_1	>3344-248_1	>3659-326_1
YETSMHQ	FQLPPWH	QTPYPHN	SKHSAML
>3327-734_1	>2121-1407_1	>3347-248_1	>3660-326_1
SAYVADK	KFTIPSW	YNPWTHC	DYTRERW
>3343-730_1	>2123-1407_1	>3350-248_1	>3667-325_1
VIPPWWS	SQISPSW	NVMVALL	WSTSISG
>3355-727_1	>2125-1405_1	>3354-248_1	>3668-325_1
AHHLAKD	IFLPSWL	NLSGWKW	VERWYLQ
>3358-726_1	>2131-1403_1	>3357-247_1	>3679-324_1
FVIPSWA	FVYPTWT	VGVHTWS	NEDWFAI
>3365-724_1	>2134-1403_1	>3360-247_1	>3680-323_1
LYRPDYD	LESASWR	SPVVGTE	VQIMWPW
>3371-722_1	>2135-1402_1	>3364-247_1	>3684-323_1
TLFSLNV	NLSSVHY	VMTPYPM	WGTRIHA
>3375-720_1	>2140-1399_1	>3377-246_1	>3686-323_1
FTSRVHV	RLVLPPW	SATPYPL	FMDAHIA
>3380-718_1	>2141-1399_1	>3382-245_1	>3688-322_1
HAAHQPP	EPARH*N	HISKGMT	STFYPPM
>3384-718_1	>2144-1397_1	>3390-245_1	>3690-322_1
TLRAPET	NRLWDSF	TRVPMVD	KLNYAGM
>3385-717_1	>2150-1390_1	>3391-245_1	>3701-320_1
AFQQRAP	TYSLPWW	HPHNGNR	APTATIF
>3389-716_1	>2156-1387_1	>3407-243_1	>3703-320_1
VYNHVG	EHLGSYF	HWEHHWR	TLGMDAE
>3394-715_1	>2159-1385_1	>3416-243_1	>3704-320_1
NEPARHS	WSEFRTK	YWWTHR	MTPEGPL
>3397-714_1	>2161-1384_1	>3421-242_1	>3708-320_1
DKMPRDS	LFSEPSW	THFRQGF	GSTRTHD
>3399-713_1	>2168-1379_1	>3425-242_1	>3714-319_1
HSSFITV	HNVSPPW	QVSQSQP	SFLLPVF
>3404-712_1	>2169-1379_1	>3426-242_1	>3721-319_1
IQWPTWL	HSLQFIG	TVSPYPM	LGSPLPI

>3405-712_1 FSLPTWL	>2170-1379_1 HSLQLIF	>3433-241_1 HINPEAR	>3760-315_1 WSVTLDH
>3406-711_1 WHTNNTV	>2179-1375_1 TSTFSDK	>3437-240_1 VTYTRYT	>3768-314_1 YGLASTA
>3410-710_1 DVPNWRT	>2181-1374_1 QHVMPSW	>3442-240_1 EQATLSP	>3769-313_1 NLFGFPI
>3415-708_1 SKFNLAY	>2182-1373_1 NFLRPSW	>3444-240_1 QWEDH**	>3774-313_1 LSNQPPS
>3416-708_1 AQTRHH	>2188-1370_1 YAPPYQR	>3449-240_1 TQQKHDL	>3780-313_1 QFNVIYV
>3420-708_1 SGLDLWT	>2189-1368_1 SYTVHIL	>3452-239_1 NNELKIM	>3789-312_1 FLWTSTS
>3429-706_1 TPDGSGL	>2190-1366_1 YIFPTWA	>3453-239_1 AHPHRGG	>3792-312_1 DRALFFP
>3432-706_1 QVSMLST	>2194-1362_1 SFTYPSW	>3465-239_1 AWSMLWS	>3795-312_1 VDITRMR
>3438-704_1 KLISWS	>2199-1360_1 KFISPPW	>3474-238_1 EPSKYPL	>3799-312_1 ALLPSNY
>3442-703_1 ITKMSG	>2200-1359_1 GQFMPPW	>3480-237_1 VVNQIFI	>3803-312_1 YPILRNA
>3451-701_1 QFMFQDN	>2203-1358_1 IYYPSWV	>3481-237_1 WTPMQDP	>3805-311_1 STCSLLR
>3456-701_1 GDRSVLW	>2204-1357_1 LTLPPWL	>3490-237_1 IYADIGF	>3812-310_1 TSLPSW
>3459-700_1 LPLTYDA	>2206-1356_1 NLIPPPW	>3491-237_1 YTSMFTY	>3820-309_1 TLYVRTK
>3460-700_1 TFASVNA	>2214-1349_1 RQLDVIA	>3496-236_1 AVRWTVP	>3838-308_1 QHFPFAF
>3464-699_1 LVARVAG	>2216-1347_1 GLMLPPW	>3502-236_1 TQGYATS	>3856-307_1 SFLLPVH
>3473-696_1 TSLRIHL	>2219-1345_1 TGTYGWT	>3506-235_1 SGLYHFH	>3867-306_1 TRPLAHM
>3475-696_1 VTSMSDN	>2222-1343_1 NPMWQGR	>3507-235_1 TMPYLRI	>3883-304_1 ESLGEEP
>3477-696_1 MTVAVTS	>2224-1340_1 SFWPFPK	>3511-235_1 VQAGVWP	>3884-304_1 TKTHSNL
>3484-695_1 MSSSLKS	>2232-1334_1 YNFPWA	>3519-235_1 GLMPYPL	>3887-304_1 SYVKIEL
>3491-693_1 SLH*TIR	>2237-1330_1 QPDSNVA	>3530-234_1 RIPMSDS	>3896-304_1 WNSIEST
>3500-692_1 WTSYLG	>2242-1327_1 HGMFLD	>3531-234_1 WHWTDLK	>3911-302_1 KTHPSCP
>3505-690_1 WTTLAFY	>2244-1326_1 VLIPPPW	>3532-234_1 GSIWFSV	>3915-302_1 VYVGPD
>3509-689_1 HPKAPEP	>2246-1325_1 YVLPPWL	>3535-233_1 VYHPTSV	>3920-302_1 HDTRKPL
>3514-689_1 QYMFQWA	>2248-1323_1 DQFYPLH	>3540-233_1 RDPVPIP	>3931-301_1 VSNPFHL
>3523-687_1 GGGLLPS	>2249-1322_1 NFIDPPW	>3543-233_1 EPGPYPL	>3953-300_1 VVPCAGL
>3524-687_1 HLTLVAS	>2250-1321_1 QFLPPWS	>3551-233_1 YRTPMID	>3957-300_1 QLHAGQY
>3525-687_1 LTMTYDS	>2255-1321_1 QTVKNTF	>3554-232_1 ASSYSMY	>3960-299_1 ASWTDYR
>3531-684_1 LTYRAAA	>2261-1318_1 IVHPTWL	>3562-232_1 WPSAYPM	>3961-299_1 FLFLMVM
>3536-683_1 GSFGIYT	>2264-1316_1 YPSDHRR	>3570-231_1 DQPRHWP	>3963-299_1 VGMRHPA
>3538-683_1 LIATPHR	>2265-1316_1 APTRDQS	>3571-231_1 ETIRNIG	>3964-299_1 VVERTIR
>3540-682_1	>2267-1316_1	>3579-231_1	>3974-298_1

AEHYCCV	DYWPFFS	QWEDQWY	A*GY PDT
>3543-681_1	>2268-1315_1	>3581-230_1	>3984-298_1
TYFAEHK	WSLSWFA	GANNERW	HT*DKPK
>3551-679_1	>2271-1315_1	>3583-230_1	>3990-297_1
SWPWYYQ	TQNVRLS	RFP MYDN	YSANHGA
>3557-678_1	>2273-1313_1	>3587-230_1	>3992-297_1
LAPRAVS	TSYVGIY	SYWNEHR	LNARDYT
>3560-678_1	>2275-1312_1	>3591-230_1	>3994-297_1
>3565-676_1	LILPTWG	RPPMTEG	GKQYVMA
GSAVPWT	>2283-1306_1	>3592-230_1	>4017-295_1
>3568-676_1	WSLIPLY	RPPMTDF	*MHYHHL
HTSLTMR	>2284-1305_1	>3595-229_1	>4020-295_1
>3569-675_1	FHNPTWN	NWMEHWH	TFKCLHN
FPRYHYL	>2291-1301_1	>3596-229_1	>4025-294_1
>3576-674_1	FWGPSWM	GL*THPL	ILHELNP
KSSYDFM	>2292-1299_1	>3599-229_1	>4026-294_1
>3578-673_1	HTAEAVR	RPPFMDG	TRVTQYL
VAHPSMP	>2293-1299_1	>3602-229_1	>4033-294_1
>3588-671_1	SQDISAY	ALKVQKT	HPGYRVH
QAAASHS	>2299-1297_1	>3627-227_1	>4042-293_1
>3589-670_1	ADNPMWQ	SQFYTFW	ITVHEHP
WPGLIPV	>2311-1290_1	>3631-227_1	>4048-292_1
>3595-669_1	EVYPPWM	IGLYTFN	TLMDLFM
NYMLPWA	>2312-1289_1	>3636-226_1	>4050-292_1
>3601-666_1	DLKSSML	FCTLSCI	AGANHLY
YWYQHRF	>2322-1282_1	>3637-226_1	>4059-291_1
>3609-664_1	NFVDPSW	NDPAGHF	ALGITQV
WTAGLQI	>2326-1280_1	>3638-226_1	>4066-291_1
>3611-664_1	ANPNGHD	H*NADLY	NGAYLRC
YWSEHRA	>2334-1273_1	>3641-226_1	>4071-290_1
>3614-662_1	TLGHSRQ	NPLKFTS	EQTPFRV
SWSDFLK	>2336-1270_1	>3642-226_1	>4081-289_1
>3619-661_1	FIGPSWD	QREDHW*	YANFTPS
VAPDHKR	>2337-1270_1	>3643-226_1	>4086-289_1
>3620-660_1	SRDFVYG	QVSWWMR	VGSSF*N
QEMGSAR	>2339-1269_1	>3648-225_1	>4090-288_1
>3631-657_1	YFTNPTW	YRSVTPL	GKLYIKS
SGRADET	>2343-1264_1	>3652-225_1	>4092-288_1
>3646-654_1	NIDRDAM	FHTAWHP	YRLAPKM
HAKVGIY	>2348-1261_1	>3655-225_1	>4095-288_1
>3647-653_1	YTPLWHM	FVANAPG	AVHGHLI
WNIEHLL	>2350-1259_1	>3659-225_1	>4119-285_1
>3680-648_1	VFWNPSW	RHLVTHV	VERWYLM
GTFGIRT	>2351-1259_1	>3668-224_1	>4123-285_1
>3681-648_1	WVQEWYD	SMDPYPL	T*MPFHP
YTHNHHD	>2358-1255_1	>3670-224_1	>4131-284_1
>3682-647_1	SFFHPSW	RLPFTDH	TQRQLTI
YTTQHLD	>2359-1254_1	>3679-223_1	>4133-284_1
>3689-645_1	FAYPTWW	LTTREML	YPPL*AW
VPDPTQQ	>2360-1254_1	>3682-223_1	>4144-283_1
>3693-644_1	FTPPPWQ	LTPTHIS	SFWLPVL
HMVTAGK	>2366-1252_1	>3684-223_1	>4145-283_1
>3696-644_1	LGANLWW	HTLV TAR	WPFASAF
QPVEIRR	>2368-1251_1	>3686-222_1	>4151-282_1
>3707-641_1	EYWPIWT	VIHY*LT	MGEYHTR
YTSSHVD	>2369-1250_1	>3689-222_1	>4155-282_1
>3713-639_1	HSFLIAQ	FEC*RDE	GTSFSGH
SSSEYTS	>2370-1250_1	>3691-222_1	>4162-281_1
>3716-639_1	LSAREPY	THTPYPM	LMVTVCP
VPATVKY	>2371-1250_1	>3696-222_1	>4165-281_1
>3725-637_1	HSLFFPS	KGMHFFN	TYKDLVQ

GSAQPSR	>2372-1249_1	>3698-222_1	>4172-280_1
>3729-637_1	FIIPSWG	SISYFIT	RCASEIP
NLFAIYM	>2374-1249_1	>3704-222_1	>4173-280_1
>3730-637_1	MFTDPPW	RQKQPYG	NISDYGR
ATDSFAT	>2375-1248_1	>3712-221_1	>4182-279_1
>3736-636_1	TLGNLSL	GRWDLHS	IDTEYKL
GTHPREW	>2381-1245_1	>3713-221_1	>4184-279_1
>3756-632_1	YPRVAST	GHVSWHK	DAQTLTR
KFFVKYT	>2382-1245_1	>3714-221_1	>4187-279_1
>3757-632_1	FMMPPWN	SPEKYPL	SWWPLFP
AWDSRML	>2383-1244_1	>3718-221_1	>4188-279_1
>3767-630_1	WWNPPWW	TLTPYPV	YTPPKAY
TLGHSRQ	>2387-1242_1	>3722-220_1	>4189-279_1
>3769-630_1	SFYDWDG	LHVAWHA	ERNVQDR
QMMENTY	>2390-1240_1	>3725-220_1	>4196-278_1
>3776-628_1	TYLPSWL	*ELKNGT	WLNKLPM
DFIPHTL	>2395-1237_1	>3726-220_1	>4203-278_1
>3785-626_1	HGQDISA	TVSMRPP	GVANARF
FVHPSWS	>2399-1234_1	>3733-220_1	>4205-277_1
>3795-624_1	YTKPMSM	PRD*SA*	FSGHPAT
DKHRPFW	>2405-1233_1	>3736-220_1	>4215-277_1
>3800-622_1	FINPTWA	SPPLAL*	RLLDGLL
YSTLH*D	>2429-1220_1	>3741-220_1	>4217-276_1
>3808-619_1	AFVVPWS	SSNASLH	SSTVNLR
NMPAGVR	>2431-1220_1	>3757-219_1	>4219-276_1
>3809-619_1	SFVDPSW	RIPMFES	NWLPMTF
AGFRLVL	>2432-1219_1	>3759-219_1	>4222-276_1
>3810-619_1	YVSPPWG	HSQVTQQ	DRALHWP
YSWKHYD	>2434-1219_1	>3761-219_1	>4224-276_1
>3815-618_1	NYVHPSW	YPPGNSW	VNPWMPH
ASTEWLA	>2440-1216_1	>3775-218_1	>4226-275_1
>3818-617_1	SSCRRGR	WPTRYPL	HQWMSLT
NTRIIVS	>2446-1213_1	>3781-218_1	>4227-275_1
>3832-614_1	YNLPSWL	KLTATPM	TRTLWTT
GWHHKLI	>2447-1212_1	>3785-217_1	>4228-275_1
>3835-613_1	HFNEPTW	*ISPYPL	SSGTRHD
ADHLNVL	>2448-1212_1	>3808-216_1	>4230-275_1
>3837-613_1	GNPQWQV	GVERPIR	DSANPVN
GWMTTFS	>2451-1211_1	>3810-216_1	>4231-275_1
>3844-611_1	SVVSYSS	WIYTDLK	FLPTPQV
YSYKGYA	>2455-1208_1	>3823-215_1	>4239-274_1
>3848-610_1	WNPFWYT	TSVGSPD	T*MPFTP
QFMYQPS	>2460-1205_1	>3824-215_1	>4246-274_1
>3850-610_1	GKTDLGI	RR*QMYP	IASSVTQ
GWMVTFM	>2462-1203_1	>3825-215_1	>4247-273_1
>3851-610_1	QVSNRFV	FPWKGHV	MPP*KLV
DRAPHQL	>2468-1199_1	>3830-214_1	>4256-273_1
>3854-609_1	TPVGPGV	DTRFVQT	KPWTLTG
LQLQVLV	>2470-1198_1	>3832-214_1	>4264-272_1
>3856-609_1	QFVSPPW	FLTPYPI	PLILRLL
SIWHLGA	>2476-1194_1	>3848-213_1	>4265-272_1
130-132	GPLWHLS	LRVPMHD	VPRADVL
>3859-608_1	>2479-1193_1	>3858-212_1	>4285-271_1
TQVFMVE	LWWYPSV	FHEPYPF	EPMDRYI
>3860-608_1	>2482-1188_1	>3862-212_1	>4288-271_1
EQTWAMR	TPMQNTN	YTKPMSM	GPPWGSA
>3864-608_1	>2483-1188_1	>3865-212_1	>4301-269_1
YPFKWLN	YPWFITE	LGLHTWS	KSPSNIR
>3872-606_1	>2486-1187_1	>3869-212_1	>4303-269_1
HLLNGVN	TFWPVWE	TPQNTRI	SFLLPVL
>3874-606_1	>2492-1184_1	>3881-211_1	>4306-269_1

CYVWRCD	SGGGRSL	VLEAYPM	GLRDICC
>3883-604_1	>2497-1182_1	>3883-211_1	>4316-268_1
EGFWLAI	NVALQSF	HPRAATD	VEWCSSQ
>3884-604_1	>2501-1180_1	>3889-210_1	>4327-268_1
AAHQKEQ	WYIQPSW	HPHNRPV	VYIIGAV
>3889-603_1	>2505-1176_1	>3892-210_1	>4336-267_1
MSVSLNA	IWHSPPW	FETYYSW	LMQQRQP
>3892-602_1	>2508-1175_1	>3897-210_1	>4338-267_1
HIITLKH	KQVTPPW	KSMVPNV	SVDLTLK
>3897-601_1	>2512-1173_1	>3899-210_1	>4342-267_1
IRQPQFN	YPWAQMG	VRHVPHA	QAMPLND
>3898-600_1	>2519-1170_1	>3909-209_1	>4349-266_1
YTSQTWR	HYKSDRT	QEYTWMN	WDIFMTY
>3915-596_1	>2521-1169_1	>3914-209_1	>4361-265_1
NVPVGTF	MPPWFAW	TFPLMRP	VGM*VLF
>3918-595_1	>2523-1168_1	>3921-209_1	>4367-265_1
CFMCYSD	FSAYWNT	*IPYPLS	NTLKAYN
>3921-595_1	>2524-1168_1	>3926-208_1	>4368-265_1
SGAERFQ	YNPFMDA	VPVHKYD	KSLEDGS
>3922-595_1	>2528-1166_1	>3927-208_1	>4369-265_1
NVVTTLR	ASTVLDL	VNLRYI	WDITQER
>3924-594_1	>2529-1165_1	>3937-208_1	>4372-265_1
VNFQIRV	SSLGTRI	HPFNRTG	WPAIRGQ
>3927-594_1	>2537-1160_1	>3939-208_1	>4374-264_1
STWELPR	HWVMPSW	SHVSWHE	YSYSRNT
>3932-593_1	>2540-1158_1	>3941-208_1	>4375-264_1
FKQHATA	NVALQLL	NTTNSHG	ASSARHY
>3934-593_1	>2543-1157_1	>3956-207_1	>4382-264_1
ASVHSVH	FVLPSWI	SPTRYPL	MLAGAPT
>3935-593_1	>2544-1156_1	>3959-207_1	>4383-264_1
ACHTRHC	SLIFPPW	APDPYPL	ATWLDRY
>3936-592_1	>2558-1150_1	>3969-206_1	>4385-264_1
TNTTGWR	FTSPSWA	NGSAKDP	HHQRQIS
>3947-589_1	>2559-1148_1	>3972-205_1	>4390-263_1
TMARIPR	FSCLLGG	SFGVRSV	ANSYYRY
>3950-589_1	>2563-1148_1	>3982-205_1	>4391-263_1
CYFKWQC	FKFPTW*	YDLIRRE	FHMETGM
>3951-589_1	>2565-1147_1	>3983-205_1	>4400-263_1
QSGILYT	VESLRMQ	NALFQTN	VTYDRER
>3954-588_1	>2566-1147_1	>3987-205_1	>4417-261_1
TQLNIWL	TLINLRV	VRWPMHD	TVTYLTV
>3955-588_1	>2570-1144_1	>3991-205_1	>4428-261_1
SPWPLMT	AWVSPSW	SFSQGRT	NWDHLTT
>3974-585_1	>2574-1142_1	>3999-204_1	>4431-260_1
SEPSMQQ	ASPMHVP	MNTIELK	LPNRLTQ
>3983-584_1	>2579-1141_1	>4000-204_1	>4432-260_1
VSYAHM	ASAPNPR	DAGMGQR	NMRSYAN
>3991-582_1	>2580-1141_1	>4014-204_1	>4433-260_1
GTRPNTV	DFSWPTW	HHT*LPK	HCTHCMP
>3993-582_1	>2581-1141_1	>4017-204_1	>4438-260_1
FTAAIYS	WWDPSWE	GVSLVVL	GVERPIR
>3996-581_1	>2582-1140_1	>4020-203_1	>4441-259_1
QALWFRV	FVDPPWR	VSSVYFP	YTELMEF
>3998-581_1	>2584-1139_1	>4027-203_1	>4445-259_1
WNSLVRS	LTDH*TW	WGAVHTT	RLKANQT
>4004-579_1	>2586-1139_1	>4030-203_1	>4457-259_1
DSFWRWP	LHKGHP	TEPSHKN	MIFSEDD
>4005-579_1	>2590-1136_1	>4036-202_1	>4458-259_1
RTDLARH	SIPLQTL	ESKLNTS	YIPDNGD
>4012-578_1	>2592-1135_1	>4038-202_1	>4460-259_1
FANPSLM	IMDLPA	RPEAYPL	TTYKSTH

	>4021-576_1 ARFVTNA >4024-575_1 MSQGNQA >4030-575_1 VFVLPSW >4031-574_1 EPQHQPM >4036-574_1 GLFTVTV >4039-573_1 GPLAFRW >4043-573_1 MLHSTVS >4052-572_1 CYMCRML >4058-571_1 WPDRLTY >4066-570_1 DGTYTHV >4079-567_1 DHRTVDN >4082-566_1 LLAKRIV >4087-565_1 KFIVKSM >4089-565_1 LAMQLSD	>2595-1134_1 FQNPTWF >2602-1128_1 CVFGYCL >2603-1126_1 LLQPLYI >2604-1126_1 ALRMGPM >2606-1125_1 VATKHDY >2607-1125_1 YPKWEGI >2611-1124_1 WYMPSWM >2617-1122_1 EIFPTWA >2619-1120_1 WSWPSWS >2624-1118_1 YDPWAQY >2628-1116_1 TITRSAH >2633-1115_1 STPRNWA >2637-1114_1 TYWPVPD >2639-1113_1 INPQWQE	>4044-202_1 VPSDEHM >4050-202_1 ETDKDSR >4052-202_1 WGLHNFA >4059-201_1 LQTHHWY >4060-201_1 HNRLHEN >4063-201_1 RVPMSDY >4074-200_1 TRFMQST >4077-200_1 SLTYPNH >4078-200_1 TWNSHWN >4085-200_1 HTYPHLL >4109-198_1 NAHNDKS >4117-198_1 MHVTLFA >4119-197_1 TGATIFK >4123-197_1 KPSPYPL	>4463-258_1 VITVVFG >4470-258_1 LDI*IAI >4472-258_1 SQSSNHL >4474-258_1 LTGTYSI >4481-257_1 QATVVVP >4483-257_1 MNWNLVV >4485-257_1 YWHPRWH >4498-256_1 YSEMHIE >4502-256_1 VHEHRQY >4507-256_1 DPRQQEP >4510-255_1 ALYDRSK >4519-255_1 TFWPDHP >4524-254_1 SSARPAT >4527-254_1 YDRDKGW
Number of peptides with matches to PAP	48	127	74	32