

Supplementary Information for

HLA and antibodies define scleroderma subtypes and risk in African and European Americans and suggest a role for molecular mimicry

Pravitt Gourh, Sarah A Safran, Theresa Alexander, Steven E Boyden, Nadia D Morgan, Ami A Shah, Maureen D Mayes, Ayo Doumateye, Amy R Bentley, Daniel Shriner, Robyn T Domsic, Thomas A Medsger Jr, Paula S Ramos, Richard M Silver, Virginia D Steen, John Varga, Vivien Hsu, Lesley Ann Saketkoo, Elena Schiopu, Dinesh Khanna, Jessica K Gordon, Brynn Kron, Lindsey A Criswell, Heather Gladue, Chris T Derk, Elana J Bernstein, S Louis Bridges Jr, Victoria K Shanmugam, Kathleen D Kolstad, Lorinda Chung, Suzanne Kafaja, Reem Jan, Marcin Trojanowski, Avram Goldberg, Benjamin D Korman, Peter J Steinbach, Settara C Chandrasekharappa, James C Mullikin, Adebowale Adeyemo, Charles Rotimi, Fredrick M Wigley, Daniel L Kastner, Francesco Boin, and Elaine F Remmers

Pravitt Gourh, MD E-mail: pravitt.gourh@nih.gov

Daniel L Kastner, MD PhD E-mail: kastnerd@mail.nih.gov

This PDF file includes:

Supplementary text Figs. S1 to S9 Tables S1 to S15 References for SI reference citations

Supporting Information Text

Supplemental Materials and Methods

African American cases and controls. AA patients with SSc were enrolled in the GRASP consortium from 23 academic centers in the US.

GRASP Centers

- National Human Genome Research Institute (NHGRI) & National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), National Institutes of Health (NIH)
- Johns Hopkins University
- University of Texas-Houston
- University of Pittsburgh
- Georgetown University
- Medical University of South Carolina
- Northwestern University
- University of Michigan
- Rutgers University
- Hospital for Special Surgery, Cornell University
- Tulane University
- Emory University
- University of California San Francisco
- University of Pennsylvania
- University of Alabama at Birmingham
- George Washington University
- Stanford University
- University of Chicago
- Columbia University
- New York University
- Boston University
- University of California Los Angeles
- Henry Ford Hospital, Detroit, Michigan
- University of Rochester

Only unrelated individuals were included as controls. Sera obtained from controls were tested for antinuclear antibodies (ANAs) by indirect immunofluorescence, and only those with a titer of <1:80 were included in this study. DNA was extracted from samples of whole blood or saliva.

Autoantibody Testing. Sera from the AA SSc patients were tested by a line immunoassay for systemic sclerosis profile autoantibodies (Euroimmun Euroline profile kit, Lübeck, Germany). For the European ancestry SSc patients, reported autoantibody data was extracted from dbGaP accession phs000357.v1.p1. Out of the 115 patients who were ATA+ only 1 was positive for ACA. More globally, 468 of the 723 tested sera (65%) were positive for a SSc-specific autoantibody (ACA, ATA or ARA) and only 8 of those 468 (1.7%) were positive for two autoantigens.

Genotyping. The AA SSc cases and controls were genotyped with the Illumina Infinium Multi-Ethnic Global Array kit. All included samples had call rate greater than 95%. Markers with call rate less than 95%, minor allele frequency less than 0.01, or controls that deviated from Hardy-Weinberg equilibrium with P less than 10⁻⁵ were excluded. High quality genotypes were imputed using the Michigan Imputation Server and the required 6,114 markers submitted to the HLA*IMP:03 server for HLA imputation. The European ancestry samples were genotyped on Illumina Human610-Quadv1_B chip and the same quality controls were executed separately on the genotypes from each study and then the samples were combined and only markers that were present in all the studies were retained. Samples were also checked for relatedness by an analysis of identity by descent of LD pruned ($r^2 < 0.5$) genotypes and no pairs were allowed with pi-hat greater than 0.085.

Principal Components Analysis (PCA). For the two ancestral populations, PCA was used to evaluate the genetic similarity of the cases with the controls, to remove outliers, and to correct for residual dissimilarity separately. LD-pruned data were used to calculate the top ten principal components for each sample. Two dimensional plots of the PCs allowed removal of outliers. Two dimensional plots of the first two PCs of the cases and the controls included in each study are shown in Supplementary Figure S1.

HLA Imputation.

African American cases and controls. We selected the HLA*IMP:03 tool to perform HLA imputation in the AA samples, because it has a multi-ethnic reference panel of 10,561 individuals that includes 568 individuals of African ancestry (1). We extracted genotypes of 19,038 markers that passed quality control from the greater MHC region (chromosome 6: 20,000,000 – 40,000,000 on GRCh37/hg19) and used the Michigan Imputation Server to perform haplotype phasing and SNP imputation using Eagle and the 1000G Phase 3 v5 reference (2). We then submitted the phased genotypes of the required 6,114 markers to the HLA*IMP:03 Server, which imputed two two-field (i.e., four-digit) classical HLA alleles of 11 HLA region genes for each individual. We used available whole exome sequence for 763 of the AA samples to determine their HLA alleles with a software tool that performs alignments to a population reference graph of the HLA region (HLA*PRG:LA) allowing comparison of the HLA*IMP:03 imputed alleles with the sequence based alleles (3).

European American cases and controls. We extracted genotypes of 2,690 SNPs that passed quality control in all the European ancestry datasets from the greater MHC region and used SNP2HLA and a mainly European ancestry reference of 5,225 individuals from the Type I Diabetes Consortium to impute classical HLA alleles and polymorphic amino acids of the HLA proteins (4).

HLA Association and Conditional Analysis. The HLA*IMP:03 imputed alleles for each gene were recoded as individual markers, each with two presence or absence alleles (PP, PA, or AA), matching the SNP2HLA imputation output. Alleles with frequency less than 0.01 were omitted. For *HLA* allele association, a logistic regression analysis was performed under a dominant model, encoding AA genotypes=0 and AP and PP genotypes=1. A dominant model was applied for analysis of *HLA* alleles because it is consistent with the method classically used for identifying *HLA* alleles associated with different human diseases (5). It is also consistent with the importance of presence or absence of a molecule endowing capability to bind and present specific peptide antigens. Regressions were corrected for genetic dissimilarity between the cases and the controls by including the top 10 PCs as covariates. To account for strong LD in the region, independent associations were identified by performing conditional analysis and at each level the strongest association accepted, and the scan rerun with all the previous 'accepted' associations included as covariates. This was done until p-value statistical significance threshold was reached. The total number of classical *HLA* alleles tested for association were 138 in both the populations and there were 5 analysis conducted. Thus a Bonferroni's multiple test corrected significance threshold of P < 0.000072 was used for association analysis.

Amino Acid Analysis. For the HLA*IMP:03 imputed data, the genotypes of all the polymorphic amino acids of the HLA proteins were extracted from the IPD-IMGT/HLA Database according to the two alleles specified for each imputed gene and each polymorphic amino acid marker was coded for absence or presence genotypes, AA, AP, or PP (6). The SNP2HLA output included similarly encoded P/A genotypes for each of the polymorphic amino acid. Amino acid associations with SSc were evaluated with a dominant model logistic regression analysis as done for the *HLA* alleles. Amino acids with frequency less than 0.01 were omitted. To account for strong LD in the region, independent associations were identified by performing conditional analysis and at each level the strongest association accepted, and the scan rerun with all the previous 'accepted' associations included as covariates. This was done until p-value statistical significance threshold was reached. The p-value threshold was set as P < 0.000013 based on 800 amino acids tested across both population samples, multiplied by 5 sets of analysis.

HLA Alpha-Beta Chain Pair Analysis. The non-polymorphic HLA-DR alpha chain combines with HLA-DRB1 alleles, thus disease-associated DR1 alpha/beta combinations are defined by the association of DRB1 alleles. However, HLA-DQ1 and HLA-DP1 molecules have polymorphic alpha and beta chains, therefore the association of DQA1/DQB1 and DPA1/DPB1 allele pairs are relevant to disease association. To identify SSc-associated HLA-DQ and -DP alpha and beta chain pairs we performed haplotype association analysis of the *HLA-DQ* and -*DP* genotypes using Haploview and then created new P/A markers for the haplotypes that were analyzed by logistic regression (7).

3D Protein Modeling. Protein Data Bank entries 1a35 for topoisomerase I, 2ipx for fibrillarin, 3nqu for centromere protein A, 6atf for HLA-DR, 1s9v for HLA-DQ, and 3lqz for HLA-DP were obtained and UCSF chimera was used to model the individual amino acid positions. Molecular graphics and analyses performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311 (8).

CART Analysis. Classification and Regression Tree (CART) analysis was performed to explore higher order interactions amongst the HLA alleles (CART 6.0, Salford Systems, San Diego, CA) (9). *HLA* alleles were used as nominal categorical variables to predict the outcome SSc. CART analysis interactively identifies the most powerful variable classifier for SSc and splits the variables into binary trees (presence/absence). The SSc-specific autoantibody subsets of SSc for both the AA and EA populations were analyzed individually and split into successive groups using ten-fold cross-validation and trees were pruned for significance at p-value (P)<0.001 using Pearson's chi-square test or, when appropriate, Fisher's exact test.

Identification of Immunodominant Peptides. The HLA-II peptide groove is open at both ends, thus it can allow peptide binding of varying lengths (10). Predominantly, 13-25 residues long peptides have been found to bind HLA-II in nature with the most abundant having an average peptide length of 15 residues (11–16). Several previous studies have used 15-mer residues as the default peptide length for predicting peptides that will bind to HLA-II molecule (17–22). The default peptide length suggested for the NetMHCIIpan 3.2 server is a 15-mer peptide (23, 24). The NetMHCIIpan 3.2 server was used to predict the binding of 15-mer peptide sequences within the protein of interest (topoisomerase I, fibrillarin, or CENPA/CENPB) to the SSc-associated *MHC* Class II α/β heterodimers within the respective autoantibody subsets. Peptides with a binding affinity \leq 500nM and observed in two of the SSc-associated *HLA* α/β heterodimers were prioritized as immunodominant peptides binders were analyzed for overlap in peptide sequences between the different *HLA* alleles, and peptide sequences were selected if they had moderate binding affinity for both the associated alleles. The prediction values are given as IC50 values in nanomolar (nM). Strong binders are defined as peptides with a binding affinity less than or equal to 500 nM (23, 25–27).

Molecular Mimicry. The prioritized immunodominant peptide sequences were entered into the NCBI Basic Local Alignment Search Tool Standard Protein BLAST with the organism set to virus (taxid:10239) and human (taxid:9606) to identify homologous sequences in viral and human sequence databases, respectively. Significant homology was defined by an E-value < 0.05 (28). To rule out homology by chance, one hundred 15-mer random sequence peptides were generated using ExPASy from the Swiss Institute of Bioinformatics (SIB) and Sequence Manipulation Suite and compared for homologous sequences in viral sequence database (29, 30). We used an arbitrarily selected 15-mer peptide from serum albumin comprising residues 152-166 as an additional comparator for self-antigen homology (31).

Web Addresses.

- dbGaP: https://www.ncbi.nlm.nih.gov/gap
- HLA-PRG-LA: https://github.com/DiltheyLab/HLA-PRG-LA
- Michigan Imputation Server: https://imputationserver.sph.umich.edu/index.html
- HLA*IMP:03 Server: http://imp.science.unimelb.edu.au/hla/
- SNP2HLA: http://software.broadinstitute.org/mpg/snp2hla/
- The IPD-IMGT/HLA Database: https://www.ebi.ac.uk/ipd/imgt/hla/
- · Haploview: https://www.broadinstitute.org/haploview/haploview
- NetMHCIIpan: http://www.cbs.dtu.dk/services/NetMHCIIpan/
- NCBI Protein Blast: https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome
- UCSF Chimera: https://www.cgl.ucsf.edu/chimera/
- Random Peptide generators: https://web.expasy.org/randseq/, http://www.bioinformatics.org/sms2/random_protein.html





662 SSc cases (blue), 946 ctrls (grey)







723 SSc cases (blue), 5,437 controls (grey)

Eigenvector 1 vs 2



Fig. S1. Principal component analysis (PCA) plots of the African and European American SSc cases and controls included in the study. PCA was performed on LD-pruned data ($r^2 < 0.5$).



Fig. S2. Association of classical HLA alleles with SSc and autoantibody subsets of SSc in African and European Americans. HLA imputed classical allele associations in SSc and autoantibody subsets determined in African Americans (left panels) and European Americans (right panels). Common AA haplotype: HLA-DRB1*08:04/DQA1*05:01/DQB1*03:01; Common EA haplotypes: HLA-DRB1*11:04/DQA1*05:01/DQB1*03:01 and HLA-DRB1*07:01/DQA1*02:01/



Conditioned Association plot of SNPs and HLA alleles on final HLA model



Fig. S3. Association Analysis in African Americans of the HLA region SNPs and Classical HLA alleles, both Unconditioned and Conditioned on the final HLA model, HLA-DRB1*08:04 and HLA-DRB1*11:02 alleles. Both analyses corrected for the top 10 Principal Components.



Conditioned Association plot of SNPs and HLA alleles on final HLA model



Fig. S4. Association Analysis in European Americans of the HLA region SNPs and Classical HLA alleles, both Unconditioned and Conditioned on the final HLA model, HLA-DQB1*02:02, HLA-DPB1*13:01 and HLA-DRB1*11:04 alleles. Both analyses corrected for the top 10 Principal Components.



Fig. S5. Population frequency of HLA-DPB1*13:01 allele and SSc prevalence with outlier removed (see Figure 1).



129 African American AFA-positive cases, 946 controls

Fig. S6. (A) Logistic regression and conditional analysis of HLA polymorphic amino acids within autoantibody subsets of SSc in African Americans in the AFA+ SSc subset (top) and ATA+ SSc subset (bottom). (SSc=systemic sclerosis; AFA=anti-fibrillarin antibody; ATA=anti-topoisomerase I antibody)



Fig. S6. (B) Logistic regression and conditional analysis of HLA polymorphic amino acids within autoantibody subsets of SSc in European Americans in the ACA+ SSc subset (left) and ATA+ SSc subset (right). (SSc=systemic sclerosis; ACA=anti-centromere antibody; ATA=anti-topoisomerase I antibody)



Fig. S7. Ribbon models of the HLA-DR and HLA-DQ proteins with independently associated amino acid residues, based on Protein Data Bank entries 6atf and 1s9v, respectively. (A) 3D model of HLA-DR in three different orientations in order to visualize peptide binding grove and associated amino acid positions in African Americans; (B) 3D model of HLA-DQ in three different orientations in order to visualize peptide binding grove and associated amino acid positions in African Americans; (C) 3D model of HLA-DR in three different orientations in order to visualize peptide binding grove and associated amino acid positions in African Americans; (C) 3D model of HLA-DR in three different orientations in order to visualize peptide binding grove and associated amino acid positions in European Americans; (D) 3D model of HLA-DQ in three different orientations in order to visualize peptide binding grove and associated amino acid positions in European Americans; (D) 3D model of HLA-DQ in three different orientations in order to visualize peptide binding grove and associated amino acid positions in European Americans; (D) 3D model of HLA-DQ in three different orientations in order to visualize peptide binding grove and associated amino acid positions in European Americans.









Sample Collection	dbGaP Accession Number	# Samples Included
Genome-Wide Association Study in Systemic Sclerosis, cases	phs000357.v1.p1	723
Cohort-Based Genome-Wide Association Study of Glioma (GliomaScan), controls	phs000652.v1.p1	29
GWAS for Genetic Determinants of Bone Fragility, controls	phs000138.v2.p1	714
Cancer Genetic Markers of Susceptibility for Bladder Cancer (CGEMS Bladder), controls	phs000346.v2.p2	1,354
The Breast and Prostate Cancer Cohort Consortium (BPC3), controls	phs000812.v1.p1	2,411
PGRN-RIKEN: Genome-Wide Association Study of Drug-Induced Long-QT Syndrome, controls	phs000331.v1.p1	169
NIA-Late Onset Alzheimer's Disease and National Cell Repository for Alzheimer's Disease Family Study, controls	phs000168.v2.p2	760

Table S1. European American SSc and control datasets obtained from dbGaP.

Table S2. S	Sample	demographics.
-------------	--------	---------------

	African Americans	European Americans
Controls		
n	946	5,437
Female, number (%)	502 (53.1)	3333 (61.3)
Male, number (%)	444 (46.9)	2103 (38.7)
SSc patients		
n	662*	723
Female, number (%)	510 (77)	643 (88.9)
Male, number (%)	150 (22.7)	80 (11.1)
Antibodies		
Anti-topoisomerase I, number (%)	183 (27.6)	115 (15.9)
Anti-fibrillarin, number (%)	129 (19.5)	Not Measured
Anti-centromere, number (%)	64 (9.7)	238 (32.9)
Anti-RNA polymerase III, number (%)	119 (18.0)	123 (17.0)

*Two samples had missing gender data.

Table S3. H	ILA*IMP:03-imputed	allele concordance with	sequence-based types of	of 763 African	Americans determined v	with HLA-PRG-LA.

HLA Locus	# Types with Frequency > 1%*	Concordance Rate
HLA-A	19	98.58%
HLA-B	24	98.62%
HLA-C	15	97.52%
HLA-DRB1	21	99.07%
HLA-DQA1	7	98.13%
HLA-DQB1	12	97.93%
HLA-DPA1	3	95.96%
HLA-DPB1	11	99.03%
	*In	the HLA-PRG-LA data.

Allolo	Frequency Cases (%)	Frequency Controls (%)
Allele	n=662	n=946
HLA-A*01:01	4.0	4.5
HLA-A*02:01	13.0	12.0
HLA-A*02:02	3.4	3.4
HLA-A*02:05	2.2	2.3
HLA-A*03:01	8.5	7.4
HLA-A*11:01	1.1	1.3
HLA-A*23:01	12.1	10.4
HLA-A*24:02	1.6	2.6
HLA-A*26:01	1.4	1.4
HLA-A*29:02	2.8	3.7
HLA-A*30:01	6.8	6.6
HLA-A*30:02	5.5	6.1
HLA-A*32:01	1.0	1.5
HLA-A*33:01	3.0	2.6
HLA-A*33:03	5.9	5.5
HLA-A*34:02	3.2	3.5
HLA-A*36:01	2.6	2.4
HLA-A*66:01	1.4	1.7
HLA-A*68:01	4.3	4.6
HLA-A*74:01	4.4	5.3
HLA-B*07:02	6.5	7.9
HLA-B*08:01	4.9	4.2
HLA-B*14:02	2.0	2.1
HLA-B*15:01	0.8	1.6
HLA-B*15:03	5.7	5.3
HLA-B*15:10	3.0	3.4
HLA-B*15:16	2.0	1.3
HLA-B*18:01	2.9	3.4
HLA-B*35:01	7.2	7.5
HLA-B*40:01	1.1	1.0
HLA-B*42:01	5.3	5.8
HLA- $B*44:02$	1.4	1.3
HLA-B*44:03	5.0	5.5
HLA- $B*45:01$	4.6	4.7
HLA-B*49:01	3.0	3.0
HLA-B*51:01	3.9	2.6
HLA-B*53:01	15.0	11.7
HLA-B*57:01	1.1	1.2
HLA-B*57:03	2.1	3.5
HLA-B*58:01	4.7	4.3
HLA-B*58:02	4.3	4.2
HLA-B*81:01	1.1	1.7

Table S4.	HLA allele	frequencies	in African	Americans.

(continued on next page)

Allele	Frequency Cases (%) n=662	Frequency Controls (%) n=946
HLA-C*02:02	6.0	6.2
HLA-C*02:10	1.7	1.2
HLA-C*03:02	2.2	1.8
HLA-C*03:03	0.9	1.4
HLA-C*03:04	5.6	4.9
HLA-C*04:01	23.0	20.3
HLA-C*05:01	2.7	2.6
HLA-C*06:02	8.8	9.4
HLA-C*07:01	11.9	13.1
HLA-C*07:02	6.8	8.1
HLA-C*08:02	3.1	3.8
HLA-C*12:03	1.0	1.8
HLA-C*14:02	2.3	1.8
HLA-C*15:05	0.8	1.0
HLA-C*16:01	10.5	8.6
HLA-C*17:01	6.9	7.1
HLA-C*18:01	2.3	3.2
HLA-DRR1*01.01	2.0	9.1
HLA-DRR1*01.00	2.5	3 9
HLA-DRR1*03.01	4.8	7 7
$HLA_DRB1*0/.01$	1 1	1.7
$HLA_DRB1*0/.05$	0.8	2.0
HLA_DRB1*07.01	5.8	10.7
$HLA_DRB1*08.01$	12 7	4.8
$HLA_DRB1*00.04$	25	2.5
$HLA_DRB1*10.01$	1.0	1.6
$HLA_DRB1*11.01$	9.4	7.9
HLA-DRB1*11.02	7.3	3.8
HLA_DRB1*12.01	2.5	3.8
HLA-DRB1*13.01	3.8	4 4
HLA-DRB1*13:02	8.8	7.1
HLA-DRB1*13:03	2.9	3.5
HLA-DRB1*13:04	3.1	1.3
HLA-DRB1*14:01	1.1	2.3
HLA-DRB1*15:01	3.0	3.2
HLA-DRB1*15:03	13.1	12.6
HLA-DRB1*16:02	1.4	1.4
HLA-DQA1*01:01	9.9	13.3
HLA-DQA1*01:02	33.9	31.1
HLA-DQA1*01:03	1.9	2.4
HLA-DQA1*02:01	5.9	10.7
HLA-DQA1*03:01	7.9	10.5
HLA-DQA1*04:01	12.2	10.1
HLA-DQA1*05:01	28.0	21.5
HLA-DQB1*02:01	9.1	13.5
HLA-DQB1*02:02	4.2	7.5
HLA-DQB1*03:01	20.5	13.6
HLA-DQB1*03:02	2.3	3.3
HLA-DOB1*03:03	2.0	2.5
HLA-DOB1*03:19	10.1	4.5
HLA-DQB1*04:02	6.3	7.7
HLA-DOB1*05:01	12.4	14.5
HLA-DOB1*05:03	0.9	1.9
HLA-DOB1*06:02	20.5	20.2
HLA-DOB1*06:03	1.2	1.4
HLA DOB1*06.07	21	2.0
$\Pi \Box \Lambda^{-} D Q D I 0 0.04$	2.1	-

Table S4. HLA allele frequencies in African Americans (continued).

(continued on next page)

Allolo Frequency Cases (%)		Frequency Controls (%)
Anele	n=662	n=946
HLA-DPA1*01:03	32.6	39.1
HLA-DPA1*02:01	40.0	31.0
HLA-DPA1*02:02	15.3	16.3
HLA-DPA1*03:01	11.6	12.4
HLA-DPB1*01:01	32.2	28.5
HLA-DPB1*02:01	12.7	13.2
HLA-DPB1*03:01	5.3	5.3
HLA-DPB1*04:01	8.4	10.9
HLA-DPB1*04:02	11.3	12.7
HLA-DPB1*11:01	1.7	3.1
HLA-DPB1*13:01	9.0	5.1
HLA-DPB1*17:01	8.9	8.1
HLA-DPB1*18:01	4.2	7.0
HLA-DPB1*40:01	1.9	1.2
HLA-DPB1*85:01	1.1	1.5

Table S4. HLA allele frequencies in African Americans (continued).

Allolo	Frequency Cases (%)	Frequency Controls (%)
Allele	n=723	n=5437
HLA-A*01:01	16.9	16.6
HLA-A*02:01	28.7	28.5
HLA-A*03:01	14.3	14.9
HLA-A*11:01	6.3	5.9
HLA-A *23:01	1.7	1.9
HLA-A*24:02	9.5	8.1
HLA-A*25:01	2.0	2.3
HLA-A*26:01	3.3	2.7
HLA-A*29:02	2.1	3.5
HLA-A*30:01	0.9	1.1
HLA-A*31:01	2.6	2.6
HLA-A*32:01	3.2	3.7
HLA-A*68:01	2.8	3.3
HLA-B*07:02	12.3	14.5
HLA-B*08:01	14.5	12.2
HLA-B*13:02	1.2	2.1
HLA-B*14:02	2.9	2.4
HLA-B*15:01	6.2	6.9
HLA-B*18:01	5.5	4.0
HLA-B*27:05	4.1	4.1
HLA-B*35:01	7.8	5.7
HLA-B*35:02	1.4	0.5
HLA-B*35:03	1.2	1.5
HLA-B*37:01	1.2	1.5
HLA-B*38:01	1.4	1.6
HLA-B*39:01	1.7	1.3
HLA-B*40:01	5.5	5.7
HLA-B*40:02	1.5	1.1
HLA-B*44:02	9.5	9.7
HLA-B*44:03	2.6	4.9
HLA-B*49:01	0.9	1.4
HLA-B*51:01	5.4	4.7
HLA-B*55:01	1.7	1.8
HLA-B*57:01	2.6	3.4
HLA-C*01:02	3.6	3.6
HLA-C*02:02	4.7	4.4
HLA-C*03:03	5.2	5.8
HLA-C*03:04	8.0	8.2
HLA-C*04:01	12.1	10.2
HLA-C*05:01	9.0	9.6
HLA-C*06:02	6.4	8.7
HLA-C*07:01	18.7	16.2
HLA-C*07:02	13.7	15.5
HLA-C*07:04	1.9	1.5
HLA-C*08:02	3.4	3.5
HLA-C*12:03	4.9	4.3
HLA-C*14:02	1.5	1.3
HLA-C*15:05	2.1	2.0
HLA-C*16:01	1.9	3.5

(continued on next page)

	Frequency Cases (%)	Frequency Controls (%)
Allele	n=723	n=5437
HLA-DRB1*01:01	12.7	9.1
HLA-DRB1*01:02	1.1	0.9
HLA-DRB1*01:03	0.8	1.4
HLA-DRB1*03:01	14.0	12.1
HLA-DRB1*04:01	9.1	10.0
HLA-DRB1*04:04	5.1	4.1
HLA-DRB1*07:01	8.0	12.7
HLA-DRB1*08:01	3.4	2.3
HLA-DRB1*09:01	0.6	1.0
HLA-DRB1*11:01	5.5	5.2
HLA-DRB1*11:04	5.3	2.4
HLA-DRB1*12:01	0.8	1.5
HLA-DRB1*13:01	4.6	6.3
HLA-DRB1*13:02	4.4	4.3
HLA-DRB1*14:01	1.1	2.5
HLA-DRB1*15:01	12.2	14.6
HLA-DRB1*16:01	1.7	1.4
HLA-DQA1*01:01	17.4	14.5
HLA-DQA1*01:02	19.3	21.3
HLA-DQA1*01:03	4.9	6.6
HLA-DQA1*02:01	8.4	13.1
HLA-DQA1*03:01	17.8	17.8
HLA-DQA1*04:01	4.2	2.7
HLA-DQA1*05:01	27.7	23.6
HLA-DQB1*02:01	14.2	12.6
HLA-DQB1*02:02	5.3	9.5
HLA-DQB1*03:01	20.5	18.0
HLA-DQB1*03:02	10.6	10.2
HLA-DQB1*03:03	3.7	4.6
HLA-DQB1*04:02	3.9	2.7
HLA-DQB1*05:01	14.9	11.9
HLA-DQB1*05:02	2.6	1.5
HLA-DQB1*05:03	2.1	2.7
HLA-DQB1*06:02	12.4	14.8
HLA-DQB1*06:03	4.5	6.4
HLA-DQB1*06:04	3.7	3.6
$HLA-DPA1^01:03$	80.4	82.1
HLA-DPA1*02:01	16.7	14.1
HLA-DPA1*02:02	2.6	3.1
HLA-DPB1*01:01	6.4 10. <i>C</i>	5.4
HLA-DPB1*02:01	10.6	12.0
$HLA-DPB1^{*}03:01$	12.4	11.1
$\Pi LA - DP B1^* 04:01$	41.ð 19.9	45.8 10.2
ПLA-DFB1*04:02 ПГА ПДД1*05.01	12.2	12.3
11LA-DFB1'00:01 ULA DDD1*02.01	2.U 9.1	1.9
ПLA-DFB1'00:01 НГА ПДД1*10.01	2.1 2.2	2.1
ПLA-DF D1 ⁻ 10:01 НГ А ПОD1*11.01	2.0 1 1	1. <i>1</i> 9.9
ПLA-DF D1 *11;01 НГА DDD1*10.01	1.1	2.2 1 7
ПLA-DF D1 13:01 НГА ПОD1*11.01	4. <i>2</i> 1.9	1.1
ПБА-DF D1 14:01 ПГА ПОР1*17.01	1.0	1.0
11LA-DF D1 11:01	1.0	1.2

Table S5. HLA allele frequencies in European Americans (continued).

Table S6. Scleroderma Prevalence and HLA-DPB1*13:01 Allele Frequency in Different Populations Around the World.

Population	SSc Prevalence (per million)	SSc Prevalence Reference	HLA-DPB1*13:01 Allele Frequency (%)	HLA-DPB1*13:01 Allele Frequency Reference
Buenos Aires, Argentina	296	(32)	2.8	http://www.allelefrequencies.net/pop6001c.asp?pop_id=1310
Tokyo, Japan	38	(33)	1.79	http://www.allelefrequencies.net/pop6001c.asp?pop_id=3279
US-Caucasian	138	(34)	2.4	http://www.allelefrequencies.net/pop6001c.asp?pop_id=2780
UK	88	(35)	2.1	http://www.allelefrequencies.net/pop6001c.asp?pop_id=2511
France	158	(36)	2.73	http://www.allelefrequencies.net/pop6001c.asp?pop_id=2517
US-African American	300	(37)	6	http://www.allelefrequencies.net/pop6001c.asp?pop_id=2779
New Zealand	30	(38)	0.3	http://www.allelefrequencies.net/pop6001c.asp?pop_id=2160
Greece	154	(39)	3	http://www.allelefrequencies.net/pop6001c.asp?pop_id=1759
Choctaw	4690	(40)	22	(41)

rabio off fizit anolo accordationo in autoantibobaly poolato ana nogativo cabooto in fantoan fantonoano

		Freq %	OR (95% CI)	P-value
		(SSc/Ctrls)	UnCondit	tioned
	HLA-DRB1*0804	24.3 / 9.3	3.2 (2.4-4.3)	3.26×10 ⁻¹⁶
	HLA-DQB1*0319	18.4 / 8.8	2.4 (1.8-3.2)	2.45×10 ⁻⁸
	HLA-DQB1*0301	37 / 25.8	1.8 (1.4-2.2)	1.41×10 ⁻⁶
All SSc v controls	HLA-DRB1*0701	11.5 / 20	0.6 (0.4-0.7)	2.72×10 ⁻⁶
SSc=662; Control=946	HLA-DQA1*0201	11.5 / 20	0.6 (0.4-0.7)	3.18×10 ⁻⁶
	HLA-DRB1*1102	13.6 / 7.1	2.2 (1.6-3)	9.39×10 ⁻⁶
	HLA-DPA1*0201	62.1 / 51.5	1.6 (1.3-1.9)	3.20×10 ⁻⁵
	HLA-DPB1*1301	16.9 / 9.7	1.9 (1.4-2.6)	3.21×10 ⁻⁵
	HLA-DRB1*0804	42.6 / 9.3	7.4 (4.9-11.3)	2.61×10 ⁻¹⁹
	HLA-DQB1*0609	20.9 / 6.6	3.8 (2.3-6.3)	1.37×10 ⁻⁶
AFA+ SSC V CONTROLS	HLA-DQB1*0301	45 / 25.8	2.5 (1.7-3.6)	9.16×10 ⁻⁶
55C=129; Control=946	HLA-DQB1*0319	22.5 / 8.8	3 (1.9-4.9)	2.53×10 ⁻⁵
	HLA-DRB1*1302	29.5 / 13.8	2.6 (1.7-4)	3.70×10 ⁻⁵
		100/02		4.05.40-8
AFA-SSC V controls	HLA-DRB1*0804	19.9/9.3	2.5 (1.8-3.3)	1.05×10°
SSC=533; Control=946	HLA-DQB1*0609	5.6 / 6.6	0.9 (0.6-1.4)	0.62
	HLA-DPB1*1301	30.6 / 9.7	4.3 (2.9-6.3)	2.35×10 ⁻¹²
	HLA-DQB1*0201	7.7 / 25.1	0.3 (0.2-0.5)	1.10×10 ⁻⁸
474.00	HLA-DRB1*0301	3.3 / 14.9	0.2 (0.1-0.5)	4.92×10 ⁻⁷
ATA+ SSC V controls	HLA-DQB1*0301	43.7 / 25.8	2.3 (1.7-3.1)	2.64×10⁻ ⁶
SSC=183; Control=946	HLA-DPA1*0201	69.9 / 51.5	2.3 (1.6-3.2)	2.73×10⁻ ⁶
	HLA-DRB1*0804	21.3 / 9.3	2.8 (1.8-4.2)	8.13×10 ⁻⁶
	HLA-DQA1*0501	51.4 / 39.1	1.6 (1.2-2.2)	2.90×10 ⁻³
ATA- SSc v controls	HLA-DPB1*1301	11.7 / 9.7	1.2 (0.9-1.7)	0.27
SSc=479; Control=946	HLA-DQB1*0201	22.1 / 25.1	0.8 (0.7-1.1)	0.21

* Frequency of individuals with one or two alleles. Common AA haplotype: HLA-DRB1*08:04/DQA1*05:01/DQB1*03:01.

		Freq %	OR (95% CI)	P-value
		(SSc/Ctrls)	UnConc	litioned
	HLA-DQB1*0202	10.2 / 18	0.5 (0.4-0.6)	3.55×10⁻ ⁹
	HLA-DRB1*0701	15.1 / 23.7	0.5 (0.4-0.7)	6.06×10 ⁻⁹
	HLA-DQA1*0201	15.8 / 24.3	0.6 (0.4-0.7)	1.04×10 ⁻⁸
All SSC V controls	HLA-DPB1*1301	8.3 / 3.3	2.6 (1.9-3.5)	1.75×10 ⁻⁸
55C-725, CONTION-5457	HLA-DRB1*1104	10.5 / 4.7	2.2 (1.7-2.9)	9.25×10⁻ ⁸
	HLA-B*4403	5.1/9.7	0.5 (0.3-0.7)	4.72×10 ⁻⁶
	HLA-DRB1*0101	23.7 / 17.3	1.5 (1.2-1.8)	4.39×10 ⁻⁵
	HLA-DPB1*1301	32.2 / 3.3	13.7 (8.9-21)	1.47×10 ⁻²⁴
ATA+ SSc v controls	HLA-DRB1*1104	25.2 / 4.7	6.3 (3.9-10)	8.62×10 ⁻¹²
SSc=115; Control=5437	HLA-DPA1*0201	48.7 / 26.3	2.9 (2-4.2)	8.70×10⁻ ⁸
	HLA-DPA1*0103	54.8 / 32.6	2.7 (1.8-3.9)	4.65×10 ⁻⁷
ATA- SSc v controls	HLA-DPB1*1301	3.8 / 3.3	1.1 (0.7-1.7)	0.65
SSc=608; Control=5437	HLA-DRB1*1104	7.7 / 4.7	1.6 (1.1-2.2)	8.81×10 ⁻³
	HLA-DRB1*0701	3.4 / 23.7	0.1 (0.05-0.2)	4.79×10 ⁻²⁰
	HLA-DQA1*0201	4.6 / 14.5	0.1 (0.1-0.2)	4.85×10 ⁻¹⁸
	HLA-DQB1*0202	2.9 / 18	0.1 (0.1-0.3)	2.44×10 ⁻¹⁴
	HLA-DQB1*0501	42.4 / 22.3	2.3 (1.8-3)	4.21×10 ⁻⁹
ACA+ SSC V CONTROIS	HLA-DQA1*0101	47.5 / 26.7	2.2 (1.7-2.9)	7.08×10 ⁻⁹
55C-256, CUILI 0I-5457	HLA-DRB1*0101	34.5 / 17.3	2.2 (1.7-3)	1.32×10 ⁻⁷
	HLA-DQA1*0401	14.3 / 5.4	2.7 (1.8-4)	4.18×10 ⁻⁵
	HLA-DQB1*0303	2.1/8.9	0.2 (0.1-0.5)	6.29×10⁻ ⁶
	HLA-DRB1*0801	11.8 / 4.6	2.6 (1.7-3.9)	4.97×10 ⁻⁵
	HLA-DRB1*0701	20.8 / 23.7	0.8 (0.7-1.1)	0.13
ACA- SSC v controls	HLA-DQB1*0501	19.4 / 22.3	0.9 (0.7-1.1)	0.28
550-465; CUILIUI-5457	HLA-DQA1*0401	5.2 / 5.4	1.0 (0.6-1.5)	0.95

* Frequency of individuals with one or two alleles.

Common EA haplotypes: HLA-DRB1*11:04/DQA1*05:01/DQB1*03:01.

Common EA haplotypes: HLA-DRB1*07:01/DQA1*02:01/DQB1*02:02.

		A	ATA+ SSc Subset	
Ancestry		Peptide	Affinit	ty (nM)
	Position	Sequence	DPA1*02:01/DPB1*13:01	DRA1*01:01/DRB1*11:04
	420	IQGSIKYIMLNPSSR	498.1	132.1
	421	QGSIKYIMLNPSSRI	357.6	81.7
	422	GSIKYIMLNPSSRIK	374.8	59.9
	423	SIKYIMLNPSSRIKG	448.4	58.4
	468	KEMKVRQRAVALYFI	492.6	94.0
	469	EMKVRQRAVALYFID	488.2	169.7
	470	MKVRQRAVALYFIDK	397.4	290.8
	473	RQRAVALYFIDKLAL	264.2	340.3
EA	474	QRAVALYFIDKLALR	320.0	195.9
	475	RAVALYFIDKLALRA	340.4	131.0
	476	AVALYFIDKLALRAG	481.1	117.5
	541	VPVEKRVFKNLQLFM	400.4	391.4
	542	PVEKRVFKNLQLFME	224.8	361.4
	543	VEKRVFKNLQLFMEN	187.4	328.0
	544	EKRVFKNLQLFMENK	187.4	306.2
	545	KRVFKNLQLFMENKQ	246.2	311.3
	546	RVFKNLQLFMENKQP	408.9	456.4
	Position	Sequence	DPA1*02:01/DPB1*13:01	DQA1*05:01/DQB1*02:01
۵۵	472	VRQRAVALYFIDKLA	268.2	489.2
~~~	473	RQRAVALYFIDKLAL	264.2	446.4

Table S9. Predicted immunodominant peptides in topoisomerase I protein; peptide in red shown in the manuscript in Figure 5.

			AFA+ SSc Subs	set	in tight LD
Ancestry		Peptide		Affinity (nM)	
	Position	Sequence	DRA1*01:01/DRB1*08:04	DQA1*01:02/DQB1*06:09	DRA1*01:01/DRB1*13:02
	130	DKIEYRAWNPFRSKL	491.0	15195.4	455.9
	131	KIEYRAWNPFRSKLA	296.5	13720.9	377.2
	132	IEYRAWNPFRSKLAA	276.3	14399.6	539.5
	133	EYRAWNPFRSKLAAA	270.5	14939.3	810.2
	134	YRAWNPFRSKLAAAI	230.8	11841.2	359.9
	135	RAWNPFRSKLAAAIL	235.1	10171.3	226.2
	136	AWNPFRSKLAAAILG	273.8	9922.7	270.4
	137	WNPFRSKLAAAILGG	315.4	9407.5	332.6
	138	NPFRSKLAAAILGGV	381.7	7104.8	390.8
	151	GVDQIHIKPGAKVLY	323.8	19226.5	90.4
	152	VDQIHIKPGAKVLYL	249.2	17250.9	69.1
	153	DQIHIKPGAKVLYLG	270.0	18134.6	86.8
	154	QIHIKPGAKVLYLGA	245.7	16850.1	106.1
	155	IHIKPGAKVLYLGAA	404.1	15835.1	285.2
	184	DGLVYAVEFSHRSGR	465.1	14957.9	1244.2
	185	GLVYAVEFSHRSGRD	435.2	18178.8	1278.0
	186	LVYAVEFSHRSGRDL	466.2	21658.7	1284.4
	194	HRSGRDLINLAKKRT	385.0	22684.5	1421.1
	195	RSGRDLINLAKKRTN	181.6	23578.4	1106.2
	196	SGRDLINLAKKRTNI	108.3	24642.8	689.9
	197	GRDLINLAKKRTNII	81.2	24260.5	131.0
AA	198	RDLINLAKKRTNIIP	88.1	24504.6	119.5
	199	DLINLAKKRTNIIPV	116.0	21083.4	83.3
	200	LINLAKKRTNIIPVI	124.4	14049.4	50.2
	201	INLAKKRTNIIPVIE	361.8	12731.9	55.4
	217	ARHPHKYRMLIAMVD	438.2	10118.4	725.3
	218	RHPHKYRMLIAMVDV	342.1	7057.2	520.4
	219	HPHKYRMLIAMVDVI	322.0	3925.2	236.0
	220	PHKYRMLIAMVDVIF	312.9	3051.9	215.8
	221	HKYRMLIAMVDVIFA	258.1	2138.2	167.5
	222	KYRMLIAMVDVIFAD	303.3	2100.7	197.0
	239	QPDQTRIVALNAHTF	485.1	9365.3	53.4
	240	PDQTRIVALNAHTFL	294.7	7407.5	23.5
	241	DQTRIVALNAHTFLR	202.7	7323.3	18.1
	242	QTRIVALNAHTFLRN	169.6	7269.3	17.6
	243	TRIVALNAHTFLRNG	197.6	8079.4	20.3
	244	RIVALNAHTFLRNGG	265.1	9136.4	25.5
	255	RNGGHFVISIKANCI	269.0	8703.3	35.4
	256	NGGHFVISIKANCID	280.2	8921.2	33.4
	257	GGHFVISIKANCIDS	243.7	8594.8	30.1
	258	GHFVISIKANCIDST	268.2	7873.6	28.8
	259	HFVISIKANCIDSTA	370.3	7776.0	31.6
	306	HAVV//GV/XBDDDK//K	166.8	20/3/ 6	1590.8

Table S10. Predicted immunodominant peptides in fibrillarin protein; peptide in red shown in the manuscript in Figure 5. (LD=linkage disequilibrium)

Table S11. (A) Predicted immunodominant peptides in Centromere Protein A; peptide in red shown in the manuscript in Figure 5. (LD=linkage disequilibrium)

			ACA+ SSc Subset (C	ENPA)	in tight ID
Ancestry		Peptide		Affinity (nM)	/
	Position	Sequence	DRA1*01:01/DRB1*07:01	DQA1*01:01/DQB1*05:01	DRA1*01:01/DRB1*01:01
	48	LKEIRKLQKSTHLLI	93.0	8523.5	68.4
	49	KEIRKLQKSTHLLIR	77.1	7917.4	57.1
	50	EIRKLQKSTHLLIRK	78.0	7992.8	60.5
	51	IRKLQKSTHLLIRKL	68.5	6146.3	56.1
	52	RKLQKSTHLLIRKLP	86.8	6388.9	77.2
	53	KLQKSTHLLIRKLPF	127.2	4942.0	105.9
	54	LQKSTHLLIRKLPFS	237.8	4827.6	142.8
	55	QKSTHLLIRKLPFSR	467.8	5302.3	140.4
	56	KSTHLLIRKLPFSRL	203.0	3881.7	68.2
	57	STHLLIRKLPFSRLA	188.1	3895.0	51.2
	58	THLLIRKLPFSRLAR	181.8	4310.4	44.4
	59	HLLIRKLPFSRLARE	211.8	4498.2	48.6
	60	LLIRKLPFSRLAREI	222.9	4538.8	54.5
	61	LIRKLPFSRLAREIC	321.5	4876.5	85.7
	62	IRKLPFSRLAREICV	378.2	4487.2	124.9
	63	RKLPFSRLAREICVK	456.5	5468.0	141.4
	64	KLPFSRLAREICVKF	391.3	4588.5	121.0
	65	LPFSRLAREICVKFT	459.3	5060.8	134.7
	70	LAREICVKFTRGVDF	181.91	6722.66	430.14
	71	AREICVKFTRGVDFN	143.5	6184.14	330.53
	72	REICVKFTRGVDFNW	115.22	3606.71	259.57
	73	EICVKFTRGVDFNWQ	123.03	3264.66	252.81
	74	ICVKFTRGVDFNWQA	130.89	2920.06	243.49
	75	CVKFTRGVDFNWQAQ	211.61	2645.54	327.23
	76	VKFTRGVDFNWQAQA	416.33	2178.19	448.15
	78	FTRGVDFNWQAQALL	447.63	540.61	187.33
	79	TRGVDFNWQAQALLA	296.61	477.89	78.89
EA	80	RGVDFNWQAQALLAL	175.7	357.28	48.24
	81	GVDFNWQAQALLALQ	186.48	433.62	42.93
	82	VDFNWQAQALLALQE	225.37	498.32	49.91
	83	DFNWQAQALLALQEA	297.42	657.02	60.8
	84	FNWQAQALLALQEAA	424.77	919.85	72.45
	88	AQALLALQEAAEAFL	499.45	735.43	24.84
	89	QALLALQEAAEAFLV	251.33	718.16	21.3
	90	ALLALQEAAEAFLVH	247.9	907.73	24.72
	91	LLALQEAAEAFLVHL	222.56	972.33	32.62
	92	LALQEAAEAFLVHLF	222.81	604.19	66.65
	93	ALQEAAEAFLVHLFE	326.61	435.3	204.64
	94	LQEAAEAFLVHLFED	445.81	304.33	299.55
	97	AAEAFLVHLFEDAYL	477.35	61.41	171.15
	98	AEAFLVHLFEDAYLL	326.8	51.99	106.63
	99	EAFLVHLFEDAYLLT	310.85	52.9	90.65
	100	AFLVHLFEDAYLLTL	267.15	55.82	82.51
	101	FLVHLFEDAYLLTLH	314.62	70.32	95.24
	102	LVHLFEDAYLLTLHA	429.45	115.73	111.14
	104	HLFEDAYLLTLHAGR	404.46	873.56	69.17
	105	LFEDAYLLTLHAGRV	87.65	1214.72	25.75
	106	FEDAYLLTLHAGRVT	68.79	2011.59	18.77
	107	EDAYLLTLHAGRVTL	50.05	2766.4	15.38
	108	DAYLLTLHAGRVTLF	40.32	2062.5	13.95
	109	AYLLTLHAGRVTLFP	42.44	2017.48	14.46
	110	YLLTLHAGRVTLFPK	51.87	2284.67	17.66
	111	LLTLHAGRVTLFPKD	97.22	3118.21	35.61
	112	LTLHAGRVTLFPKDV	318.83	4232.43	148.92

# Table S11. (B) Predicted immunodominant peptides in Centromere Protein B (LD=linkage disequilibrium) (continued on next page).

		<b>.</b>	ACA+ SSc Subset (	CENPB)	in tight LD
Ancestry	Desition	Peptide	DDA1*01.01/DDD1*07.01	Affinity (nM)	DPA1*01/01/DPP1*01/01
	Position 3	PKRROI TEREKSRII	375 95	11822.26	353 19
	4	KRRQLTFREKSRIIQ	316.44	11434.78	225.98
	5	RRQLTFREKSRIIQE	352.09	11172.33	218.04
	6	RQLTFREKSRIIQEV	290.65	8483.02	175.26
	7	QLTFREKSRIIQEVE	419.57	7673.62	246.06
	30	EIARRFNIPPSTLST	495.35	3670.44	104.18
	31	IARRFNIPPSTLSTI	313.12	5271.36	62.96
	32	ARRFNIPPSTLSTIL	253.04	8886.34	52.78
	33	RRFNIPPSTLSTILK	294.55	15951.31	62.32
	34	RFNIPPSTLSTILKN	497.14	15988.46	118.19
	39	PSTLSTILKNKRAIL	467.87	13153.38	93.81
	40	STLSTILKNKRAILA	344.35	10784.31	56.92
	41		338.26	10203.69	47.51
	42		393.67	9622.73	50.87
	45 68		307 16	9650.9	91.01
	69		212.84	3614 64	43.45
	70	KI SPYDKI EGI LIAW	193 34	3322.86	34 99
	71	LSPYDKLEGLLIAWF	166.91	2104.14	30.51
	72	SPYDKLEGLLIAWFQ	203.33	1718.01	34.41
	73	PYDKLEGLLIAWFQQ	263.81	1137.75	43.55
	74	YDKLEGLLIAWFQQI	407.91	528.68	74.25
	77	LEGLLIAWFQQIRAA	442.11	211.2	121.31
	78	EGLLIAWFQQIRAAG	434.52	249.2	94.61
	79	GLLIAWFQQIRAAGL	25.94	279.05	12.04
	80	LLIAWFQQIRAAGLP	25.38	373.15	9.69
	81	LIAWFQQIRAAGLPV	19.47	663.45	8.11
	82	IAWFQQIRAAGLPVK	19.03	1687.92	7.39
	83	AWFQQIRAAGLPVKG	22.38	4397.83	7.83
	84 95		25.70	5498.38	8.55
	85 86		47.05	8575 58	38.69
	87		212.25	9471 19	61 41
	88	IRAAGLPVKGIILKE	364.96	11466.5	120.22
EA	89	RAAGLPVKGIILKEK	389.85	12931.13	125.73
	93	LPVKGIILKEKALRI	91.31	7369.78	37.98
	94	PVKGIILKEKALRIA	73.42	6923.9	21.76
	95	VKGIILKEKALRIAE	82.33	6914.17	21.17
	96	KGIILKEKALRIAEE	96.42	5246.95	22.59
	97	GIILKEKALRIAEEL	133.48	3609.48	29.39
	98		197.54	3621.92	44.93
	99 111		397.54	20/2.15	121.47
	112	GMDDETASNGWLDR	231 32	3800 53	233.65
	113	MDDETASNGWIDRER	205.27	2887.85	180 47
	114	DDFTASNGWLDRFRR	282.66	2939.08	249.32
	115	DFTASNGWLDRFRRR	305.7	2927.53	274.66
	128	RRHGVVSCSGVARAR	488.78	13005.77	127.13
	129	RHGVVSCSGVARARA	465.62	11914.08	105
	133	VSCSGVARARARNAA	494.88	16283.33	92.98
	186	YASQDVFSATETSLW	489.38	5457.13	354.48
	187	ASQDVFSATETSLWY	261.92	5093.95	368.91
	188	SQDVFSATETSLWYD	359.7	5261.22	460
	189		279.01	3574.89	410.28
	197		432.00	200.50	75 /3
	199		480.42	907 77	72 19
	266	KGGVTTQALAKYLKA	497.63	10583.06	165.87
	267	GGVTTQALAKYLKAL	403.36	7981.26	134.97
	268	GVTTQALAKYLKALD	471.8	4837.88	159.07
	269	VTTQALAKYLKALDT	485.14	3868.99	112.51
	271	TQALAKYLKALDTRM	239.47	2736.25	33.03
	272	QALAKYLKALDTRMA	134.65	2182.63	16.75
	273	ALAKYLKALDTRMAA	105.21	2054.29	11.96
	274		118.34	2361.04	12.03
	275	KAI RVI DIBININAED	143.1 183 77	2203.22	15.11
	270	YLKALDTRMAAFSRP	338 52	4806 47	24 49
	280	ALDTRMAAESRRVLL	278.71	10228.89	157.6

Angestry         Peritie         Affield (MM)           Position         Sequence (DM/17/014/02/01/02/01/02/01/02/01/02/01/02/01/02/01/02/01/02/01/02/01/05/01/05/01/02/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01/05/01				ACA+ SSc Subset (C	CENPB)	in tight LD
Position         Sequence         Distance Statul         199 (2017)         12017 (2017)         12017 (2017)         12017 (2017)         12017           281         DTRMACSRYULLA         213 76         6444 37         155 62           283         TRMACSRYULLA         212 75         6781 77         162 24           284         RMALSSRYULLAGR         462 1         7070 9         247 82           286         AASSRYULLAGRIA         244 88         4448 08         32 3           286         SRIVLLAGRIAQ         219 73         2792 3         19           288         SRIVLLAGRIAQS         215 46         215 60 1         18 77           290         RIVLLAGRIAQS         215 46         215 60 1         18 77           303         LDTSGLHYQLAFFP         418 29         206 74 2         268 7           304         LDTSGLHYQLAFFP         418 3         206 74 2         268 7           305         TSGLHYQLAFFPF         418 3         106 75 6         125 13           305         LHYQLAFFPFCTVH         37 2         13 10 71 7         122 41           306         LHYQLAFFPFCTVH         37 2         13 10 71 7         12 48 21           310         LHYQLAFFPFCTVH	Ancestry		Peptide		Affinity (nM)	
281         LDTRMARSINULL         283, 276         6448,37         155,62           283         TRMAASSINULLAGR         462,1         7047,09         207,62           284         MAASSINULLAGR         462,1         7047,09         207,62           286         AACSINULLAGRILAQ         244,95         3290,41         21,79           287         ASSINULLAGRILAQ         249,5         3290,41         21,79           298         SERNULLAGRILAQ         249,5         3290,41         21,79           291         BKVILLAGRILAQ         240,5         329,50         27,73           304         LDTSGRIHVQLAFPP         380,21         263,63,22         304,03           305         LDTSGRIHVQLAFPP         484,4         499,69         221,7           305         DTSGRIHVQLAFPPCTVH         37,3         105,7,5         140,51           306         DTSGRIHVQLAFPPCTVH         37,3         105,7,5         140,51           313         QUAFPPCTVHPL         35,21         917,7,6         148,15           313         QUAFPPCTVHPL         35,21         917,5         140,51           313         QUAFPPCTVHPL         35,21         917,50         143,33           313 <th></th> <th>Position</th> <th>Sequence</th> <th>DRA1*01:01/DRB1*07:01</th> <th>DQA1*01:01/DQB1*05:01</th> <th>DRA1*01:01/DRB1*01:01</th>		Position	Sequence	DRA1*01:01/DRB1*07:01	DQA1*01:01/DQB1*05:01	DRA1*01:01/DRB1*01:01
282         DIRMAASSINULLAG         27.3 /b         648.4 /r         125.5 /r           283         RMAASSINULLAG         27.1 /s         12.2 /r         127.2 /r           284         RMAASSINULLAGR         24.4 /r         127.9 /r         12.7 /r           285         ASSINULLAGRIAA         24.4 /r         17.9 /r         12.7 /r           286         SINULLAGRIAADS         12.1 /r         12.7 /r         12.7 /r           286         SINULLAGRIAADS         21.5 /r         25.6 /r         17.8 /r           290         RIVILLAGRIAADS         21.5 /r         25.6 /r         17.8 /r           303         LDTSGLINVQLAFPP         42.2 /r         25.6 /r         27.7 /r         17.5 /r           304         LDTSGLINVQLAFPPGV         38.3 /r         271.7 /r         14.3 /r         30.7 /r           305         TSGLINVQLAFPPGV         38.7 /r         10.7 /r         14.4 /r         30.7 /r         31.3 /r         31.1 /r         32.7 /r         31.1 /r		281	LDTRMAAESRRVLLL	189.26	7053.35	119.67
243         INVOLUSION LLAG         2/1.9-3         0-0.1/7         102.24           284         AAESRIVILLAGINA         244.95         320.41         2.1.79           285         ASERNULLAGINA         224.95         320.41         2.1.79           286         ESRIVILLAGINACIS         210.31         2624.66         17.81           290         RIVULAGINACIS         210.31         2624.66         17.81           291         RIVULAGINACIS         213.39         2577.05         27.83           303         SLDTSGERHVQLAFFP         360.21         263.83.2         304.03           304         DTSGERHVQLAFFP         488.9         2067.42         268.67           305         DTSGERHVQLAFFP         477.36         1057.65         125.13           305         DTSGERHVQLAFFPCTIVHP         357.22         97.7.65         125.13           305         DTSGERHVQLAFFPCTIVHP         357.22         97.7.65         125.13           305         DTSGERHVQLAFFPCTIVHP         357.22         97.7.65         125.13           306         UAEFPCTIVHPE         357.22         97.66         128.62           307         VQLAFFPCTIVHPE         352.1         417.7         128.42		282		253.76	6448.37	155.62
286         AUDICATE         224.32         224.33         227           287         AESRAVILLAGRIAA         224.35         2270.41         21.79           288         ESRAVILLAGRIAAQ         21.97         279.23         19           289         SRAVILLAGRIAAQD         21.54.6         2156.01         18.77           290         RAVILLAGRIAAQD         21.54.6         2156.01         18.77           291         RAVILLAGRIAAQD         21.54.6         2156.07         27.13           303         LDTSGRIHVOLAFFP         41.88.9         2067.42         268.87           305         DTSGRIHVOLAFFP         41.8.9         2067.42         268.87           306         TSGLIHVOLAFFPGV         33.72         971.76         148.15           308         GTSGLIHVOLAFFPGV         347.36         1057.65         125.13           310         CHAVQLAFFPGTVHPE         246.97         1508.04         106.6           312         VOQUKGHYROAMLIKA         467.92         487.66         328.62           313         QUKGHYROAMLIKA         47.72         284.62         13.6           313         QUKGHYROAMLIKA         47.72         284.50         13.6           314		283		271.95	5/81.77	102.24
207         ALSRIVILLAGRIAA         224.95         3290.41         21.79           228         ESRIVILLAGRIAAD         220.31         226.46         91.78.11           229         RRVILLAGRIAADS         220.31         226.46         91.78.11           230         RRVILLAGRIAADS         235.46         215.601         18.77           241         RRVILLAGRIAADS         233.99         257.05         27.83           303         SLDTSGIRHVQLAFF         300.21         236.32         304.03           304         IDTSGIRHVQLAFFP         484.4         199.69         221.7           309         ILGRIAMQLAFFP         484.4         199.69         221.7           309         ILGRIAMQLAFFP         397         1305.75         140.51           310         RHVQLAFFPEOTVH         37.76         148.15         115.13           313         QUAFFPEOTVHPIL         246.97         1508.64         10.66           313         QUAFFPEOTVHPIL         246.97         1508.64         10.66           313         QUAFFPEOTVHPIL         242.21         470.55         12.51.31           313         QUAFFPEOTVHPIL         246.97         1508.61         22.41           329 <td></td> <td>286</td> <td></td> <td>204 88</td> <td>7047.09</td> <td>247.62</td>		286		204 88	7047.09	247.62
288         ESRPULLAGRUAAGS         219.73         272.3         19           299         SRVLLAGRUAAGS         203.1         266.46.9         17.81           290         SRVLLAGRUAAGS         203.1         266.46.9         17.81           291         NVLLAGRUAAGS         215.46         2156.01         17.87           303         LDTSGLIHVQLAFF         360.21         286.87         27.83           304         LDTSGLIHVQLAFFP         418.89         2067.42         268.87           305         TSGLIHVQLAFFPGV         434.4         1499.69         221.7           306         TSGLIHVQLAFFPGV         37.36         1057.65         125.13           306         ILMQLAFFPGTVHP         246.97         1308.04         106.6           311         HVQLAFFPGTVHPLE         246.97         1308.04         106.6           312         VQLAFFPGTVHPLE         246.97         1308.04         106.6           313         QLAFFPGTVHPLE         246.97         1308.04         106.6           313         QLAFFPGTVHPLE         246.97         236.47         236.26           333         QLAFFPGTVHPLE         247.92         237.81         13.6           333		287	AFSRRVIIIAGRIAA	224.00	3290 41	21 79
289         SRVLLAGRI, AAQSL         220.31         257.65         17.81           291         RVLLLAGRI, AAQSL         215.46         215.60         13.77           291         RVLLLAGRI, AAQSL         329.9         2577.05         27.83           303         SLDTSGLRHVQLAFFP         360.21         2265.32         304.03           304         LDTSGLRHVQLAFFP         472.25         195.07         221.5           305         DTSGLRHVQLAFFP         472.25         195.07         21.5           306         GLRHVQLAFFPFGTVH         33.3.72         971.76         148.15           301         RHVQLAFFPFGTVHP         397         1305.75         140.51           313         RHVQLAFFPFGTVHPLE         246.97         150.80         10.66           312         VQLAFFPFGTVHPLE         282.31         201.51         115.43           313         QUXFFPFGTVMPLE         282.32         425.04         286.64           313         QUXFFPFGTVMPLE         282.31         231.66         333.16           313         QUXFFFFGTVMMLKAM         177.71         393.494         22.24           313         QUXFFFFGTVMMLKAMALE         178.11         173.13         16.16		288	ESRRVLLLAGRLAAQ	219.73	2792.3	19
200         RRVLLAGRI-ANQSL         215.46         215.01         18.77           303         SLOTSGI, HVQLAFF         360.21         2263.32         304.33           304         LDTSGI, HVQLAFF         360.21         2263.32         304.33           305         DTSGI, HVQLAFFP         418.89         2057.02         221.5           305         DTSGI, HVQLAFFP         412.25         1955.07         221.5           306         GLNHVQLAFFPGTV         33.72         971.76         (148.15           303         LRIVQLAFFPGTVHPE         246.97         1508.04         106.6           312         VUQLAFFPGTVHPLE         246.97         1508.04         106.6           313         QLAFFPGTVHPLE         33.21         417.78         122.41           313         QLAVGHYRQAMLLKA         467.22         428.62         33.22           313         QUVGVGHYRQAMLKA         127.17         394.94         23.24           313         QUVGVGHYRQAMLKAM         167.77         394.94         23.24           314         GUVGVGHYRQAMLKAMA         167.71         394.94         23.24           313         QUVGVGHYRQAMLKAMA         167.71         394.94         23.24		289	SRRVLLLAGRLAAQS	220.31	2624.69	17.81
291         RVLLAGRLAAQSLD         339.9         2577.05         27.83           303         LDTSGLHWQLAFF         86.02.1         2363.27         268.87           305         DTSGLHWQLAFFP         418.89         2067.42         268.87           306         DTSGLHWQLAFFPGT         487.42         1499.69         221.7           306         DTSGLHWQLAFFPGTW         387.22         971.76         148.15           301         HWQLAFFPGTWH         347.36         1057.65         123.13           311         HWQLAFFPGTWHP         397         1306.75         140.51           313         RHWQLAFFPGTWHP         246.97         1308.02         194.9           313         QUAFFPGTWHER         353.21         111.51         135.43           313         QUAFFPGTWAMLIKA         247.82         4250.51         53.26           313         QUAFFPGTWAMLIKAM         167.72         394.94         23.24           313         QUAVGHTRQAMLIKAMA         167.71         394.94         23.24           314         QUAVGHTRQAMLIKAMAA1         152.47         237.92         13.35           315         GHTRQAMLIKAMAA1         152.47         237.92         13.85           31		290	RRVLLLAGRLAAQSL	215.46	2156.01	18.77
303         SLDTSGRHVQLAFF         360.21         2363.32         304.03           304         LDTSGRHVQLAFFP         472.25         1933.07         271.5           305         DTSGRHVQLAFFPR         472.25         1933.07         271.5           308         GLRHVQLAFFPRGTV         353.72         971.76         144.15           309         LRHVQLAFFPRGTVH         347.36         1057.65         1425.13           310         RHVQLAFFPRGTVHPL         246.97         1568.04         10.66           312         VQLAFFPRGTVHPL         246.97         1568.04         10.66           313         QLAFFPRGTVHPL         246.97         1568.04         10.66           313         QLAFFPRGTVHPL         246.97         287.66         122.86           331         QLAFFPRGTVHPL         246.97         286.05         153.26           332         VUQUKGHYRQAMLIKA         467.92         487.66         122.86           333         QLWGHYRQAMLIKAM         179.93         3845.08         31.66           333         VGGWKGHYRQAMLIKAM         167.77         393.94.50         13.66           334         KGHYRQAMLIKAMAAL         127.1         207.77.27         153.27		291	RVLLLAGRLAAQSLD	339.9	2577.05	27.83
304         LDTSGLRHVQLAFFP         418.89         2067.42         288.7           305         DTSGLRHVQLAFFPC         484.4         1499.69         221.7           306         TSGLRHVQLAFFPPCTV         353.72         97.17         6         148.15           309         LEHVQLAFFPPCTVH         347.36         1057.65         125.13           310         RIVQLAFFPPCTVHP         397         1306.75         140.51           313         RIVQLAFFPPCTVHPLE         282.31         2011.51         115.4           313         QLAFFPCTVHPLE         283.21         471.78         122.41           320         VQQVKGHYRQAMLLKA         242.82         4876.61         228.62           331         QLAFFPCTVHPLE         323.11         471.78         322.41           320         VQQVKGHYRQAMLLKAM         127.27         393.49         23.24           333         VKGHYRQAMLLKAM         127.17         293.49         23.24           333         VKGHYRQAMLLKAMA         147.18         208.76         14.86           333         VKGHYRQAMLLKAMAALEG         128.17         13.16         133.8           334         PSGLQLGTEALHF         75.41         197.31         161.73		303	SLDTSGLRHVQLAFF	360.21	2363.32	304.03
305         DTSGLRHVQLAFFPP         472.25         1953.07         271.5           306         DTSGLRHVQLAFFPPGTVH         333.72         971.76         148.15           309         LRHVQLAFFPPGTVH         337.75         125.13           310         RHVQLAFFPPGTVHP         246.97         1508.04         106.6           311         HVQLAFFPPGTVHPL         246.97         1508.04         106.6           312         VQLAFFPPGTVHPL         235.21         4171.78         122.41           292         VVQLVKGHYRQAMLLK         467.92         487.661         228.62           330         VQLVKGHYRQAMLLKA         467.92         487.661         228.62           331         QUXGHYRQAMLLKAMA         167.77         393.454         23.24           333         VQUKGHYRQAMLLKAMAA         167.71         393.454         23.26           333         VKGHTRQAMLLKAMAA         152.47         2837.92         15.35           333         KGHTRQAMLLKAMAA         152.71         2037.84         12.87           334         KGHTRQAMLLKAMAALE         152.71         2037.84         12.87           335         GGHTRQAMLLKAMAALE         152.71         2037.84         12.87		304	LDTSGLRHVQLAFFP	418.89	2067.42	268.87
306         TSGIRHVQLAFFPPG         484.4         1499.69         221.7           308         CIRNVQLAFFPPGTVH         347.36         1057.65         125.13           310         RIVQLAFFPGTVHP         246.97         1508.04         106.6           311         HVQLAFFPGTVHPE         282.31         2011.51         115.43           312         VQLAFFPGTVHPE         283.21         4171.78         122.41           329         VVQQVKGHYRQAMLLK         467.22         4876.61         228.62           330         VQQVKGHYRQAMLLKA         242.82         4250.51         53.26           331         QQVKGHYRQAMLLKA         242.82         4250.51         53.26           333         VKGHYRQAMLLKAMA         167.77         3934.94         23.24           335         GHYRQAMLLKAMAALE         153.59         153.53         13.61           336         GHYRQAMLLKAMAALE         147.18         2037.68         15.03           337         YRQAMLLKAMAALE         147.18         204.82.76         14.86           340         MALLKAMAALEGOD         224.18         248.26         14.86           341         MLKAMAALEGODP         20.36         243.94         15.66		305	DTSGLRHVQLAFFPP	472.25	1953.07	271.5
308         GLRHVQLAFFPGTVH         333.72         971.76         148.13           309         LRHVQLAFFPGTVHP         397         1305.75         140.51           310         RIVQLAFFPGTVHPL         326.97         1508.04         106.6           312         VQLAFFPGTVHPLE         282.31         2011.51         115.43           313         QLAFFPGTVHPLE         282.31         2011.51         125.42           329         VVQQVKGHYRQAMLLK         467.92         4876.61         228.62           330         QUXGKHYRQAMLLKA         467.92         4876.61         228.62           331         QUXGKHYRQAMLKA         422.82         420.51         53.25           332         QVKGHYRQAMLKAMA         147.21         241.92         17.13           335         GHYRQAMLKAMAALE         152.47         2337.92         15.33           336         HKOMMLKAMAALEG         153.39         198.19         13.61           337         YRQAMLKAMAALEGOD         224.18         2247.26         14.84           340         AMLLKAMAALEGODP         224.12         17.73         33.14           345         SGLQLGITEALHFV         35.13         683.6         2.6           355 </td <td></td> <td>306</td> <td>TSGLRHVQLAFFPPG</td> <td>484.4</td> <td>1499.69</td> <td>221.7</td>		306	TSGLRHVQLAFFPPG	484.4	1499.69	221.7
309         LRHVQLAFFPGTVHP         397         1305.75         140.51           311         HVQLAFFPGTVHPL         246.97         1508.04         106.6           312         VQLAFFPGTVHPLE         282.31         2011.51         115.43           313         OLAFFPGTVHPLR         353.21         4171.78         122.41           329         VVQQVKGHYRQAMLL         467.92         4876.61         228.62           330         VQQVKGHYRQAMLLKA         424.82         4250.51         53.326           331         QUKGHYRQAMLLKA         242.82         4250.51         53.326           333         VQGVKGHYRQAMLLKAM         167.77         3934.94         23.24           334         KGHYRQAMLLKAMAALE         147.18         2037.68         15.03           335         GHYRQAMLKAMAALE         17.18         244.26         14.86           340         AMLLKAMAALEGQ         162.71         2037.68         15.03           338         RQAMLKAMAALEGQ         224.18         2442.26         14.86           341         MLLKAMAALEGQDPS         382.12         2788.16         21.12           355         GLQLGITEALHFVA         3.14         1973.12         61.73 <t< td=""><td></td><td>308</td><td>GLRHVQLAFFPPGTV</td><td>353.72</td><td>971.76</td><td>148.15</td></t<>		308	GLRHVQLAFFPPGTV	353.72	971.76	148.15
310         HHQUAFPPG1VHP         397         1305./5         105./5           311         HVQLAFPPG1VHPLE         226.97         1508.04         106.6           312         VQLAFPPGTVHPLE         223.11         211.51         115.43           312         VQLAFFPGTVHPLE         223.11         417.78         122.41           329         VQQVKGHYRQAMLLK         467.28         4830.82         194.9           331         QUKKHYRQAMLLKAMA         179.93         3845.08         31.66           332         QVKGHYRQAMLLKAMA         147.21         3419.21         17.13           333         VKGHYRQAMLLKAMAA         147.21         3419.21         17.13           333         VKGHYRQAMLLKAMAA         147.21         2319.2         15.35           334         KGHYRQAMLLKAMAA         152.47         2837.92         15.35           335         HYRQAMLLKAMAALEG         152.35         1958.19         13.61           336         HYRQAMLLKAMAALEG         152.35         1958.19         13.61           337         YRQAMLLKAMAALEG         152.35         293.64         21.48           340         MULKAMAALEGOP         224.18         2248.26         14.486		309	LRHVQLAFFPPGTVH	347.36	1057.65	125.13
311         NUCLAPPPENTWPLE         248.37         1308.04         106.8           313         QUAPPPENTWPLE         823.31         201.15.1         115.43           313         QUVGCWSHYRQAMLLK         427.92         476.61         228.62           330         VQQVKGHYRQAMLLK         427.82         430.82         194.9           331         QUVKGHYRQAMLLKA         242.82         4250.51         53.26           332         QUVKGHYRQAMLLKA         147.21         341.92         17.13           334         KGHYRQAMLLKAMA         167.77         3934.94         23.24           333         KGHYRQAMLLKAMAALE         147.18         2037.68         15.03           337         YRQAMLLKAMAALEGQ         167.71         2037.68         15.03           337         YRQAMLLKAMAALEGQ         167.71         2037.68         15.03           338         ROAMLLKAMAALEGQ         167.71         2037.68         15.03           341         MLLKAMAALEGQDP         280.36         2439.41         16.68           341         MLLKAMAALEGQDP         280.36         2439.41         16.68           341         MLLKAMAALEGQDP         280.36         2439.41         16.68		310	RHVQLAFFPPGTVHP	397	1305.75	140.51
312         VQLAPPPGVMPLER         283.21         2011.31         11.343           313         QLAFPPGVMPLER         353.21         417.78         122.41           329         VQQVKGHYRQAMLLK         467.28         483.08.2         194.9           331         QQVKGHYRQAMLLK         22.42         4250.51         53.26           332         QVKGHYRQAMLLKAM         179.93         3845.08         31.66           333         VKGHYRQAMLLKAMAA         147.71         3419.21         17.13           334         KGHYRQAMLLKAMAA         147.71         2837.92         15.35           335         GHYRQAMLLKAMAALE         152.47         2837.92         15.35           336         HYRQAMLLKAMAALEGQ         152.17         2037.68         15.03           337         YRQAMLLKAMAALEGQ         162.71         2037.64         12.87           339         QAMLLKAMAALEGQD         224.18         2248.26         14.486           340         AMLKAMAALEGQDP         224.18         2248.26         13.486           357         LQLGITEALHFVA         37.17         1262.23         33.01           356         GLQLGITEALHFVA         37.17         1262.24         33.44		311		246.97	1508.04	106.6
313         VQQVKGHYRQAMILK         467.92         477.61         228.62           330         VQQVKGHYRQAMILK         467.28         4830.82         194.9           331         QQVKGHYRQAMILKA         422.82         4250.51         53.36           332         QVKGHYRQAMILKAM         179.93         3845.08         31.66           333         VKGHYRQAMILKAMA         167.77         3934.94         23.24           333         KGHYRQAMILKAMAAL         152.47         283.792         15.35           336         GHYRQAMILKAMAALEG         153.59         1958.19         13.61           333         RQAMILKAMAALEG         152.71         2037.68         15.03           333         RQAMILKAMAALEG         152.71         2037.64         12.87           333         QAMILKAMAALEGQDP         280.36         2439.41         16.68           341         MILKAMAALEGQDP         280.36         2439.41         16.68           341         MILKAMAALEGQDP         382.12         2788.15         21.12           353         DPSCLQLGITEA.HF         75.41         197.73         344.3           354         GLQLGITEA.HF         75.41         97.8         366           355<		31Z 212		282.31	2011.51	115.43
12.0         VQQVKHYRQAMILKA         10.12         10.01         12.01           331         QQVKGHYRQAMILKA         242.82         4250.51         53.26           332         QVKGHYRQAMILKA         179.93         3845.08         31.66           333         VKGHYRQAMILKAMAA         167.77         3934.94         23.24           334         KGHYRQAMILKAMAA         147.21         3419.21         17.13           335         GHYRQAMILKAMAALE         147.18         2037.68         15.03           337         YRQAMILKAMAALEGQD         122.11         2037.84         12.87           338         RQAMILKAMAALEGQD         122.11         2037.84         12.87           339         QAMILKAMAALEGQD         224.18         2248.26         14.86           340         AMILKAMAALEGQDP         280.36         2439.41         16.68           341         MILKAMAALEGQDP         280.36         1236.62         33.48           355         SGLQLGITEALHFV         45.14         1408.73         38.76           355         SGLQLGITEALHFVA         37.17         1262.23         33.48           357         LUGITEALHFVAAAWQA         10.12         914.02         66.47		329	VVOOVKGHYROAMU	467.92	4171.78	228.62
121         1241.00         1242.9         1250.51         53.26           332         QVKGHYRQAMLLKAM         179.93         3845.08         31.66           333         QVKGHYRQAMLLKAM         167.77         393.44         23.24           334         KGHYRQAMLLKAMAA         147.21         3419.21         17.13           335         GHYRQAMLLKAMAALE         152.47         2837.92         15.35           336         HYRQAMLLKAMAALEG         153.59         1958.19         13.61           337         YRQAMLLKAMAALEG         152.59         1958.19         13.61           338         RQAMLLKAMAALEGQDP         280.36         2439.41         16.68           341         MLKAMAALEGQDP         280.36         1236.62         33.41           353         DPSCLOLGUTEALHF         75.41         197.12         61.73           354         SGLUGUTEALHFVA         37.17         126.23         33.01           355         SGLUGUTEALHFVA         37.17         126.23         33.01           355         SGLUGUTEALHFVA         37.17         126.23         33.61           356         GLUGUTEALHFVAA         36.5         123.66         26.45           357		320	VOOVKGHVROAMULK	467.28	4830 82	194 9
323         QVKGHYRQAMILKAM         179.93         3945.08         31.66           333         VKGHYRQAMILKAMA         167.77         3934.94         23.24           334         KGHYRQAMILKAMAA         147.21         3419.21         17.13           335         GHYRQAMILKAMAAL         152.47         2837.92         15.35           336         HYRQAMILKAMAALE         147.18         2037.68         15.03           337         YRQAMILKAMAALEGQ         162.71         2037.84         12.87           338         RQAMILKAMAALEGQ         122.18         2248.26         14.48           340         AMILKAMAALEGQP         220.36         2439.41         16.68           341         MILKAMAALEGQPS         382.12         2788.16         21.12           353         DPSGLQIGUTEALHFV         45.14         1408.73         38.75           355         SCIQLGUTEALHFVA         37.17         1262.23         33.01           355         SCIQLGUTEALHFVAA         34.65         1236.62         33.48           357         LQUGUTEALHFVAAA         36.65         126.62         33.48           356         GLQUGUTEALHFVAA         36.65         126.12           356         G		330		2407.28	4030.02	53.26
333         VKGHYRQAMILKAMAA         167.77         3934.94         23.24           334         KGHYRQAMILKAMAA         147.21         3419.21         17.13           335         GHYRQAMILKAMAA         152.47         2837.92         15.35           336         HYRQAMILKAMAALE         147.18         2037.68         15.03           337         YRQAMILKAMAALEGQ         162.71         2037.84         12.87           339         QAMILKAMAALEGQD         224.18         2248.26         14.86           340         AMULKAMAALEGQD         280.36         2439.41         16.68           341         MILKAMAALEGQDPS         382.12         2788.16         21.12           353         DPSGLQLGITEALHFV         45.14         4008.73         38.76           355         SGLQLGITEALHFVA         37.17         1262.23         33.01           356         GLQLGITEALHFVA         37.17         1264.73         44.3           357         LQLGITEALHFVAA         40.73         1241.73         44.3           358         GLQLGITEALHFVAA         40.73         1241.73         44.3           360         GLTEALHFVAAAWQAVE         52.96         659.33         22.174           3		332	OVKGHYROAMLLKAM	179.93	3845.08	31.66
334         KGHYRQAMLLKAMAA         147.21         3419.21         17.13           335         GHYRQAMLLKAMAAL         152.47         2837.92         15.35           336         HYRQAMLLKAMAALEG         153.59         1958.19         13.61           337         YRQAMLLKAMAALEGQ         162.71         2037.84         12.87           338         RQAMLLKAMAALEGQ         224.18         2248.26         14.86           340         AMLLKAMAALEGQDP         220.36         2439.41         16.68           341         MLLKAMAALEGQDP         280.36         2439.41         16.86           353         DPSGLQLGITEALHFVA         37.17         1262.23         33.01           356         GLQLGITEALHFVAA         40.73         1241.73         44.3           357         LQLGITEALHFVAA         40.73         1241.73         44.3           360         GITEALHFVAA         40.73         1241.73         44.3           361 </td <td></td> <td>333</td> <td>VKGHYROAMLLKAMA</td> <td>167.77</td> <td>3934.94</td> <td>23.24</td>		333	VKGHYROAMLLKAMA	167.77	3934.94	23.24
335         GHYRQAMLLKAMAALE         152.47         2837.92         15.35           336         HYRQAMLLKAMAALE         147.18         2037.68         15.03           337         YRQAMLLKAMAALEGQ         153.59         1958.19         13.61           338         RQAMLLKAMAALEGQ         162.71         2037.84         12.87           339         QAMLLKAMAALEGQD         224.18         2248.26         14.86           340         AMLLKAMAALEGQD         280.36         2439.41         16.68           341         MLLKAMAALEGQD         280.36         2439.41         16.68           341         MLLKAMAALEGQD         280.36         2439.41         16.68           341         MLLKAMAALEGQD         280.36         2439.41         16.68           341         MLKAMAALEGQD         280.37         383.76         338.76           355         SGLQIGITEALHFVA         37.17         1262.23         33.01           356         GLQIGITEALHFVA         40.73         1241.73         44.3           350         GITEALHFVAAAWQ         161.26         838.91         61.96           361         LITEALHFVAAAWQAVE         52.96         659.33         21.74           363 <td></td> <td>334</td> <td>KGHYRQAMLLKAMAA</td> <td>147.21</td> <td>3419.21</td> <td>17.13</td>		334	KGHYRQAMLLKAMAA	147.21	3419.21	17.13
336         HYRQAMULIKAMAALE         147.18         2037.68         15.03           337         YRQAMULIKAMAALEG         153.59         1958.19         13.61           338         RQAMULIKAMAALEG         162.71         2037.68         12.87           339         QAMULKAMAALEGQD         224.18         2248.26         14.86           340         AMULKAMAALEGQDP         382.12         2788.16         21.12           353         DPSGLQLGUTEALHFV         45.14         1408.73         38.76           354         PSGLQLGUTEALHFV         45.14         1408.73         38.76           355         SGLQLGUTEALHFVA         37.17         1262.23         33.01           356         GLQLGUTEALHFVA         40.73         1241.73         44.3           357         LQIGUTEALHFVA         40.73         1241.73         44.3           358         QLGUTEALHFVAAAWQ         121.27         875.41         97.8           360         GUTEALHFVAAAWQ         121.27         875.41         97.8           361         LTEALHFVAAAWQAVE         52.96         659.3         21.74           362         TEALHFVAAAWQAVE         52.96         659.3         21.74           363		335	GHYRQAMLLKAMAAL	152.47	2837.92	15.35
337         YRQAMLLKAMAALEG         153.59         1958.19         13.61           338         RQAMLLKAMAALEGQ         162.71         2037.84         12.87           343         RQAMLLKAMAALEGQD         224.18         2248.26         14.86           340         AMLLKAMAALEGQDP         280.36         2439.41         16.68           341         MLLKAMAALEGQDPS         382.12         2788.16         21.12           353         DPSGLQIGITEALHFV         45.14         1408.73         38.76           355         SGLQIGITEALHFVA         37.17         1262.23         33.01           356         GLQIGITEALHFVAA         34.65         1236.62         33.48           357         LQIGITEALHFVAA         40.73         1241.73         44.3           358         QLGITEALHFVAAAWQA         151.2         914.02         66.47           359         LGITEALHFVAAAWQAV         53.13         683.6         26           360         GITEALHFVAAAWQAV         53.13         683.6         26           361         LTEALHFVAAAWQAVE         57.93         603.24         21.78           364         ALHFVAAAWQAVEPSD         89.03         28.693         28.39           365		336	HYRQAMLLKAMAALE	147.18	2037.68	15.03
338         RQAMLIKAMAALEGQ         162.71         2037.84         12.87           339         QAMLIKAMAALEGQD         224.18         2248.26         14.86           340         AMULKAMAALEGQDP         280.36         2439.41         16.68           341         MILKAMAALEGQDPS         382.12         2788.16         21.12           353         DPSGLQLGITEALHFV         45.14         1408.73         38.76           355         SGLQLGITEALHFVA         37.17         1262.23         33.01           356         GLQLGITEALHFVAA         40.73         1211.73         44.33           358         QLGITEALHFVAA         40.73         1241.73         44.33           358         QLGITEALHFVAAAWQ         121.87         875.41         97.83           360         GITEALHFVAAAWQA         161.26         838.91         61.96           361         LITEALHFVAAAWQAVE         52.96         659.3         21.74           363         EALHFVAAAWQAVE         52.96         659.3         21.78           364         ALHFVAAAWQAVEPSD         89.03         826.93         28.39           365         HFVAAAWQAVEPSD         201.46         1046.22         96.05           366 <td></td> <td>337</td> <td>YRQAMLLKAMAALEG</td> <td>153.59</td> <td>1958.19</td> <td>13.61</td>		337	YRQAMLLKAMAALEG	153.59	1958.19	13.61
339         QAMLLKAMAALEGQDP         224.18         2248.26         14.86           340         AMLLKAMAALEGQDP         280.36         2439.41         16.68           341         MULKAMAALEGQDP         382.12         2788.16         21.12           353         DPSGLQLGLTEALHFV         45.14         1973.12         61.73           354         PSGLQLGLTEALHFVA         37.17         1262.23         33.01           356         GLQLGLTEALHFVAA         34.65         1236.62         33.48           357         LQLGITEALHFVAA         40.73         1241.73         44.3           358         QLGITEALHFVAA         40.73         1241.73         44.3           360         GLTEALHFVAAAWQA         121.87         875.41         97.8           360         GLTEALHFVAAAWQA         53.13         683.6         26           362         TEALHFVAAAWQAVEP         57.93         603.24         21.74           363         ELHFVAAAWQAVEPSD         89.03         826.93         28.39           364         HFVAAAWQAVEPSD         89.03         826.93         28.39           365         LHFVAAAWQAVEPSD         201.46         1046.22         96.05           368		338	RQAMLLKAMAALEGQ	162.71	2037.84	12.87
340         AMILIKAMAALEGQDP         280.36         2439.41         16.68           341         MILIKAMAALEGQDPS         382.12         2788.16         21.12           351         DPSGLQLGITEALHF         75.41         1973.12         61.73           355         SGLQLGITEALHFVA         37.17         1262.23         33.01           355         SGLQLGITEALHFVA         34.65         1236.62         33.48           357         LQLGITEALHFVAA         40.73         1241.73         44.3           358         QLGITEALHFVAAAWQ         61.12         914.02         66.47           359         IGITEALHFVAAAWQA         161.26         838.91         61.96           360         GLITEALHFVAAAWQAV         53.13         683.6         26           361         ITEALHFVAAAWQAVEP         57.93         603.24         21.78           364         ALHFVAAAWQAVEPS         63.45         618.86         22.45           365         HFVAAAWQAVEPSDI         114.93         890.53         45.34           367         FVAAAWQAVEPSDIA         201.46         1046.22         96.05           368         VAAWQAVEPSDIA         206.31         306.07.91         164.34           370<		339	QAMLLKAMAALEGQD	224.18	2248.26	14.86
341         MLLKAMAALEGQDPS         382.12         2788.16         21.12           353         DPSGLQLGITEALHFV         45.14         10973.12         61.73           355         SGLQLGITEALHFV         45.14         10973.12         38.76           355         SGLQLGITEALHFVA         37.17         1262.23         33.01           356         GLQLGITEALHFVAA         40.73         1241.73         44.3           358         QLGITEALHFVAAA         40.73         1241.73         44.3           358         QLGITEALHFVAAAWQ         161.26         838.91         61.96           361         LITEALHFVAAAWQAVE         52.96         659.3         21.74           362         TEALHFVAAAWQAVE         57.93         603.24         21.78           364         ALHFVAAAWQAVEPS         63.45         618.86         22.45           365         HFVAAAWQAVEPSD         114.93         890.53         45.34           367         FVAAAWQAVEPSD         308.43         5990.71         204.19           368         VAAAWQAVEPSDIA         201.46         1046.22         96.05           370         AAWQAVEPSDIAA         308.33         5090.71         204.19           388		340	AMLLKAMAALEGQDP	280.36	2439.41	16.68
353         DPSGLQLGLTEALHF         75.41         1973.12         61.73           EA         354         PSGLQLGLTEALHFV         45.14         1408.73         38.76           355         SGLQLGLTEALHFVA         37.17         1262.23         33.01           356         GLQLGLTEALHFVAA         40.73         1241.73         44.3           357         LQLGITEALHFVAAA         40.73         1241.73         44.3           359         LGLTEALHFVAAAWQ         121.87         875.41         97.8           360         GLTEALHFVAAAWQA         161.26         838.91         61.96           361         LTEALHFVAAAWQAVE         52.96         659.3         21.74           363         EALHFVAAAWQAVE         52.96         603.24         21.78           364         ALHFVAAAWQAVEP         57.93         603.24         21.78           365         HFVAAAWQAVEPSD         89.03         826.93         28.39           366         HFVAAAWQAVEPSD         89.03         826.93         28.39           366         HFVAAAWQAVEPSDIA         206.4         138.31         369           367         FVAAAWQAVEPSDIA         205.34         1990.64         138.31 <td< td=""><td></td><td>341</td><td>MLLKAMAALEGQDPS</td><td>382.12</td><td>2788.16</td><td>21.12</td></td<>		341	MLLKAMAALEGQDPS	382.12	2788.16	21.12
EA         354         PSGLQLGLTEALHFV         45.14         1408.73         38.76           355         SGLQLGLTEALHFVA         37.17         1262.23         33.01           356         GLQLGLTEALHFVAA         34.65         1236.62         33.48           357         LQLGLTEALHFVAAA         40.73         1241.73         44.3           358         QLGLTEALHFVAAAWQ         121.87         875.41         97.8           360         GLTEALHFVAAAWQA         161.26         838.91         61.96           361         LTEALHFVAAAWQAV         53.13         683.6         26           362         TEALHFVAAAWQAVE         52.96         659.3         21.74           364         ALHFVAAAWQAVEPS         63.45         618.86         22.45           365         HFVAAAWQAVEPSD         89.03         826.93         28.39           366         HFVAAAWQAVEPSDIA         201.46         1046.22         96.05           368         VAAAWQAVEPSDIA         201.46         108.31         366           369         AAAWQAVEPSDIAA         256.34         1990.64         138.31           369         VAAWQAVEPSDIAA         256.34         1990.71         204.19           <		353	DPSGLQLGLTEALHF	75.41	1973.12	61.73
355       SGLQL GLTEALHFVA       37.17       1262.23       33.01         356       GLQL GLTEALHFVAA       34.65       1236.62       33.48         357       LQL GLTEALHFVAAA       40.73       1241.73       44.3         358       QLGLTEALHFVAAAWQ       61.12       914.02       66.47         359       LGLTEALHFVAAAWQA       161.26       838.91       61.96         361       LTEALHFVAAAWQAV       53.13       683.6       26         362       TEALHFVAAAWQAVE       52.96       659.3       21.74         363       EALHFVAAAWQAVE       57.93       603.24       21.78         364       ALHFVAAAWQAVEPS       63.45       618.86       22.45         365       LHFVAAAWQAVEPSD       89.03       826.93       28.39         366       HFVAAAWQAVEPSDIA       201.46       1046.22       96.05         368       VAAAWQAVEPSDIA       256.34       1990.64       138.31         369       AAAWQAVEPSDIAAC       308.33       506.71       204.19         538       EVPVPSFGEAMAYFA       348.08       3153.32       146.22         539       VPVPSFGEAMAYFA       348.08       3153.32       146.1         541 <td>EA</td> <td>354</td> <td>PSGLQLGLTEALHFV</td> <td>45.14</td> <td>1408.73</td> <td>38.76</td>	EA	354	PSGLQLGLTEALHFV	45.14	1408.73	38.76
356       GLQL GLTEALHFVAAA       34.65       1236.62       33.48         357       LQL GLTEALHFVAAAW       40.73       1241.73       44.3         358       QL GLTEALHFVAAAWQ       121.87       875.41       97.8         360       GLTEALHFVAAAWQA       161.26       888.91       61.96         361       LTEALHFVAAAWQAV       53.13       683.6       26         362       TEALHFVAAAWQAVE       52.96       659.3       21.74         363       EALHFVAAAWQAVEP       57.93       603.24       21.78         364       ALHFVAAAWQAVEPS       63.45       618.86       22.45         365       LHFVAAAWQAVEPSD       89.03       826.93       28.39         366       HFVAAAWQAVEPSD       114.93       890.53       45.34         367       FVAAAWQAVEPSDIA       201.46       1046.22       96.05         368       VAAWQAVEPSDIAA       206.83       3606.79       164.34         370       AAWQAVEPSDIAAC       308.33       5606.79       164.34         370       AAWQAVEPSDIAAC       380.43       5990.71       204.19         58       EVPVPSFGEAMAYFAM       181.1       2399.1       204.19         54		355	SGLQLGLTEALHFVA	37.17	1262.23	33.01
357       LQLGLTEALHFVAAAW       40.73       1241.73       44.3         358       QLGLTEALHFVAAAWQ       61.12       914.02       66.47         359       LGLTEALHFVAAAWQA       161.26       838.91       61.96         360       GLTEALHFVAAAWQAV       53.13       683.6       26         361       LTEALHFVAAAWQAV       53.13       683.6       26         362       TEALHFVAAAWQAVE       52.96       659.3       21.74         364       ALHFVAAAWQAVEPS       63.45       618.86       22.45         365       LHFVAAAWQAVEPSD       89.03       826.93       28.39         366       HFVAAAWQAVEPSDI       114.93       890.53       45.34         367       FVAAAWQAVEPSDIA       201.46       1046.22       96.05         368       VAAWQAVEPSDIA       201.46       1046.22       96.05         368       VAAWQAVEPSDIAACF       380.43       5990.71       204.19         538       EVPVPSFGEAMAYFA       348.08       3153.32       146.22         539       VPVPSFGEAMAYFAMV       125.36       1206.91       46.1         540       PVPSFGEAMAYFAMVK       105.57       88.358       36.01         541		356	GLQLGLTEALHFVAA	34.65	1236.62	33.48
358       QLGL FEALHFVAAAWQ       61.12       914.02       66.47         359       LGL TEALHFVAAAWQA       121.87       875.41       97.8         360       GL TEALHFVAAAWQAV       53.13       683.6       26         361       L TEALHFVAAAWQAV       52.96       659.3       21.74         363       EALHFVAAAWQAVE       52.96       603.24       21.78         364       ALHFVAAAWQAVEPS       63.45       618.86       22.45         365       LHFVAAAWQAVEPSD       89.03       826.93       28.39         366       HFVAAAWQAVEPSD       201.46       1046.22       96.05         368       VAAAWQAVEPSDIA       201.46       1046.22       96.05         368       VAAAWQAVEPSDIA       256.34       1990.64       138.31         369       AAAWQAVEPSDIAA       256.34       1990.64       138.31         369       AAAWQAVEPSDIAA       256.34       1990.64       138.31         370       AAWQAVEPSDIAA       25.36       1206.91       46.1         370       AAWQAVEPSDIAA       181.1       239.91       146.22         533       VPVPSFGEAMAYFAMVK       105.57       883.58       36.01         541		357	LQLGLTEALHFVAAA	40.73	1241.73	44.3
359         LGT FALHF WAAAWQ         121.87         87.41         97.8           360         GITEALHF WAAAWQA         161.26         383.91         61.96           361         LTEALHF WAAAWQAVE         52.96         659.3         21.74           363         EALHF WAAAWQAVE         52.96         659.3         21.74           363         EALHF WAAAWQAVE         53.13         603.24         21.78           364         ALHF WAAAWQAVEPS         63.45         618.86         22.45           365         LHF WAAAWQAVEPSD         89.03         826.93         28.39           366         HF WAAAWQAVEPSD         89.03         890.53         45.34           367         F WAAAWQAVEPSDIA         201.46         1046.22         96.05           368         WAAAWQAVEPSDIAA         256.34         1990.64         138.31           369         AAAWQAVEPSDIAAC         308.33         3606.79         164.34           370         AAWQAVEPSDIAACF         380.43         5990.71         204.19           538         EVPVPSFGEAMAYFAM         181.1         2399.13         65.53           540         PVPSFGEAMAYFAM         181.1         239.17           543         SFGEAM		358	QLGLTEALHEVAAAW	61.12	914.02	66.47
360         GLTEALTHYVAAWQAV         101.20         363.91         01.50           361         LTEALTHYVAAWQAV         53.13         683.6         26           362         TEALTHYVAAWQAVE         52.96         659.3         21.74           363         EALHFVAAAWQAVEP         57.93         603.24         21.78           364         ALHFVAAAWQAVEPS         63.45         618.86         22.45           365         LHFVAAAWQAVEPSD         89.03         826.93         28.39           366         HFVAAAWQAVEPSDI         114.93         890.53         45.34           367         FVAAAWQAVEPSDIA         201.46         1046.22         96.05           368         VAAAWQAVEPSDIA         256.34         1990.64         138.31           369         AAAWQAVEPSDIAAC         380.33         3606.79         164.34           370         AAWQAVEPSDIAACF         380.43         5990.71         204.19           538         EVPVPSFGEAMAYFAM         181.1         2399.13         65.53           540         PVPSFGEAMAYFAM         181.1         2399.17         543         546         62AMAYFAMVKR         105.57         88.58         36.01           541         VPSFGEAMAYFAMV		359		121.87	8/5.41	97.8
S01         ETEALH VAAAWQAVE         52.96         659.3         21.74           363         EALHFVAAAWQAVE         52.96         659.3         21.74           364         ALHFVAAAWQAVEPS         63.45         618.86         22.45           365         LHFVAAAWQAVEPSD         89.03         826.93         28.39           366         HFVAAAWQAVEPSD         89.03         826.93         28.39           366         HFVAAAWQAVEPSD         89.03         826.93         28.39           366         HFVAAAWQAVEPSD         89.03         826.93         28.39           366         HFVAAAWQAVEPSDIA         201.46         1046.22         96.05           368         VAAWQAVEPSDIA         201.46         1046.22         96.05           368         VAAWQAVEPSDIAA         256.34         1990.64         138.31           369         AAAWQAVEPSDIAAC         308.33         3606.79         164.34           370         AAWQAVEPSDIAAC         308.33         3605.73         146.22           539         VPVPSFGEAMAYFA         348.08         3153.32         146.22           541         VPSFGEAMAYFAMVK         105.57         883.58         36.01           542		361		53 13	683.6	26
361         EALHFVAAAWQAVEP         57.33         603.24         21.78           363         EALHFVAAAWQAVEPS         63.45         618.86         22.45           364         ALHFVAAAWQAVEPSD         89.03         826.93         28.39           366         HFVAAAWQAVEPSDI         114.93         890.53         45.34           367         FVAAAWQAVEPSDIA         201.46         1046.22         96.05           368         VAAAWQAVEPSDIA         256.34         1990.64         138.31           369         AAAWQAVEPSDIA         256.34         1990.64         138.31           369         AAAWQAVEPSDIAA         256.34         1990.64         138.31           369         AAAWQAVEPSDIAA         256.34         1990.64         138.31           369         AAAWQAVEPSDIAAC         308.33         3606.79         164.34           370         AAWQAVEPSDIAACF         380.43         5990.71         204.19           538         EVPVPSFGEAMAYFA         348.08         3153.32         146.22           539         VPVPSFGEAMAYFAM         181.1         2399.13         65.53           540         PVPSFGEAMAYFAMVK         105.57         883.58         36.01           <		362		52.06	659.3	20
364         ALHI-WAARWQAVEPS         63.45         603.24         22.45           364         ALHFVAAAWQAVEPSD         89.03         826.93         28.39           366         HFVAAAWQAVEPSD         89.03         826.93         28.39           366         HFVAAAWQAVEPSDI         114.93         890.53         45.34           367         FVAAAWQAVEPSDIA         201.46         1046.22         96.05           368         VAAAWQAVEPSDIA         256.34         1990.64         138.31           369         AAAWQAVEPSDIAAC         308.33         3606.79         164.34           370         AAWQAVEPSDIAAC         308.33         5990.71         204.19           538         EVPVPSFGEAMAYFA         348.08         3153.32         146.22           539         VPVPSFGEAMAYFA         348.08         3153.32         146.22           539         VPVPSFGEAMAYFA         348.08         3153.32         146.22           539         VPVPSFGEAMAYFA         348.08         3153.32         146.22           533         540         PVPSFGEAMAYFAMVK         105.57         883.58         36.01           541         VPSFGEAMAYFAMVKR         119.08         535.35         50.39 <td></td> <td>363</td> <td></td> <td>57.93</td> <td>603.24</td> <td>21.74</td>		363		57.93	603.24	21.74
Adv         Adv         Adv         Adv         Adv           365         LHFVAAAWQAVEPSD         89.03         826.93         28.39           366         HFVAAAWQAVEPSDI         114.93         890.53         45.34           367         FVAAAWQAVEPSDIA         201.46         1046.22         96.05           368         VAAAWQAVEPSDIAA         256.34         1990.64         138.31           369         AAAWQAVEPSDIAAC         308.33         3606.79         164.34           370         AAWQAVEPSDIAAC         308.33         3606.79         164.34           370         AAWQAVEPSDIAACF         380.43         5990.71         204.19           538         EVPVPSFGEAMAYFA         348.08         3153.32         146.22           539         VPVPSFGEAMAYFA         348.08         3153.32         146.22           539         VPVPSFGEAMAYFAM         181.1         2399.13         65.53           540         PVPSFGEAMAYFAMVK         105.57         883.58         36.01           541         VPSFGEAMAYFAMVKR         119.34         705.1         39.17           543         SFGEAMAYFAMVKRY         128.18         566.98         44.58           544		364	ALHEVAAAWOAVEPS	63.45	618.86	22.45
And Market         Data         Data         Data           366         HFVAAAWQAVEPSDI         114.93         890.53         45.34           367         FVAAAWQAVEPSDIA         201.46         1046.22         96.05           368         VAAAWQAVEPSDIAA         256.34         1990.64         138.31           369         AAAWQAVEPSDIAAC         308.33         3606.79         164.34           370         AAWQAVEPSDIAACF         380.43         5990.71         204.19           538         EVPVPSFGEAMAYFA         348.08         3153.32         146.22           539         VPVPSFGEAMAYFA         348.08         3153.32         146.22           539         VPVPSFGEAMAYFAM         181.1         2399.13         65.53           540         PVPSFGEAMAYFAMVK         105.57         883.58         36.01           541         VPSFGEAMAYFAMVKR         109.357         883.58         36.01           542         PSFGEAMAYFAMVKR         109.357         883.58         36.01           542         PSFGEAMAYFAMVKRY         128.18         566.98         44.58           544         FGEAMAYFAMVKRYL         19.08         535.35         50.39           545         GEA		365	LHEVAAAWOAVEPSD	89.03	826.93	28,39
367         FVAAAWQAVEPSDIA         201.46         1046.22         96.05           368         VAAAWQAVEPSDIAA         256.34         1990.64         138.31           369         AAAWQAVEPSDIAA         308.33         3606.79         164.34           370         AAWQAVEPSDIAAC         308.33         3606.79         164.34           370         AAWQAVEPSDIAACF         380.43         5990.71         204.19           538         EVPVPSFGEAMAYFA         348.08         3153.32         146.22           539         VPVPSFGEAMAYFA         348.08         3153.32         146.22           539         VPVPSFGEAMAYFAM         181.1         2399.13         65.53           540         PVPSFGEAMAYFAMV         125.36         1206.91         46.1           541         VPSFGEAMAYFAMVK         105.57         883.58         36.01           542         PSFGEAMAYFAMVKR         119.34         705.1         39.17           543         SFGEAMAYFAMVKRY         128.18         566.98         44.58           544         FGEAMAYFAMVKRYL         19.08         535.35         50.39           545         GEAMAYFAMVKRYLT         97.64         688.09         43.9           5		366	HFVAAAWQAVEPSDI	114.93	890.53	45.34
368         VAAAWQAVEPSDIAA         256.34         1990.64         138.31           369         AAAWQAVEPSDIAAC         308.33         3606.79         164.34           370         AAWQAVEPSDIAACF         380.43         5990.71         204.19           538         EVPVPSFGEAMAYFA         348.08         3153.32         146.22           539         VPVPSFGEAMAYFA         348.08         3153.32         146.22           539         VPVPSFGEAMAYFAM         181.1         2399.13         65.53           540         PVPSFGEAMAYFAMV         125.36         1206.91         46.1           541         VPSFGEAMAYFAMVK         105.57         883.58         36.01           542         PSFGEAMAYFAMVKR         119.34         705.1         39.17           543         SFGEAMAYFAMVKRY         128.18         566.98         44.58           544         FGEAMAYFAMVKRYL         119.08         535.35         50.39           545         GEAMAYFAMVKRYL         176.64         698.09         43.9           546         EAMAYFAMVKRYLTS         96.14         891.11         36.51           547         AMAYFAMVKRYLTS         91.74         14212.15         36.81           54		367	FVAAAWQAVEPSDIA	201.46	1046.22	96.05
369         AAAWQAVEPSDIAAC         308.33         3606.79         164.34           370         AAWQAVEPSDIAACF         380.43         5990.71         204.19           538         EVPVPSFGEAMAYFA         348.08         3153.32         146.22           539         VPVPSFGEAMAYFAM         181.1         2399.13         65.53           540         PVPSFGEAMAYFAM         181.1         2399.13         65.53           540         PVPSFGEAMAYFAMV         125.36         1206.91         46.1           541         VPSFGEAMAYFAMVK         105.57         883.58         36.01           542         PSFGEAMAYFAMVKR         119.34         705.1         39.17           543         SFGEAMAYFAMVKRY         128.18         566.98         44.58           544         FGEAMAYFAMVKRY         128.18         566.98         43.9           545         GEAMAYFAMVKRYL         19.08         535.35         50.39           546         EAMAYFAMVKRYL         19.08         535.35         50.39           545         GEAMAYFAMVKRYLT         97.64         698.09         43.9           546         EAMAYFAMVKRYLTS         96.14         891.11         36.51           547		368	VAAAWQAVEPSDIAA	256.34	1990.64	138.31
370         AAWQAVEPSDIAACF         380.43         5990.71         204.19           538         EVPVPSFGEAMAYFA         348.08         3153.32         146.22           539         VPVPSFGEAMAYFAM         181.1         2399.13         65.53           540         PVPSFGEAMAYFAM         181.1         2399.13         65.53           540         PVPSFGEAMAYFAMV         125.36         1206.91         46.1           541         VPSFGEAMAYFAMVK         105.57         883.58         36.01           542         PSFGEAMAYFAMVKR         119.34         705.1         39.17           543         SFGEAMAYFAMVKRY         128.18         566.98         44.58           544         FGEAMAYFAMVKRYL         119.08         535.35         50.39           545         GEAMAYFAMVKRYLT         97.64         698.09         43.9           546         EAMAYFAMVKRYLTS         96.14         891.11         36.51           547         AMAYFAMVKRYLTS         96.14         891.11         36.51           548         MAYFAMVKRYLTSFP         107.41         2142.15         36.81           549         AYFAMVKRYLTSFP         107.41         2142.15         36.81           549		369	AAAWQAVEPSDIAAC	308.33	3606.79	164.34
538         EVPVPSFGEAMAYFA         348.08         3153.32         146.22           539         VPVPSFGEAMAYFAM         181.1         2399.13         65.53           540         PVPSFGEAMAYFAMV         125.36         1206.91         46.1           541         VPSFGEAMAYFAMV         125.36         1206.91         46.1           541         VPSFGEAMAYFAMVK         105.57         883.58         36.01           542         PSFGEAMAYFAMVKR         119.34         705.1         39.17           543         SFGEAMAYFAMVKRY         128.18         566.98         44.58           544         FGEAMAYFAMVKRYL         119.08         535.35         50.39           545         GEAMAYFAMVKRYLT         97.64         698.09         43.9           546         EAMAYFAMVKRYLTS         96.14         891.11         36.51           547         AMAYFAMVKRYLTSF         93.24         1466.01         32.8           548         MAYFAMVKRYLTSFP         107.41         2142.15         36.81           549         AYFAMVKRYLTSFPI         46.74         1850.09         28.5           550         YFAMVKRYLTSFPID         54.2         1556.56         31.19           551		370	AAWQAVEPSDIAACF	380.43	5990.71	204.19
539         VPVPSFGEAMAYFAM         181.1         2399.13         65.53           540         PVPSFGEAMAYFAMV         125.36         1206.91         46.1           541         VPSFGEAMAYFAMVK         105.57         883.58         36.01           542         PSFGEAMAYFAMVK         105.57         883.58         39.17           543         SFGEAMAYFAMVKR         119.34         705.1         39.17           543         SFGEAMAYFAMVKRY         128.18         566.98         44.58           544         FGEAMAYFAMVKRYL         119.08         535.35         50.39           545         GEAMAYFAMVKRYLT         97.64         698.09         43.9           546         EAMAYFAMVKRYLTS         96.14         891.11         36.51           547         AMAYFAMVKRYLTSF         93.24         1466.01         32.8           548         MAYFAMVKRYLTSF         93.24         1466.01         32.8           549         AYFAMVKRYLTSF         93.24         1466.01         32.8           548         MAYFAMVKRYLTSFP         107.41         2142.15         36.81           549         AYFAMVKRYLTSFPID         54.2         1556.56         31.19           551         <		538	EVPVPSFGEAMAYFA	348.08	3153.32	146.22
540         PVPSFGEAMAYFAMVV         125.36         1206.91         46.1           541         VPSFGEAMAYFAMVK         105.57         883.58         36.01           542         PSFGEAMAYFAMVKR         119.34         705.1         39.17           543         SFGEAMAYFAMVKRY         128.18         566.98         44.58           544         FGEAMAYFAMVKRY         119.08         535.35         50.39           545         GEAMAYFAMVKRYL         119.08         535.35         50.39           545         GEAMAYFAMVKRYLT         97.64         698.09         43.9           546         EAMAYFAMVKRYLTS         96.14         891.11         36.51           547         AMAYFAMVKRYLTSF         93.24         1466.01         32.8           548         MAYFAMVKRYLTSFP         107.41         2142.15         36.81           549         AYFAMVKRYLTSFPI         46.74         1850.09         28.5           550         YFAMVKRYLTSFPID         54.2         1556.56         31.19           551         FAMVKRYLTSFPID         54.2         1556.56         39.94           552         AMVKRYLTSFPIDR         70.67         620.35         39.94           553         <		539	VPVPSFGEAMAYFAM	181.1	2399.13	65.53
541       VPSFGEAMAYFAMVK       105.57       883.58       36.01         542       PSFGEAMAYFAMVKR       119.34       705.1       39.17         543       SFGEAMAYFAMVKRY       128.18       566.98       44.58         544       FGEAMAYFAMVKRY       119.08       535.35       50.39         545       GEAMAYFAMVKRYL       119.08       535.35       50.39         546       EAMAYFAMVKRYLT       97.64       698.09       43.9         547       AMAYFAMVKRYLTS       96.14       891.11       36.51         547       AMAYFAMVKRYLTSF       93.24       1466.01       32.8         548       MAYFAMVKRYLTSFP       107.41       2142.15       36.81         549       AYFAMVKRYLTSFP       107.41       1850.09       28.5         550       YFAMVKRYLTSFPID       54.2       1556.56       31.19         551       FAMVKRYLTSFPIDD       69.94       927.84       45.09         552       AMVKRYLTSFPIDDR       70.67       620.35       39.94         553       MVKRYLTSFPIDDRV       69.52       424.57       37.95         554       VKRYLTSFPIDDRVQ       88.71       453.78       46.03         555		540	PVPSFGEAMAYFAMV	125.36	1206.91	46.1
542       PSFGEAMAYFAMVKR       119.34       705.1       39.17         543       SFGEAMAYFAMVKRY       128.18       566.98       44.58         544       FGEAMAYFAMVKRYL       119.08       535.35       50.39         545       GEAMAYFAMVKRYL       97.64       698.09       43.9         546       EAMAYFAMVKRYLTS       96.14       891.11       36.51         547       AMAYFAMVKRYLTSF       93.24       1466.01       32.8         548       MAYFAMVKRYLTSFP       107.41       2142.15       36.81         549       AYFAMVKRYLTSFP       107.41       1850.09       28.5         550       YFAMVKRYLTSFPID       54.2       1556.56       31.19         551       FAMVKRYLTSFPIDD       69.94       927.84       45.09         552       AMVKRYLTSFPIDDR       70.67       620.35       39.94         553       MVKRYLTSFPIDDRV       69.52       424.57       37.95         554       VKRYLTSFPIDDRVQ       88.71       453.78       46.03         555       KRVLTSFPIDDRVQS       156.74       524.42       77.53		541	VPSFGEAMAYFAMVK	105.57	883.58	36.01
543         SFGEAMAYFAMVKRY         128.18         566.98         44.58           544         FGEAMAYFAMVKRYL         119.08         535.35         50.39           545         GEAMAYFAMVKRYL         19.08         535.35         50.39           545         GEAMAYFAMVKRYLT         97.64         698.09         43.9           546         EAMAYFAMVKRYLTS         96.14         891.11         36.51           547         AMAYFAMVKRYLTSF         93.24         1466.01         32.8           548         MAYFAMVKRYLTSFP         107.41         2142.15         36.81           549         AYFAMVKRYLTSFPI         46.74         1850.09         28.5           550         YFAMVKRYLTSFPID         54.2         1556.56         31.19           551         FAMVKRYLTSFPIDD         69.94         927.84         45.09           552         AMVKRYLTSFPIDDR         70.67         620.35         39.94           553         MVKRYLTSFPIDDRV         69.52         424.57         37.95           554         VKRYLTSFPIDDRVQ         88.71         453.78         46.03           555         KRVLTSFPIDDRVQS         156.74         524.42         77.53		542	PSFGEAMAYFAMVKR	119.34	705.1	39.17
544         FGEAMAYFAMVKRYL         119.08         535.35         50.39           545         GEAMAYFAMVKRYLT         97.64         698.09         43.9           546         EAMAYFAMVKRYLTS         96.14         891.11         36.51           547         AMAYFAMVKRYLTS         93.24         1466.01         32.8           548         MAYFAMVKRYLTSFP         107.41         2142.15         36.81           549         AYFAMVKRYLTSFP         107.41         1850.09         28.5           550         YFAMVKRYLTSFPID         54.2         1556.56         31.19           551         FAMVKRYLTSFPIDD         69.94         927.84         45.09           552         AMVKRYLTSFPIDDR         70.67         620.35         39.94           553         MVKRYLTSFPIDDRV         69.52         424.57         37.95           554         VKRYLTSFPIDDRVQ         88.71         453.78         46.03           555         KRYLTSFPIDDRVQS         156.74         524.42         77.53		543	SFGEAMAYFAMVKRY	128.18	566.98	44.58
545         GEAMAYFAMVKRYLT         97.64         698.09         43.9           546         EAMAYFAMVKRYLTS         96.14         891.11         36.51           547         AMAYFAMVKRYLTS         93.24         1466.01         32.8           548         MAYFAMVKRYLTSFP         107.41         2142.15         36.81           549         AYFAMVKRYLTSFPI         46.74         1850.09         28.5           550         YFAMVKRYLTSFPID         54.2         1556.56         31.19           551         FAMVKRYLTSFPID         69.94         927.84         45.09           552         AMVKRYLTSFPIDDR         70.67         620.35         39.94           553         MVKRYLTSFPIDDRV         69.52         424.57         37.95           554         VKRYLTSFPIDDRVQ         88.71         453.78         46.03           555         KRYLTSFPIDDRVQ         156.74         524.42         77.53		544	FGEAMAYFAMVKRYL	119.08	535.35	50.39
546         EAMAYFAMVKRYLIS         96.14         891.11         36.51           547         AMAYFAMVKRYLISF         93.24         1466.01         32.8           548         MAYFAMVKRYLTSFP         107.41         2142.15         36.81           549         AYFAMVKRYLTSFPI         46.74         1850.09         28.5           550         YFAMVKRYLTSFPID         54.2         1556.56         31.19           551         FAMVKRYLTSFPIDD         69.94         927.84         45.09           552         AMVKRYLTSFPIDDR         70.67         620.35         39.94           553         MVKRYLTSFPIDDRV         69.52         424.57         37.95           554         VKRYLTSFPIDDRVQ         88.71         453.78         46.03           555         KRVLTSFPIDDRVQS         156.74         574.42         77.53		545	GEAMAYFAMVKRYLT	97.64	698.09	43.9
547         AMAYFAMVKRYLISF         93.24         1466.01         32.8           548         MAYFAMVKRYLISFP         107.41         2142.15         36.81           549         AYFAMVKRYLISFPI         46.74         1850.09         28.5           550         YFAMVKRYLISFPID         54.2         1556.56         31.19           551         FAMVKRYLISFPIDD         69.94         927.84         45.09           552         AMVKRYLISFPIDDR         70.67         620.35         39.94           553         MVKRYLISFPIDDRV         69.52         424.57         37.95           554         VKRYLISFPIDDRVQ         88.71         453.78         46.03           555         KRVLTSFPIDDRVQS         156.74         574.42         77.53		546	LAMAYFAMVKRYLTS	96.14	891.11	36.51
548         IMATFAMIVKRYLTSFP         107.41         2142.15         36.81           549         AYFAMVKRYLTSFPI         46.74         1850.09         28.5           550         YFAMVKRYLTSFPID         54.2         1556.56         31.19           551         FAMVKRYLTSFPID         69.94         927.84         45.09           552         AMVKRYLTSFPIDDR         70.67         620.35         39.94           553         MVKRYLTSFPIDDRV         69.52         424.57         37.95           554         VKRYLTSFPIDDRVQ         88.71         453.78         46.03           555         KRYLTSFPIDDRVQS         156.74         524.42         77.53		547	AMAYFAMVKRYLTSF	93.24	1466.01	32.8
349         ATFANVIVNETLISFPI         40.74         1850.09         28.5           550         YFAMVKRYLTSFPID         54.2         1556.56         31.19           551         FAMVKRYLTSFPIDD         69.94         927.84         45.09           552         AMVKRYLTSFPIDDR         70.67         620.35         39.94           553         MVKRYLTSFPIDDRV         69.52         424.57         37.95           554         VKRYLTSFPIDDRVQ         88.71         453.78         46.03           555         KRVLTSFPIDDRVQS         156.74         524.42         77.53		548		107.41	2142.15	36.81
530         TEAMVKRYLTSFPID         54.2         1556.56         31.19           551         FAMVKRYLTSFPIDD         69.94         927.84         45.09           552         AMVKRYLTSFPIDDR         70.67         620.35         39.94           553         MVKRYLTSFPIDDRV         69.52         424.57         37.95           554         VKRYLTSFPIDDRVQ         88.71         453.78         46.03           555         KRVLTSFPIDDRVQS         156.74         524.42         77.53		549		46.74	1820.09	28.5
551         FAMILISFIED         65.34         927.84         45.09           552         AMVKRYLTSFPIDDR         70.67         620.35         39.94           553         MVKRYLTSFPIDDRV         69.52         424.57         37.95           554         VKRYLTSFPIDDRVQ         88.71         453.78         46.03           555         KRVLTSFPIDDRVQS         156.74         524.42         77.53		550		54.2	1556.56	31.19
552         AWVENTLTSFFIDDR         70.07         620.35         39.94           553         MVKRYLTSFPIDDRV         69.52         424.57         37.95           554         VKRYLTSFPIDDRVQ         88.71         453.78         46.03           555         KRYLTSFPIDDRVQS         156.74         574.42         77.53		221		09.94 70.67	927.84	45.09
553         WKRYLTSFPIDDRVQ         88.71         453.78         46.03           555         KRYLTSFPIDDRVQS         156.74         574.42         77.53		552		/U.D/	020.35 101 E7	23.94 27 OE
555 KRYLTSFPIDDRVOS 156.74 524.42 77.53		550		09.52 88 71	424.37	46 NR
		555	KRYLTSFPIDDRVOS	156.74	524.42	77.53

# Table S11. (B) Predicted immunodominant peptides in Centromere Protein B (continued). (LD=linkage disequilibrium)

Table S12. Bioinformatically derived immunodominant peptides and homologous viral protein identification. (A) Peptide sequences from microbial proteins homologous to topoisomerase I immunodominant peptide; (B) Peptide sequences from microbial proteins homologous to fibrillarin immunodominant peptide; and (C) Peptide sequences from microbial proteins homologous to CENPA immunodominant peptide.

Α			Topoisomerase I																
	Source	Protein	NCBI Accession	E-Value	% Homology							Sequ	uen	ce					
	Human	Topoisomerase I	AAL10791	-	100	R	Q	R	Α	۷	Α	L	Y	F	1	рκ	L	Α	L
	Hokovirus HKV1	DNA topoisomerase IB	ARF10236.1	0.000003	93	R	Q	T	А	۷	А	L	Y	F		D K	L	А	L
	Megavirus chiliensis	DNA topoisomerase 1b	YP_004894910.1	0.003	80	R	Q	T	А	Т	А	L	Y	F		D K	F	А	L
	Megavirus vitis	DNA topoisomerase 1b	AVL94169.1	0.003	80	R	Q	T	А	Т	А	L	Y	F	I I	D K	F	А	L
	Powai lake megavirus	Hypothetical protein	ANB50940.1	0.003	80	R	Q	T	А	Т	А	L	Y	F	I I	D K	F	А	L
	Klosneuvirus KNV1	DNA topoisomerase IB	ARF11421.1	0.003	80	R	Q	Т	А	т	А	L	Y	F	I I	D K	F	Α	L
	Catovirus CTV1	DNA topoisomerase IB	ARF08721.1	0.013	80	R	Q	Т	А	т	А	L	Y	F		р к	F	А	L

В

		Fibrillarin															
Source	Protein	NCBI Accession	E-Value	% Homology						Sec	lne	nce					
Human	Fibrillarin	EAW56925	-	100	G	R	DL	Т	Ν	L	Α	К	К	R	ΤN	I	Т
Acanthamoeba polyphaga moumouvirus	Cell division cycle 123 protein	YP_007354044.1	0.004	65			DL	Т	Ν	L	А	К	К	•	-N N	- 1	Т
Saudi moumouvirus	Cell division cycle	AQN67932	0.004	65			DL	1	Ν	L	А	К	К	I ·	N <mark>N</mark>	- 1	- I
Moumouvirus Monve	Hypothetical protein mv_R1060	AEX63262	0.004	65			DL	Т	Ν	L	А	К	К		-N N	I	1

С

		CENPA															
Source	Protein	NCBI Accession	E-Value	% Homology						Sec	que	nce					
Human	Centromere protein A	AAH02703	-	100	L	Q	E a	A A	È	Α	F	L	v	н	L	F	E D
Dishui lake phycodnavirus 1	Hypothetical protein DSLPV1_013	YP_009465730.1	0.01	80	L	Q	E	A A	E	А	Υ	L	т	S	L	F	E D

Exact match Conservative amino acid substitution Table S13. (A) Fifty 15-mer Random sequence peptides generated using ExPASy compared with viral sequence database for homology.

Peptide Sequence         Viruses with E-value <1		Number of Results		Results with E-value <1	
VKFIQNSEALDROFT         0         0         0           EIVTSAPXNGGFLRD         0         0         0           EIVTSAPXNGGFLRD         0         0         0           EINSTPGQAVKLFRD         0         0         0           FRASCEKRINTLVQD         0         0         0           FRASCEKRINTLVQD         0         0         0           STRARGEKPNVLFBQ         0         0         0           STARGEKPNVLFBQ         0         0         0           VQGRSDLIENTAPK         0         0         0           STARGEKPNVLFBQ         0         0         0           RUGAPEONVLFBQ         0         0         0	Peptide Sequence	Viruses with E-value <1	Mimiviridae with E-value <10	Virus	E-Value
EIVTSAPKNOGFLRD         0         0           SQLPARDTGKEVINF         0         0           SQLPARDTGKEVINF         0         0           EINSTPGGAVKIFRD         0         0           TKRLPSNEAFIDGQV         0         0           SRVPAKVFRD         0         0           SRVPAKVFLQD         0         0           VQGRSDLENTFAPK         0         0           VQGRSDLENTFAPK         0         0           SITARGEKPNVFEGIN         0         0           ADLSTRQKIGNVFE         1         0         Turkeypox virus           GLIANTRAPEGSIN         0         0         1           FUXGRARTPGSIN         0         0         1           VMAEPPSKETING         0         0         1           ROMORPAVESTIG         0         0         1           RVSUBARCHTROR         0         0         1           RVSUBARCHTOR         0         0         1           RVSUBARCHTOR	VKFIONSEALDRGPT	0	0		
SQLPARDTGKEVINF         0         0           EINSTPGQAVKLFRD         0         0           FPASGEKRINTIVQD         0         0           FPASGEKRINTIVQD         0         0           SITARQEKRINTIVQD         0         0           SITARQEKRINTERPR         0         0           VQGRSDLEINTAR         0         0           RUAGAPEGNULDTFSR         0         0           ADISTFRQXIGARDN         0         0           IVANEDPSGKFTLRQ         0         0           ROBLPQETRIFRSNAV         0         0           ROBLPGETRENAV         0         0           ROBLPARESETERD         0         <	EIVTSAPKNOGFLRD	0	0		
EINSTPGQAVKLFRD         0         0           TKRLPSNEAFIDGQV         0         0           TKRLPSNEAFIDGQV         0         0           KLGDPNSIATFVERQ         0         0           SITARQEKPNVLFDG         0         0           SITARQEKPNVLFDG         0         0           SITARQEKPNVLFDG         0         0           SITARQEKPNVLFDG         0         0           KIGAPEQNVLDTFSR         0         0           ADISTFRQUGNVFE         1         0         Turkeypox virus           GELMNTRAPEQVIDS         0         0         I           VARADEPSKERTUR         0         0         I           VANEDPSKERTUR         0         0         I           NANDEVRSKERTR         0         0         I           RUNDARAKESFTILG         0         0         I           RUNSEX	SOLPARDTGKEVINF	0	0		
TKRLPSNEAFIDGQV         0         0           FPASGEKRINTIVQD         0         0         0           FRASGEKRINTIVQD         0         0         0           SRVPAKDFLQNGIET         0         0         0           SITARQEKPNVLPGG         0         0         0           SITARQEKPNVLPGG         0         0         0           OKGRSDIENTFAPK         0         0         0           ADLSTFRQXIGNVPE         1         0         Turkeypox virus         0.61           FGLKITRAPEQVIDS         0         0         0         0         0           IVANEDPSGKFTLRQ         0         0         0         0         0         0           IVANEDPSGKFTLRQ         0         0         0         0         0         0         0           IVANEDPSGKFTLRA         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0	FINSTPGOAVKIERD	0	0		
FPASGEKRINTIVQD         0         0           KLGDPNSIATEVERQ         0         0           SVPARDE/LONGIET         0         0           SVTARDE/LONGIET         0         0           SVPARDE/LONGIET         0         0           VQGRSDUENTFAPK         0         0           DXVEQATRPSSIN         0         0           ALDSTRRQKIGNPE         1         0         Turkeypox virus           ADJSTRRQKIGNPE         1         0         Turkeypox virus           ADJSTRRQKIGNPE         0         0         1           VANEDPSGKTLRQ         0         0         1           IAVANEDPSGKTLRQ         0         0         1           NANEDPSGKTLRQ         0         0         1           NANEDPSGKETTIQ         0         0         1           NUSLPKGAETFORD         0         0         1           ORKENDGATVLPKFS         1         0         Streptomyces phage         0.93           NALDWRISCERFTIG         0         0         1         1         1           ONEVEPAVESTKLAP         0         0         1         1         1           QUFVERINARDOLINE         0         0	TKRI PSNEAEIDGOV	0	0		
KLSDPNSIATEVERQ         0         0           SRVPAXDFLQNGIET         0         0           SRVPAXDFLQNGIET         0         0           VQGRSDLIENTFAPK         0         0           DKVEQATRFPGSLNI         0         0           ADLSTFRQKIGNVPE         1         0         Turkeypox virus           GLIKNTRAPEQVIDS         0         0           IEVPSTKQLGARDN         0         0           IEVPSTKQLGARDN         0         0           IEVPSTKQLGARDN         0         0           IEVPSTKQLGARDN         0         0           NADLSPSGKFTIRQ         0         0           RNRDAPCSDELIGF         0         0           NUSLPGAETFQRD         0         0           NMSLPKGAETFQRD         0         0           NMSLPKGAETFQRD         0         0           NALDVRISQEKPTFG         0         0           SQPPITRLEINKVDGA         0         0           QMFEVGATVLRFS         1         0         Streptomyces phage           QJQGRVACSPTITL         0         0         1           QMFEVGARQULV         0         0         1           QMFEVGAGUNK         2	FPASGEKRINTI VOD	0	0		
Instruction         D         D           SITPAROFLONGIET         0         0           SITARQERPNUEDG         0         0           VQGRSDIENTFAPK         0         0           DKVEQATRFPGSLNI         0         0           NGAPEQNVLDTSR         0         0           ADISTFRQKIGNVPE         1         0         Turkeypox virus         0.61           FGLKINTRAPEQVIDS         0         0         0         0         0           IEFVPSTKQLGARDN         0         0         0         0         0         0           NANEDPSCKFTLRQ         0         0         0         0         0         0         0           NANAPOPSCKFTLRQ         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0 <t< td=""><td>KIGDPNSIATEVERO</td><td>0</td><td>0</td><td></td><td></td></t<>	KIGDPNSIATEVERO	0	0		
STRAQER/NVLPDG         0         0           VQGRSDLIENTFAPK         0         0           KIGAPEQNVLDTFSR         0         0           ADLSTFRQKIGNVPE         1         0         Turkeypox virus         0.61           FGLKNTRAPEQVIDS         0         0         1         0         1         0           FGLKNTRAPEQVIDS         0         0         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         1         0         1         1         0         1         1         0         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1 <td></td> <td>0</td> <td>0</td> <td></td> <td></td>		0	0		
VQGRSDLIENTFAPK         0         0           DKVEQATRFPSSINI         0         0           AUSTFRQKIGNVPE         1         0         Turkeypox virus         0.61           ADLSTFRQKIGNVPE         1         0         Turkeypox virus         0.61           FGLKNTRAPEQVIDS         0         0         Intrevention         0.61           FGLKNTRAPEQVIDS         0         0         Intrevention         0.61           KORAPCIGAURGARAN         0         0         Intrevention         0.61           KNRAVPTGSDELIGF         0         0         Intrevention         0.61           KNRAVPTGSDELIGF         0         0         Intrevention         0.61           VISLPKGARETCRRD         0         0         Intrevention         0.61           NUSLPKGARETCRRD         0         0         Intrevention         0.93           NALDVRISGEKPTFG         0         0         Intrevention         0.93           NALDVRISGEKPTFG         0         0         Intrevention         0.66           EINSUPGERENFQR         0         0         Intrevention         0.66           ELANKTDSPRQGVI         1         0         Hepacivirus         0.66	SITAROFKPNVLFDG	0	0		
DKVEQATRPRGSLNI         D         D           KIGAPEQNVLDTFSR         0         0         0           FGLKNTRAPEQVIDS         0         0         0           FGLKNTRAPEQVIDS         0         0         0           IEFVPSTKQLGARDN         0         0         0           IVANEDPSGKPTIAQ         0         0         0           KNGAPKFIAQ         0         0         0           KNAVPTCSDELIGF         0         0         0           RUDDNAPGKRSET         0         0         0           GNEIGAT/URKFS         1         0         Streptomyces phage         0.93           NALDVRISQEKPTFG         0         0         0         0         0           QANELOKATURKFS         1         0         Streptomyces phage         0.93           NALDVRISQEKPTFG         0         0         0         0         0           QANELOKATURKS         1         0         Hepacivirus F         0         0           QUGVERAVEXDLYKDG         0         0         0         0         0         0           QUGVERAVEXDLYKDA         0         0         0         0         0         0         0	VOGRSDI JENTEAPK	0	0		
NIGAPEQNVLDTFSR         0         0           ADLSTRQKIGNVPE         1         0         Turkeypox virus         0.61           FGLKNTRAPEQVIDS         0         0         0         0           IEYPSTKQLGARDN         0         0         0         0           IVANEDPSGKFTLRQ         0         0         0         0           IVANEDPSGKFTLQ         0         0         0         0           KNRAVPTQSDELIGF         0         0         0         0           NUSLPKGAETFQRD         0         0         0         0           NUSLPKGAETFQRD         0         0         0         0           QRNEIDGATVLPKFS         1         0         Streptomyces phage         0.93           NALDVRISQEKPTFG         0         0         0         0         0           QPTITLENKVOGA         0         0         0         0         0         0         0           QUYGFNKAEPDLTRS         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0		0	0		
Instruction         O         O         Turkeypox virus         0.61           FGLKNTRAPEQVIDS         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0	KIGAPEONVIDTESR	0	0		
ADJENTIQUE         2         0         TORKYPOX MOD         000           IEFVPSTKQLGARDN         0         0         0         0         0           IVANEDPSGKFLRQ         0         0         0         0         0           IVANEDPSGKFLRQ         0         0         0         0         0           NQDPSGKFLRQ         0         0         0         0         0           NQDRPAVESFILG         0         0         0         0         0           NUSLPKGAETFORD         0         0         0         0         0         0           SQFPITRLENKVDGA         0         0         0         0         0         0         0           SQFPITRLENKVDGA         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0 <t< td=""><td></td><td>1</td><td>0</td><td>Turkeypox virus</td><td>0.61</td></t<>		1	0	Turkeypox virus	0.61
ICKNITUS         D         D           IEVPSTKQLGARDN         0         0         0           IVANEDPSGKFTLRQ         0         0         0           KNRAVPTOSDELIGF         0         0         0           NQDKPRAVESFTLG         0         0         0           NVISLPKGASETTORD         0         0         0           QRNEIDGATVLPKFS         1         0         Streptomyces phage         0.93           NALDVRISQEKPTFG         0         0         0         0         0           QRNEIDGATVLPKFS         1         0         Streptomyces phage         0.93           SQFPITRLENKVDGA         0         0         0         0         0           QRNEIDGATVLPKFS         1         0         O         0         0           QRNEIDGATVLPKFS         1         0         0         0         0           QREVRIDGATILENKVDGA         0         0         0         0         0         0           QVGFINKAEPDLTRS         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0 <td>EGLKNTRAPEOVIDS</td> <td>0</td> <td>0</td> <td></td> <td>0.01</td>	EGLKNTRAPEOVIDS	0	0		0.01
ID: TO TRADITION         D         D           IVANEDPSGKFTLRQ         0         0         0           KNRAVPTQSDELIGF         0         0         0           KNRAVPTQSDELIGF         0         0         0           RVIDUDAPGKRSET         0         0         0           QNVERPKAETFQRD         0         0         0           QNNEIDGATVLPKFS         1         0         Streptomyces phage         0.93           NALDVRISQEKPTFG         0         0         0         0         0           QNEVIDGATVLPKFS         1         0         Streptomyces phage         0.93           NALDVRISQEKPTFG         0         0         0         0         0           QNEVRDATVLPKFS         1         0         0         0         0           LTKSVIDGENPQLVE         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0		0	0		
INNELD JOK         0         0           INGIPQEFITRSNAV         0         0           INGLPCEFITRSNAV         0         0           NRAVPTQSDELIGF         0         0           NURAVPTQSDELIGF         0         0           NVISLPKGAETFQRD         0         0           NVISLPKGAETFQRD         0         0           QRNEIDGATVLPKFS         1         0         Streptomyces phage           NALDVRISQEKPTFG         0         0         1           QPIPTIELINKVDGA         0         0         1           QMFEVROSTKIAP         0         0         1           QVGFNKAEPDLTRS         0         0         1           QVGFNKAEPDLTRS         0         0         1           QVGFNKAEPDLTRS         0         0         1           QVGFNKAEPDLTRS         0         0         1           QUGFNKAEPDLTRS         0         0         1           QUSGFNKAP         0         0         1           GRNEUTSPRIQADVK         2         0         Rodent hepacivirus         0.66           ELANKTFDSPRQGVI         1         0         Hepativirus         0.66         1           <		0	0		
NDLOE QLTHINAVE         O         O         O           NRRAVPTQSDELIGF         0         0         0           NVISLPKGAETFQRD         0         0         0           NVISLPKGAETFQRD         0         0         0           RNEIDGATVLPKFS         1         0         Streptomyces phage         0.93           NALDVRISQEKPTFG         0         0         0         0           QRNEIDGATVLPKFS         1         0         Streptomyces phage         0.93           NALDVRISQEKPTFG         0         0         0         0           ITRITNKDGASPQLVE         0         0         0         0           ICKSVIDGPENFQAR         0         0         0         0           IQVGFNKAEPDLTRS         0         0         0         0           IQVGFNKAEPDLTGK         0         0         0         0.66           ELANKTFDSPRQGVI         1         0         Hepatitis C virus subtype 1b         0.33           QFSEIPRTVANDLGK         0         0         1         0         1           GRKDFEIQSLVAN         0         0         1         0         1           GRKDFEIQSLVAN         0         0         1		0	0		
NNMY FUGUE         0         0           RODKRAVESFITLG         0         0           FLVQIDNAPGKRSET         0         0           QRNEUDGATVLPKFS         1         0           SQPERAVLPKFS         1         0           SQPETRLENKVDGA         0         0           SQPETRLENKVDGA         0         0           SQPETRLENKVDGA         0         0           CINSVIDOPENÇAR         0         0           QNFEVRDGLSTKIAP         0         0           QVGFNKAEPDLTRS         0         0           QVGFNKAEPDLTRS         0         0           EGNFLTSPRIQADVK         2         0         Hepativirus F Rodent hepacivirus           ELANKTEDSPRQGVI         1         0         Hepatitis C virus subtype 1b         0.33           QFSEIPRTVANDLGK         0         0         1         0         Hepatitis C virus subtype 1b         0.33           GRSGRALVKITPNEDQ         0         0         1         1         0         1         0         1           PAEINVTFSDGQLR         0         0         1         1         0         1         1         0         1         1         1         1 <td< td=""><td></td><td>0</td><td>0</td><td></td><td></td></td<>		0	0		
INCORTANUES         0         0           IEVACIDNAPGKRSET         0         0           NVISLPKGAETFQRD         0         0           QRNEIDGATVLPKFS         1         0         Streptomyces phage         0.93           NALDVRISQEKPTFG         0         0         0         0         0           SQFPITRLENKVDGA         0         0         0         0         0           FRITNKDGSAPQLVE         0         0         0         0         0           QNFEVRDGLSTKIAP         0         0         0         0         0           QVFEVRDGLSTKIAP         0         0         0         0         0         0           UQVSFNKAEPDLTRS         0         0         0         0         0         0.66           ELANKTEDSPRQGVI         1         0         Hepacivirus F         0.66         0.33         0         0         0         0         0         0         0.66         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0 <td></td> <td>0</td> <td>0</td> <td></td> <td></td>		0	0		
PLVGL00ARGETFQRD       0       0         QRNEIDGATVLPKFS       1       0       Streptomyces phage       0.93         NALDVRISQEKPTFG       0       0       0       0         SQFPITRLENKVDGA       0       0       0       0         SQFPITRLENKVDGA       0       0       0       0         ITKSVIDGPENFQAR       0       0       0       0         ITKSVIDGPENFQAR       0       0       0       0         IQVGFNKAEPDLTRS       0       0       0       0.66         ELANKTFDSPRQGVI       1       0       Hepacivirus       0.66         ELANKTFDSPRQGVI       1       0       Hepatitis C virus subtype 1b       0.33         QFSEIPRTVANDLGK       0       0       0       0         AEIPTDLFQGKNSVR       0       0       0       0         GGRAUKITPNEDQ       0       0       0       0         AEIPTDLFQGKISDLRP       0       0       0       0       0         GFRAUKITPNEDQ       0       0       0       0       0       0         AAEIPTDLFQGKINSVR       0       0       0       0       0       0       0       0		0	0		
INVESTIGATURE         0         0         0         0           QRNEIDGATULPKFS         1         0         Streptomyces phage         0.93           NALDVRISQEKPTFG         0         0         Image: Construction of the streptomyces of the streptomyce		0	0		
CANALIDAR VERYS         1         0         Subplicities prage         0.53           NALDVRISQENTFG         0         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         1         0         1         1         0         1         1         0         1         1         0         1         1         0         1         1         0         1         1         0         1         1         0         1         1         0         1         1         0         1         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         1         0         1         1         0         1         1         0         1         1         0         1         1         0         1         1         1		0	0	Strontomycoc phago	0.02
NMLDVNISQENTIPS         0         0         0           GQFPITRLENKVDGA         0         0         0           FRITNKDGSAPQLVE         0         0         0           LTKSVIDGPENFQAR         0         0         0           QNFEVRDGLSTKIAP         0         0         0           IQVGFNKAEPDLTRS         0         0         0           EGNFLTSPRIQADVK         2         0         Hepacivirus F Rodent hepacivirus         0.66           ELANKTFDSPRQGVI         1         0         Hepatitis C virus subtype 1b         0.33           QFSEIPRTVANDLGK         0         0         1         0         Hepatitis C virus subtype 1b         0.33           QFSEIPRTVANDLGK         0         0         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         1         0         1         1         0         1         1         0         1         1         1         1 </td <td></td> <td>1</td> <td>0</td> <td>Streptomyces phage</td> <td>0.95</td>		1	0	Streptomyces phage	0.95
SUPFIRIENT/ODA         0         0         0           ERITNKDGSAPQLVE         0         0         0           LTKSVIDGPENFQAR         0         0         0           QNFEVRDGLSTKIAP         0         0         0           IQVGFNKAEPDLTRS         0         0         0           EGNFLTSPRIQADVK         2         0         Rodent hepacivirus         0.66           ELANKTFDSPRQGVI         1         0         Hepatitis C virus subtype 1b         0.33           QFSEIPRTVANDLGK         0         0         0         1         0           AEIPTDLFQGKNSVR         0         0         0         1         0         Hepatitis C virus subtype 1b         0.33           QFSEIPRTVANDLGK         0         0         0         1         0         1         0         1         0         1         0         1         0         1         0         1         0         3         3         1         0         1         0         1         0         1         0         1         0         1         0         1         0         1         1         0         1         1         0         1         1 <td< td=""><td></td><td>0</td><td>0</td><td></td><td></td></td<>		0	0		
FNTINDOSAPQUYE         0         0         0           LTKSVIDGPENFQAR         0         0         0           IQVFEVRDGLSTKIAP         0         0         0           IQVGFNKAEPDLTRS         0         0         0           EGNFLTSPRIQADVK         2         0         Hepacivirus F Rodent hepacivirus         0.66           ELANKTFDSPRQGVI         1         0         Hepatitis C virus subtype 1b         0.33           QFSEIPRTVANDLGK         0         0         0         1           GRKDFEIQPSLVAN         0         0         1         0           TGRKDFEIQPSLVAN         0         0         1         0           SGFRALVKITPNEDQ         0         0         1         1         0           SGFRALVKITPNEDQ         0         0         1         1         1         1         1         1         1         1         1         0         1         1         0         1         1         0         1         1         0         1         1         0         1         1         1         1         1         1         1         1         1         1         1         1         1         1		0	0		
Litsvilderendark         0         0         0           QNFEVRDGLSTKIAP         0         0         0           EGNFLTSPRIQADVK         2         0         Hepacivirus F Rodent hepacivirus         0.66           ELANKTFDSPRQGVI         1         0         Hepatitis C virus subtype 1b         0.33           QFSEIPRTVANDLGK         0         0         1         0         Hepatitis C virus subtype 1b         0.33           QFSEIPRTVANDLGK         0         0         1         0         1         0         1         0.66           ELANKTFDSPRQGVI         1         0         0         1         0.33         1         0.66         1         0.66         1         0.66         1         0.66         1         0.33         1         0.66         1         0.66         1         0.66         1         0.66         1         0.66         1         1         0.66         1         0.66         1         1         0.66         1         0.66         1         1         0.66         1         1         0.66         1         1         0.66         1         1         0.66         1         1         0.66         1         1		0	0		
Liver VRDGISTRAP         0         0         0           IQVGFNKAEPDLTRS         0         0         Hepacivirus F         0.66           EGNFLTSPRIQADVK         2         0         Hepacivirus F         0.66           ELANKTFDSPRQGVI         1         0         Hepatitis C virus subtype 1b         0.33           QFSEIPRTVANDLGK         0         0         0         1           AEIPTDLFQGKNSVR         0         0         1         1           YALIPCQKNSVR         0         0         1         1         1         1         1         0         1         1         0         1         1         0         1         1         0         1         1         0         1         1         0         1         1         0         1         0         1         1         0         1         1         0         1         1         0         1         1         0         1         1         0         1         1         0         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1 <td></td> <td>0</td> <td>0</td> <td></td> <td></td>		0	0		
Individent ACP DETRS         0         0         0         Hepacivirus F Rodent hepacivirus         0.66           ELANKTFDSPRQGVI         1         0         Hepatitis C virus subtype 1b         0.33           QFSEIPRTVANDLGK         0         0         0         0           AEIPTDLFQGKNSVR         0         0         0         0           AEIPTDLFQGKNSVR         0         0         0         0           TGRKDFEIQPSLVAN         0         0         0         0           SGFRALVKITPNEQQ         0         0         0         0           SGFRALVKITPNEQQ         0         0         0         0           TANISQFRGLDKEPV         0         0         0         0           TAVNREQGFISDLKP         0         0         0         0           AVEDGPQFKSLIRNT         0         0         0         0           DNEVKIPSTRAQGIF         0         0         0         0         0           SQVDPTRFLNIGEAK         0         0         0         0         0         0           QKFAPISLTGERVDN         0         0         0         0         0         0         0           QKFAPISLTGERVDN		0	0		
EGNFLTSPRIQADVK         2         0         Inepativitus P Rodent hepacivirus         0.66           ELANKTFDSPRQGVI         1         0         Hepatitis C virus subtype 1b         0.33           QFSEIPRTVANDLGK         0         0         0         0           AEIPTDLFQGKNSVR         0         0         0         0           TGRKDFEIQPSLVAN         0         0         0         0           PAEINKVTFSDGQLR         0         0         0         0           SGFRALVKITPNEDQ         0         0         0         0           TNAISQFRGLDKEPV         0         0         0         0           TAVNREQGFISDLKP         0         0         0         0           DVEXKIPSTRAQGLF         0         0         0         0           VGPFTLENADKQISR         0         0         0         0           QKFAPISLTGERVDN         0         0         0         0           NGFLTVEPIDSRQKA         0 <td>IQVGFINKAEPDLIKS</td> <td>0</td> <td>0</td> <td>Honocivirus F</td> <td></td>	IQVGFINKAEPDLIKS	0	0	Honocivirus F	
ELANKTFDSPRQGVI10Hepatitis C virus subtype 1b0.33QFSEIPRTVANDLGK000AEIPTDLFQGKNSVR000TGRKDFEIQPSLVAN000PAEINKVTFSDGQLR000SGFRALVKITPNEDQ000SGFRALVKITPNEDQ000TNAISQFRGLDKEPV000TAVNREQGFISDLKP000DQKAGELVTRPFNIS000DVGPFTLENADKQISR000SQVDPTRFLNIGEAK000QKFAPISLTGERVDN000RADITVQSEKFNGLP000RADITVQSEKFNGLP000FKDTGANIQSRLV000FKDTGANIQSPRLV000FKDTGANIQSPRLV000TRADITVQSEKFNGLP000TRADITVQSEKFNGLP000TRADITVQSEKFNGLP000TRADITVQSEKFNGLP000TRADITVQSEKFNGLP000TRADITVQSEKFNGLP000TRADITVQSEKFNGLP000TRADITVQSEKFNGLP000TRADITVQSEKFNGLP000TRADITVQSEKFNGLP000TRADITVQSEKFNGLP000TRADITVARS000TRADITVARS000TRADITVARS00	EGNFLTSPRIQADVK	2	0	Rodent benacivirus	0.66
Description         Description <thdescription< th=""> <thdescription< th=""></thdescription<></thdescription<>	FLANKTEDSPROGVL	1	0	Hepatitis C virus subtype 1b	0.33
AEIPTDLFQGKNSVR         0         0         0           AEIPTDLFQGKNSVR         0         0         0           TGRRDFEIQPSLVAN         0         0         0           PAEINKVTFSDGQLR         0         0         0           SGFRALVKITPNEDQ         0         0         0           SGFRALVKITPNEDQ         0         0         0           ERNDATFILKGPQVS         0         0         0           TAVINEQGFISDLKP         0         0         0           TAVNREQGFISDLKP         0         0         0           AVEDGPQFKSLIRNT         0         0         0           DQKAGELVTRPFNIS         0         0         0           DNEVKIPSTRAQGLF         0         0         0           VGPFTLENADKQISR         0         0         0           SQVDPTRFLNIGEAK         0         0         0           ASGNLPTVQDKIFER         0         0         0           QKFAPISLTGERVDN         0         0         0           RADITVQSEKFNGLP         0         0         0           FLGDETPQNRIVAKS         0         0         0           EFKDTGANIQSPRLV         0         0		0	0		0.00
TGR HOL QONOM         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         <		0	0		
PAEINKVTFSDGQLR         0         0           SGFRALVKITPNEDQ         0         0           SGFRALVKITPNEDQ         0         0           ERNDATFILKGPQVS         0         0           TNAISQFRGLDKEPV         0         0           TAVNREQGFISDLKP         0         0           AVEDGPQFKSLIRNT         0         0           DQKAGELVTRPFNIS         0         0           DVGPFTLENADKQISR         0         0           VGPFTLENADKQISR         0         0           VGPFTLENADKQISR         0         0           QKFAPISLTGERVDN         0         0           QKFAPISLTGERVDN         0         0           NGFLTVEPIDSRQKA         0         0           RADITVQSEKFNGLP         0         0           FLGDETPQNRIVAKS         0         0           FKDTGANIQSPRLV         0         0	TGRKDEFIOPSIVAN	0	0		
SGFRALVKITPNEDQ         0         0           SGFRALVKITPNEDQ         0         0           ERNDATFILKGPQVS         0         0           TNAISQFRGLDKEPV         0         0           TAVNREQGFISDLKP         0         0           AVEDGPQFKSLIRNT         0         0           DQKAGELVTRPFNIS         0         0           DVGPFTLENADKQISR         0         0           VGPFTLENADKQISR         0         0           VGPFTLENADKQISR         0         0           QKFAPISLTGERVDN         0         0           QKFAPISLTGERVDN         0         0           NGFLTVEPIDSRQKA         0         0           RADITVQSEKFNGLP         0         0           FLGDETPQNRIVAKS         0         0           EFKDTGANIQSPRLV         0         0	PAFINKVTFSDGOLR	0	0		
Scholzvirinitzbod         0         0           ERNDATFILKGPQVS         0         0           TNAISQFRGLDKEPV         0         0           TAVNREQGFISDLKP         0         0           AVEDGPQFKSLIRNT         0         0           DQKAGELVTRPFNIS         0         0           DNEVKIPSTRAQGLF         0         0           VGPFTLENADKQISR         0         0           SQVDPTRFLNIGEAK         0         0           ASGNLPTVQDKIFER         0         0           QKFAPISLTGERVDN         0         0           NGFLTVEPIDSRQKA         0         0           RADITVQSEKFNGLP         0         0           FLGDETPQNRIVAKS         0         0           EFKDTGANIQSPRLV         0         0           TPKOLNFAVRISDGE         1         0         Human immunodeficiency virus 1         0.33	SGERALVKITPNEDO	0	0		
TNAISQFRGLDKEPV         0         0           TAVNREQGFISDLKP         0         0           AVEDGPQFKSLIRNT         0         0           DQKAGELVTRPFNIS         0         0           DNEVKIPSTRAQGLF         0         0           VGPFTLENADKQISR         0         0           SQVDPTRFLNIGEAK         0         0           ASGNLPTVQDKIFER         0         0           QKFAPISLTGERVDN         0         0           RADITVQSEKFNGLP         0         0           FLGDETPQNRIVAKS         0         0           EFKDTGANIQSPRLV         0         0           TRKOLMEAVERSURGE         1         0	FRNDATEII KGPOVS	0	0		
TAVNREQGFISDLKP         0         0           AVEDGPQFKSLIRNT         0         0           DQKAGELVTRPFNIS         0         0           DNEVKIPSTRAQGLF         0         0           VGPFTLENADKQISR         0         0           SQVDPTRFLNIGEAK         0         0           ASGNLPTVQDKIFER         0         0           QKFAPISLTGERVDN         0         0           NGFLTVEPIDSRQKA         0         0           RADITVQSEKFNGLP         0         0           FLGDETPQNRIVAKS         0         0           EFKDTGANIQSPRLV         0         0           TPKOLNEAVRISDGE         1         0         Human immunodeficiency virus 1         0.33		0	0		
AVEDGPQFKSLIRNT         0         0           DQKAGELVTRPFNIS         0         0           DNEVKIPSTRAQGLF         0         0           VGPFTLENADKQISR         0         0           SQVDPTRFLNIGEAK         0         0           ASGNLPTVQDKIFER         0         0           QKFAPISLTGERVDN         0         0           RADITVQSEKFNGLP         0         0           FLGDETPQNRIVAKS         0         0           EFKDTGANIQSPRLV         0         0           TPKOLINEAVRISDGE         1         0		0	0		
NUCLOI QL NOLINITI         0         0           DQKAGELVTRPFNIS         0         0           DNEVKIPSTRAQGLF         0         0           VGPFTLENADKQISR         0         0           SQVDPTRFLNIGEAK         0         0           ASGNLPTVQDKIFER         0         0           QKFAPISLTGERVDN         0         0           NGFLTVEPIDSRQKA         0         0           RADITVQSEKFNGLP         0         0           FLGDETPQNRIVAKS         0         0           EFKDTGANIQSPRLV         0         0           TPKOLINFAVRISDGE         1         0         Human immunodeficiency virus 1         0.33		0	0		
DREVKIPSTRAQGLF         0         0           VGPFTLENADKQISR         0         0           SQVDPTRFLNIGEAK         0         0           ASGNLPTVQDKIFER         0         0           QKFAPISLTGERVDN         0         0           NGFLTVEPIDSRQKA         0         0           RADITVQSEKFNGLP         0         0           FLGDETPQNRIVAKS         0         0           EFKDTGANIQSPRLV         0         0           TPKOLINEAVRISDGE         1         0         Human immunodeficiency virus 1         0.33		0	0		
Direction         0         0         0           VGPFTLENADKQISR         0         0         0           SQVDPTRFLNIGEAK         0         0         0           ASGNLPTVQDKIFER         0         0         0           QKFAPISLTGERVDN         0         0         0           NGFLTVEPIDSRQKA         0         0         0           RADITVQSEKFNGLP         0         0         0           FLGDETPQNRIVAKS         0         0         0           EFKDTGANIQSPRLV         0         0         0           TPKOLNEAVRISDGE         1         0         Human immunodeficiency virus 1         0.33		0	0		
VGITTLEINRORQISIT         0         0         0           SQVDPTRFLNIGEAK         0         0         0           ASGNLPTVQDKIFER         0         0         0           QKFAPISLTGERVDN         0         0         0           NGFLTVEPIDSRQKA         0         0         0           RADITVQSEKFNGLP         0         0         0           FLGDETPQNRIVAKS         0         0         0           EFKDTGANIQSPRLV         0         0         0           TPKOLNEAVRISDGE         1         0         Human immunodeficiency virus 1         0.33		0	0		
ASGNLPTVQDKIFER     0     0       QKFAPISLTGERVDN     0     0       QKFAPISLTGERVDN     0     0       NGFLTVEPIDSRQKA     0     0       RADITVQSEKFNGLP     0     0       FLGDETPQNRIVAKS     0     0       EFKDTGANIQSPRLV     0     0       TPKOLNEAVRISDGE     1     0		0	0		
OGNELT VQUNEX         O         O           QKFAPISLTGERVDN         0         0           NGFLTVEPIDSRQKA         0         0           RADITVQSEKFNGLP         0         0           FLGDETPQNRIVAKS         0         0           EFKDTGANIQSPRLV         0         0           TPKOLNEAVRISDGE         1         0         Human immunodeficiency virus 1         0.33		0	0		
NGFLTVEPIDSRQKA     0     0       RADITVQSEKFNGLP     0     0       FLGDETPQNRIVAKS     0     0       EFKDTGANIQSPRLV     0     0       TPKOLNEAVRISDGE     1     0		0	0		
RADITVQSEKFNGLP     0     0       FLGDETPQNRIVAKS     0     0       EFKDTGANIQSPRLV     0     0       TPKOLNEAVRISDGE     1     0		0	0		
FLGDETPQNRIVAKS     0     0       EFKDTGANIQSPRLV     0     0       TPKOLNEAVRISDGE     1     0		0	0		
EFKDTGANIQSPRLV     0     0       TPKOLNEAVRISDGE     1     0	FIGDETPONRIVAKS	0	0		1
TPKOLNEAVRISDGE 1 0 33	FEKDTGANIOSPRIV	0	0		<u> </u>
		1	0	Human immunodeficiency virus 1	0 33

# https://web.expasy.org/randseq/

Table S13. (B) Fifty 15-mer Random sequence peptides generated using SMS compared with viral sequence database for homology.

	Number of Results		Results with E-value <1	
Pentide Sequence	Viruses with F-value <1	Mimiviridae with F-value <10	Virus	F-Value
			Virus	L Vulue
	0	0		
	0	0		
PECEIEDNECYWTEP	0	0		
	1	0	Saimiriina batabarnasvirus 4	0.61
	1	0	Sammine becane pessitus 4	0.01
	0	0		
	0	0		
LINKTHVINFAATEDES	0	0	Escherichia phage EcS1	
ICIPQYYMPATSKIQ	2	0	Serratia phage CHI14	0.43
SAACVPSVKVSIMPC	0	0		
DSNSRKVCQYSQCMG	0	0		
EWAWNRDGISHYIAM	1	0	Pithovirus sibericum	0.87
QFSWPDSLGMPTAKS	0	0		
MWKYGHDITGNFSTC	0	0		
AHFVAEIVQSFFPIG	0	0		
RPFKYMGMFHQHWLQ	0	0		
STHDFMATLFCDGSS	0	0		
ENWDSNHCYGRHYQL	0	0		
DSKYMTFLAICANCS	0	0		
CNHCCFWEPKRGLLP	0	0		
ETFGPLNFLANTQEW	0	0		
AMIFCFLRHEASFSP	0	0		
VKFMWHRLAHIFFRA	0	0		
QPDFWKSYPCQPKLQ	0	0		
CISPYFESTGACPTF	1	0	Aphis glycines virus 3	0.33
KDATIVDHVCTVVWQ	0	0		
NIAMDGLAMSLKARR	0	0		
CSDFPHIRGLVILLP	0	0		
RPYKSDHHILCAKNS	0	0		
DHKIKYLSSYYTDPN	0	0		
QNLQQPKCHTEFVDF	0	0		
MTYMDRNNEVNELTK	1	0	Cvanophage S-RIM14	0.93
TKGTWRRGPMARWES	0	0		
PYNKWMSDSDRMTNR	0	0		
TSAHMHCKAGKDALE	0	0		
RWHVMVNWLYIADDG	0	0		
GTEICPMIVYHFIWW	1	0	White spot syndrome virus	0.33
YYTWAOPGKRKCSET	0	0		
WFMWVPAEMQRWHRV	0	0		
CESOHPNMSRPYMSA	0	0		
IYVEPTIWSARKHLG	0	0		
GGLVTWYAFRDMTDY	0	0		
SPVANHIMPPTFYNG	0	0		
CGGLNELIEWQPPFE	0	0		
CVNQGYVHFLHYOWC	0	0		
GFSPTYSHEENWVAI	0	0		
RCEDPLTILHPAKLC	0	0		
EFPHIMIAHWGAGGM	0	0		
FDSPKELDPTNNDRR	0	0		
ATRNMPPESRTSCDV	0	0		1

# http://www.bioinformatics.org/sms2/random_protein.html

#### Table S14. (A) BLAST results for topoisomerase I immunodominant peptide.

RQRAVALYFIDKLAL in Topoisomerase I			
Protein	E-Value	% Homology	
topoisomerase (DNA) I, isoform CRA_b	0.00000006	100%	
Topoisomerase (DNA) I	0.00000006	100%	
topoisomerase I	0.00000006	100%	
DNA topoisomerase 1	0.00000006	100%	
DNA topoisomerase I	0.00000006	100%	
DNA topoisomerase I	0.00000006	100%	
topoisomerase (DNA) I, isoform CRA_c	0.00000006	100%	
DNA topoisomerase I, mitochondrial isoform X1	0.00000006	100%	
TOP1MT protein	0.00000006	100%	
DNA topoisomerase I, mitochondrial isoform 1 precursor	0.00000006	100%	
Topoisomerase (DNA) I, mitochondrial	0.00000006	100%	
mitochondrial DNA topoisomerase I	0.00000006	100%	
Chain A, Human Dna Topoisomerase I In Covalent Complex With A 22 Base Pair Dna Duplex	0.00000006	100%	
Chain A, Human Dna Topoisomerase I (70 Kda) In Complex With The Indolocarbazole Sa315f And Covalent Complex With A 22 Base Pair Dna Duplex	0.00000006	100%	
Chain A, Human Topoisomerase I (Topo70) Double Mutant K532rY723F	0.00000006	100%	
Chain A, Human Topoisomerase IDNA COMPLEX	0.00000006	100%	
Chain A, Human Reconstituted Dna Topoisomerase I In Covalent Complex With A 22 Base Pair Dna Duplex	0.00000006	100%	
Chain A, Structural Mechanisms Of Camptothecin Resistance By Mutations In Human Topoisomerase I	0.00000006	100%	
Chain C, Structural Mechanisms Of Camptothecin Resistance By Mutations In Human Topoisomerase I	0.00000006	100%	
Chain A, Human Dna Topoisomerase I (70 Kda) In Non-Covalent Complex With A 22 Base Pair Dna Duplex Containing An 8-Oxog Lesion	0.00000006	100%	
Chain A, Human Topoisomerase I Ara-C Complex	0.00000006	100%	
Chain A, Crystal Structure Of Human Topoisomerase I Dna Complex	0.00000006	100%	
unnamed protein product	0.00000006	100%	
unnamed protein product	0.00000006	100%	
DNA topoisomerase I, mitochondrial isoform 2	0.00000006	100%	
unnamed protein product	0.00000006	100%	
unnamed protein product	0.00000006	100%	
topoisomerase (DNA) I, mitochondrial, isoform CRA_e	0.000006	100%	
topoisomerase (DNA) I, mitochondrial, isoform CRA_f	0.000006	100%	
topoisomerase (DNA) I, mitochondrial, isoform CRA_c	0.000006	100%	
topoisomerase (DNA) I, mitochondrial, isoform CRA_a	0.000006	100%	
topoisomerase (DNA) I, mitochondrial, isoform CRA_d	0.000006	100%	
topoisomerase (DNA) I, mitochondrial, isoform CRA_b	0.000006	100%	

GRDLINLAKKRTNII in Fibrillarin					
Protein	E-Value	% Homology			
rRNA 2'-O-methyltransferase fibrillarin	0.00000008	100%			
fibrillarin	0.00000008	100%			
rRNA 2'-O-methyltransferase fibrillarin isoform X1	0.00000008	100%			
FBRL_HUMAN	0.00000008	100%			
unnamed protein product	0.00000008	100%			
rRNA 2'-O-methyltransferase fibrillarin isoform X2	0.00000008	100%			
Unknown (protein for IMAGE:4538098)	0.00000008	100%			
Chain A, Human Fibrillarin	0.00000008	100%			
FBRL_HUMAN [AA 1- 227]	0.00000008	100%			
rRNA/tRNA 2'-O-methyltransferase fibrillarin-like protein 1	0.000001	87%			
hCG1641252	0.000001	87%			

Table S14. (B) BLAST results for fibrillarin immunodominant peptide.

#### Table S14. (C) BLAST results for centromere protein A immunodominant peptide.

LQEAAEAFLVHLFED in Centromere Protein A				
Protein	E-Value	% Homology		
Chain A. Crvo-EM structure of CENP-A nucleosome in complex with kinetochore protein CENP-N	0.00000004	100%		
Chain A. Structure Of A Cenp-A-Histone H4 Heterodimer In Complex With Chaperone Hiuro	0.00000004	100%		
Chain A, The Structure Of The Centromeric Nucleosome Containing Cenp-A	0.00000004	100%		
histone H3-like centromeric protein A isoform a	0.00000004	100%		
centromere protein A, 17kDa, isoform CRA b	0.00000004	100%		
Chain A, Crystal Structure Of (Cenp-AH4)2 HETEROTETRAMER	0.00000004	100%		
Chain A, Complex Structure 6 Of DaxxH3.3(SUB7)H4	0.000001	93%		
Chain A, Complex Structure 3 Of DaxxH3.3(SUB5,G90A)H4	0.000001	93%		
Chain A, Complex Structure 2 Of DaxxH3.3(SUB5,G90M)H4	0.000001	93%		
Chain A, Complex Structure 1 Of DaxxH3.3(SUB5)H4	0.000001	93%		
Chain A, Crystal Structure Of Human Nucleosome Containing H3.y	0.000006	80%		
histone cluster 2, H3c pseudogene	0.000006	80%		
H3L-like histone	0.000008	80%		
histone H3-like centromeric protein A isoform b	0.00001	100%		
hCG1749005	0.0002	80%		
Chain E, Crystal Structure of the heterotypic nucleosome containing human CENP-A and H3.3	0.0002	80%		
Chain A, The Crystal Structure Of The Nucleosome Containing H3.6	0.0002	80%		
Chain A, The Nucleosome Structure Containing H2b-k120 And H4-k31 Monoubiquitinations	0.0002	80%		
Chain A, The Human Nucleosome Structure Containing The Histone Variant H3.3	0.0002	80%		
Chain A, The Human Nucleosome Structure Containing The Histone Variant H3.2	0.0002	80%		
histone H3.3	0.0002	80%		
Chain A, Crystal Structure Of The Nucleosome Containing Histone H3 With The Crotonylated Lysine 122	0.0002	80%		
unnamed protein product	0.0002	80%		
histone H3.2	0.0002	80%		
unnamed protein product	0.0002	80%		
H3 histone, family 3A	0.0002	80%		
Chain A, Crystal structure of EBV tegument protein BNRF1 in complex with histone chaperone DAXX and histones H3.3-H4	0.0002	80%		
Chain M, CryoEM Structure INO80core Nucleosome complex	0.0002	80%		
hCG1639761	0.0002	80%		
Chain A, Crystal Structure Of Human Mcm2 Hbd And Asf1b Chaperoning A Histone H3.2-h4 Dimer	0.0002	80%		
Chain A, Crystal structure of Human MCM2 HBD chaperoning a histone H3-H4 tetramer	0.0002	80%		
Chain A, Crystal Structure Of Human Nucleosome Core Particle Containing H3k122q Mutation	0.001	80%		
Chain A, Crystal Structure Of Human Nucleosome Core Particle Containing H3k115q Mutation	0.001	80%		
Chain A, Crystal Structure Of Human Nucleosome Core Particle Containing H3k79q Mutation	0.001	80%		
Chain A, Crystal Structure Of Human Nucleosome Core Particle Containing H3k64q Mutation	0.001	80%		
Chain A, Crystal Structure Of Human Nucleosome Core Particle Containing H3k56q Mutation	0.001	80%		
Chain A, The Human Nucleosome Structure	0.001	80%		
Histone cluster 1, H3i	0.001	80%		
histone H3.1	0.001	80%		
Histone cluster 1, H3i	0.001	80%		
histone H3	0.001	80%		
Chain A, Crystal Structure Of The Nucleosome Core Particle	0.001	80%		
Chain A, Crystal Structure Of Human Nucleosome Core Particle Lacking H3.1 N- Terminal Region	0.001	80%		
Chain G, Crystal Structure Of Human Spt16 Mid-aid/h3-h4 Tetramer Fact Histone Complex	0.001	80%		
histone cluster 2 H3 pseudogene 2	0.002	80%		
Chain A, The Nucleosome Containing A Testis-Specific Histone Variant, Human H3t	0.012	73%		
Chain A, The Nucleosome Containing Human H3.5	0.012	73%		
histone H3.1t	0.012	73%		
Histone H3-like	0.012	73%		
histone H3.3C	0.012	73%		
hCG1789848	0.048	71%		

		ATA+	SSc Subset	
Ancestry		Peptide		Affinity(nM)
	Position	Sequence	Core	DRA1*01:01/DRB1*15:02
Japanese Thai	473	RQRAVALYFIDKLAL	VALYFIDKL	299.7
	Position	Sequence	Core	DPA1*02:01/DPB1*13:01
Turkish Iranian	473	RQRAVALYFIDKLAL	VALYFIDKL	264.2
	Position	Sequence	Core	DRA1*01:01/DRB1*11:04
Turkish Iranian	473	RQRAVALYFIDKLAL	VALYFIDKL	340.3
	Position	Sequence	Core	DRA1*01:01/DRB1*08:02
Mexican	473	RQRAVALYFIDKLAL	VALYFIDKL	600.1
	Position	Sequence	Core	DRA1*01:01/DRB1*16:02
Choctaw Indian	473	RQRAVALYFIDKLAL	VALYFIDKL	268.8

Table S15. Predicted immunodominant peptides in other ancestries for (A) topoisomerase I and (B) centromere protein A.

_	
	٠
-	E.

Α

		ACA	+ SSc Subset	
Ancestry	Peptide			Affinity(nM)
	Position	Sequence	Core	DQA1*01:01/DQB1*05:01
Japanese Chinese	94	LQEAAEAFLVHLFED	AEAFLVHLF	304.3
	Position	Sequence	Core	DQA1*03:03/DQB1*03:01 [@]
Japanese	94	LQEAAEAFLVHLFED	AEAFLVHLF	400.3
	Position	Sequence	Core	DQA1*05:03/DQB1*03:01®
Japanese	94	LQEAAEAFLVHLFED	AEAFLVHLF	421.7
	Position	Sequence	Core	DQA1*06:01/DQB1*03:01 [@]
Japanese	94	LQEAAEAFLVHLFED	AEAFLVHLF	578.6
		@ Three dif	ferent haplotypes	

#### References

- 1. Motyer A, et al. (2016) Practical use of methods for imputation of HLA alleles from SNP genotype data. bioRxiv.
- 2. Das S, et al. (2016) Next-generation genotype imputation service and methods. Nat Genet 48(10):1284–87.
- 3. Dilthey AT, et al. (2019) Hla*la—hla typing from linearly projected graph alignments. Bioinformatics 35(21):4394–4396.
- 4. Jia X, et al. (2013) Imputing amino acid polymorphisms in human leukocyte antigens. PloS One 8(6):e64683.
- 5. Thomson G (1995) Hla disease associations: models for the study of complex human genetic disorders. Critical Reviews in Clinical Laboratory Sciences 32(2):183–219.
- Robinson J, et al. (2015) The IPD and IMGT/HLA database: allele variant databases. Nucleic Acids Res 43(Database issue):D423–31.
- Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21(2):263–65.
- 8. Pettersen EF, et al. (2004) UCSF Chimera–a visualization system for exploratory research and analysis. J Comput Chem 25(13):1605–12.
- 9. Zhang H, Bonney G (2000) Use of classification trees for association studies. Genet Epidemiol 19(4):323-32.
- 10. Brown JH, et al. (2015) Pillars article: three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Journal of Immunology 194(1):5–11.
- 11. Chicz RM, et al. (1992) Predominant naturally processed peptides bound to HLA-DR1 are derived from MHC-related molecules and are heterogeneous in size. *Nature* 358(6389):764–68.
- 12. Bassani-Sternberg M, et al. (2016) Direct identification of clinically relevant neoepitopes presented on native human melanoma tissue by mass spectrometry. *Nature communications* 7(13404).
- 13. Chong C, et al. (2018) High-throughput and Sensitive Immunopeptidomics Platform Reveals Profound Interferongamma-Mediated Remodeling of the Human Leukocyte Antigen (HLA) Ligandome. *Molecular Cell Proteomiscs* 17(3):533–48.
- 14. Ritz D, et al. (2018) Membranal and Blood-Soluble HLA Class II Peptidome Analyses Using Data-Dependent and Independent Acquisition. *Proteomics* 18(12):e1700246.
- 15. Collado JA, et al. (2013) Composition of the HLA-DR-associated human thymus peptidome. European journal of immunology 43(9):2273–82.
- 16. Ciudad MT, et al. (2017) Analysis of the HLA-DR peptidome from human dendritic cells reveals high affinity repertoires and nonconventional pathways of peptide generation. *Journal of Leukocyte Biology* 101(1):15–27.
- 17. Godkin AJ, et al. (1998) Use of complete eluted peptide sequence data from HLA-DR and -DQ molecules to predict T cell epitopes, and the influence of the nonbinding terminal regions of ligands in epitope selection. *Journal of Immunology* 161(2):850–58.
- 18. Breous E, Somanathan S, Wilson JM (2009) Identification of the immunodominant cytotoxic T-cell epitope of human -1 antitrypsin. *Gene Therapy* 16(11):1380–82.
- 19. Hartman IZ, et al. (2010) A reductionist cell-free major histocompatibility complex class II antigen processing system identifies immunodominant epitopes. *Nature Medicine* 16(11):1333–40.
- 20. Stickler M, et al. (2004) The HLA-DR2 haplotype is associated with an increased proliferative response to the immunodominant CD4(+) T-cell epitope in human interferon-beta. *Genes and immunity* 5(1):1–7.
- 21. Okhuri T, et al. (2009) Identification of novel helper epitopes of MAGE-A4 tumour antigen: useful tool for the propagation of Th1 cells. *British journal of cancer* 100(7):1135–43.
- Zhao w, Sher X (2018) Systematically benchmarking peptide-MHC binding predictors: From synthetic to naturally processed epitopes. *PLoS computational biology* 14(11):e1006457.
- 23. Jensen KK, et al. (2018) Improved methods for predicting peptide binding affinity to MHC class II molecules. *Immunology* 154(3):394–406.
- 24. Nielsen M, et al. (2008) Quantitative predictions of peptide binding to any hla-dr molecule of known sequence: Netmhciipan. *PLoS Comput Biol* 4(7):e1000107.
- Singh T, et al. (2018) Epitope-binding characteristics for risk versus protective DRB1 alleles for visceral leishmaniasis. J Immunol 200(8):2727–37.
- 26. Degoot AM, Chirove F, Ndifon W (2018) Trans-allelic model for prediction of peptide:MHC-II interactions. *Front Immunol* 9:1410.
- 27. Nielsen M, Lundegaard C, Lund O (2007) Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. *BMC Bioinformatics* 8:238.
- 28. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215(3):403-10.
- 29. Stothard P (2000) The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *BioTechniques* 28(6):1102–04.
- 30. Artimo P, et al. (2016) ExPASy: SIB bioinformatics resource portal. Nucleic Acids Res 40(Web Server issue):W597-603.
- 31. Hogeboom C, et al. (2015) Peptide motif analysis predicts lymphocytic choriomeningitis virus as trigger for multiple sclerosis. *Mol Immun* 67(2 Pt B):625–35.
- 32. Rosa JE, et al. (2011) Incidence and prevalence of systemic sclerosis in a healthcare plan in Buenos Aires. J Clin Rheumatol 17(2):59–63.
- Tamaki T, Mori S, Takehara K (1991) Epidemiological study of patients with systemic sclerosis in Tokyo. Arch Dermatol Res 283(6):366–71.

- 34. Michet Jr. CJ, McKenna CH, Elveback LR, Kaslow RA, Kurland LT (1985) Epidemiology of systemic lupus erythematosus and other connective tissue diseases in Rochester, Minnesota, 1950 through 1979. Mayo Clin Proc 60(2):105–13.
- 35. Allcock RJ, Forrest I, Corris PA, Crook PR, Griffiths ID (2004) A study of the prevalence of systemic sclerosis in northeast England. *Rheumatology* 43(5):596–602.
- 36. Le Guern V, et al. (2004) Prevalence of systemic sclerosis in a French multi-ethnic county. Rheumatology 43(9):1129–37.
- Nietert PJ, Silverstein MD, Silver RM (2001) Hospital admissions, length of stay, charges, and in-hospital death among patients with systemic sclerosis. J Rheumatol 28(9):2031–37.
- 38. Eason RJ, Tan PL, Gow PJ (1981) Progressive systemic sclerosis in Auckland: a ten year review with emphasis on prognostic features. Aust N Z J Med 11(6):657–62.
- 39. Alamanos Y, et al. (2005) Epidemiology of systemic sclerosis in northwest Greece 1981 to 2002. Semin Arthritis Rheum 34(5):714–20.
- 40. Arnett FC, et al. (1996) Increased prevalence of systemic sclerosis in a Native American tribe in Oklahoma. Association with an Amerindian HLA haplotype. *Arthritis Rheum* 39(8):1362–70.
- 41. Tan FK, et al. (1999) HLA haplotypes and microsatellite polymorphisms in and around the major histocompatibility complex region in a Native American population with a high prevalence of scleroderma (systemic sclerosis). *Tissue Antigens* 53(1):74–80.