

Immunogenic Amino Acid Motifs and Linear Epitopes of COVID-19 mRNA Vaccines

Adam V Wisnewski^{1*}, Carrie A Redlich¹, Kathy Kamath², Queenie-Ann Abad¹, Richard F Smith¹, Louis Fazen¹, Romero Santiago¹, Jian Liu¹, Julian Campillo Luna¹, Brian Martinez², Elizabeth Baum-Jones², Rebecca Waitz², Winston A Haynes², John C Shon²

¹Department of Medicine, Yale University School of Medicine, New Haven, CT 06520

²Serimmune, Inc., Goleta, CA, 93117, USA

*Corresponding Author: Adam Wisnewski, PO Box 208057, Yale University School of Medicine, New Haven, CT 06520, email: adam.wisnewski@yale.edu, phone: (203) 737-4054, fax: (203) 785-3826

Table of Contents	Page
Table S1. Table of study subjects	2
Table S2. Table of fold enrichment in amino acid motifs	4
Figure S1. Correlation of enrichment for IgG with specificity for overlapping amino acid motifs	5
Figure S2. Correlation of enrichment for IgG with specificity for non-overlapping amino acid motifs and age	6
Correlation Matrix 1. Heat map of correlation between recognition of different amino acid motifs	7
Figure S3. Correlation of PIWAS score for LE-1 vs. enrichment for amino acid motif [FY]XX[QE]LXXWF	8
Figure S4. Linear epitope mapping of individual subject's IgG, after COVID/pre-vaccine vs. after COVID/post vaccine	9
List S1. Peptides with strongest identity to SARS-CoV-2 spike mRNA vaccine epitopes identified though BLAST® analysis	10
Table S3. PIWAS values for all subjects.	11

Immunogenic Amino Acid Motifs and Linear Epitopes of COVID-19 mRNA Vaccines: Supplemental Materials

Table S1. Table of Study Subjects

Subject #	Sex	Age Range	COVID Vaccination	Vaccine Manufacturer	COVID Vaccination Initial	COVID Vaccination Booster	Timepoint	SARS-CoV-2 PCR Status	SARS-CoV-2 IgG Status	Spike IgG ELISA OD
1	Male	35-40					pre-vaccine	N/A	Negative	not detected
2	Female	30-35					pre-vaccine	N/A	Negative	not detected
3	Female	46-50					pre-vaccine	N/A	Negative	not detected
4	Male	25-30					pre-vaccine	N/A	Negative	not detected
5	Female	46-50					pre-vaccine	N/A	Negative	not detected
6	Female	40-45					pre-vaccine	N/A	Negative	not detected
7	Female	56-60					pre-vaccine	N/A	Negative	not detected
8	Male	35-40					pre-vaccine	N/A	Negative	not detected
9	Female	35-40					pre-vaccine	N/A	Negative	not detected
10	Female	40-45					pre-vaccine	N/A	Negative	not detected
11	Male	30-35					pre-vaccine	N/A	Negative	not detected
12	Female	26-30					pre-vaccine	N/A	Negative	not detected
13	Female	26-30					pre-vaccine	N/A	Negative	not detected
14	Female	30-35					pre-vaccine	N/A	Negative	not detected
15	Female	26-30					pre-vaccine	N/A	Negative	not detected
16	Female	30-35					pre-vaccine	N/A	Negative	not detected
17	Female	26-30					pre-vaccine	N/A	Negative	not detected
18	Male	26-30					pre-vaccine	N/A	Negative	not detected
19	Male	56-60					pre-vaccine	N/A	Negative	not detected
20	Female	26-30					pre-vaccine	N/A	Negative	not detected
1	Male	35-40	yes	Pfizer	yes	yes	D36 Post Vaccine	N/A	Positive	1.6
2	Female	30-35	yes	Pfizer	yes	yes	D36 Post Vaccine	N/A	Positive	1.7
3	Female	46-50	yes	Pfizer	yes	yes	D36 Post Vaccine	N/A	Positive	2.0
4	Male	25-30	yes	Pfizer	yes	yes	D36 Post Vaccine	N/A	Positive	2.5
5	Female	46-50	yes	Pfizer	yes	yes	D36 Post Vaccine	N/A	Positive	2.4
6	Female	40-45	yes	Pfizer	yes	yes	D36 Post Vaccine	N/A	Positive	1.3
7	Female	56-60	yes	Pfizer	yes	yes	D36 Post Vaccine	N/A	Positive	1.4
8	Male	35-40	yes	Pfizer	yes	yes	D35 Post Vaccine	N/A	Positive	1.0
9	Female	35-40	yes	Pfizer	yes	yes	D34 Post Vaccine	N/A	Positive	2.0
10	Female	40-45	yes	Pfizer	yes	yes	D34 Post Vaccine	N/A	Positive	1.9
11	Male	30-35	yes	Pfizer	yes	yes	D34 Post Vaccine	N/A	Positive	1.8
12	Female	26-30	yes	Pfizer	yes	yes	D34 Post Vaccine	N/A	Positive	1.3
13	Female	26-30	yes	Pfizer	yes	yes	D34 Post Vaccine	N/A	Positive	1.8
14	Female	30-35	yes	Pfizer	yes	yes	D34 Post Vaccine	N/A	Positive	2.0
15	Female	26-30	yes	Pfizer	yes	yes	D34 Post Vaccine	N/A	Positive	2.4
16	Female	30-35	yes	Pfizer	yes	yes	D33 Post Vaccine	N/A	Positive	2.0
17	Female	26-30	yes	Pfizer	yes	yes	D33 Post Vaccine	N/A	Positive	1.8
18	Male	26-30	yes	Pfizer	yes	yes	D31 Post Vaccine	N/A	Positive	3.0
19	Male	56-60	yes	Pfizer	yes	yes	D30 Post Vaccine	N/A	Positive	0.9
20	Female	26-30	yes	Pfizer	yes	yes	D30 Post Vaccine	N/A	Positive	2.1
21	Male	50-55	yes	Moderna	yes	yes	D36 Post Vaccine	N/A	Positive	2.2
22	Female	56-60	yes	Moderna	yes	yes	D38 Post Vaccine	N/A	Positive	2.0
23	Male	30-35	yes	Moderna	yes	yes	D35 Post Vaccine	N/A	Positive	2.5

Immunogenic Amino Acid Motifs and Linear Epitopes of COVID-19 mRNA Vaccines: Supplemental Materials

24	Male	20-25	yes	Moderna	yes	yes	D35 Post Vaccine	N/A	Positive	2.4
25	Female	95-100	yes	Moderna	yes	yes	D35 Post Vaccine	N/A	Positive	2.7
26	Female	50-55	yes	Moderna	yes	yes	D35 Post Vaccine	N/A	Positive	1.9
27	Female	20-25	yes	Moderna	yes	yes	D35 Post Vaccine	N/A	Positive	2.4
28	Male	26-30	yes	Moderna	yes	yes	D35 Post Vaccine	Negative	Positive	2.5
29	Female	40-45	yes	Moderna	yes		D13-Post vaccine	positive	Positive	1.8
30	Female	40-45	yes	Pfizer	yes		D11-Post Vaccine	positive	Positive	>3.5
31	Male	46-50	yes	Moderna	yes		D10-Post vaccine	positive	Positive	>3.5
32	Male	30-35	yes	Moderna	yes		D14-Post Vaccine	positive	Positive	>3.5
33	Female	60-65	yes	Pfizer	yes		D17-Post Vaccine	positive	Positive	>3.5
32	Male	30-35	no				pre-vaccine	positive	Positive	1.6
33	Female	60-65	no				pre-vaccine	positive	Positive	1.1

Table 2. Fold Enrichment in Amino Acid Motif Recognition by Different Subjects

Subject ID #	[ACP]XXXDXSGI	DXSGIXXX[DGESN]	FXXELXX[WY]F	KWXXFQQ	[FY]XX[QE]LXXWF	P[EDT]WRXY	RXPX[ST]L[DE]	K[FY][LM]XFXQ	PXWR[IV]Y	[AC]K[FY]XXFXQ
1	24.12188	28.89815	14.16188	2.90406	10.03086	1.63869	1.8324	2.86602	0.97399	1.14955
2	1.3568	2.80499	0.79658	8.14162	0.91802	2.94699	1.60712	4.3079	2.53118	2.10418
3	45.54796	42.14705	7.65677	15.33523	4.19016	10.15588	11.67805	6.29133	9.65864	4.88007
4	74.64611	52.41181	35.46319	23.11651	17.59506	3.7325	12.62332	9.11356	2.20935	6.17086
5	86.52744	75.57166	35.07616	7.11044	4.17209	2.62366	0.83056	2.41959	2.41142	1.62939
6	9.08279	9.20852	20.26051	4.23238	3.19377	4.01896	7.89146	2.39675	3.20445	3.09941
7	11.775	9.4349	0.34627	2.19791	0.53115	1.09487	1.67598	1.86742	1.91283	0.83012
8	12.75622	16.91751	26.06295	0.85512	10.78932	0.87751	1.168	1.58294	0.42967	1.7407
9	16.87036	18.89413	22.21202	31.01272	1.58234	22.15339	4.71435	14.63689	18.01009	9.6201
10	7.991	7.36731	24.80848	9.95705	23.19741	6.76044	7.69714	2.83034	4.01064	2.25895
11	8.81782	8.24756	6.1367	3.31935	1.65641	7.76581	5.46426	2.23179	5.67561	1.65069
12	33.42265	30.07481	17.76317	18.24522	7.05165	4.54149	3.89708	6.58812	4.62662	4.24812
13	21.65294	21.84343	58.46848	19.18538	26.32006	9.91419	20.31965	5.5046	5.86914	3.792
14	56.01096	48.77159	47.76832	9.07255	23.93239	11.28625	20.43179	3.24127	5.76119	2.20879
15	27.55929	28.09445	27.27589	28.66981	5.85507	23.93945	13.42879	11.62889	16.50785	7.58091
16	69.39203	72.35511	40.49051	82.80327	6.35543	5.81869	10.75896	31.03799	3.28103	25.63274
17	3.64422	4.90357	10.3055	2.92826	1.50749	2.41707	6.58392	2.84551	1.50774	1.42116
18	39.47938	30.3566	20.42899	22.36679	10.4027	11.71427	2.73141	7.92746	13.83002	5.05539
19	2.27761	3.50932	8.46049	2.84196	4.86742	1.2403	1.71339	1.13049	2.1335	1.25074
20	11.09478	10.25416	42.05175	28.66604	13.59703	34.38442	1.30943	5.73332	20.88919	3.42484
21	0.66991	1.17311	6.45055	2.99584	2.0799	0.64202	0.87316	1.43845	0.73027	1.2127
22	0.78098	0.56674	1.44845	2.59875	1.38039	3.30676	2.21652	1.26755	2.51584	1.28278
23	8.76277	9.59592	28.73647	10.94687	4.11469	2.28089	4.93592	5.0789	1.31525	2.97093
24	14.24855	16.75375	5.25025	45.38644	2.94386	23.11886	12.23777	21.57018	17.69763	15.08657
25	28.48567	27.45478	21.08092	20.42211	2.53462	16.57896	4.38707	9.4571	13.02877	5.76048
26	1.45674	1.00078	7.97533	3.71951	1.18545	2.23491	5.45684	1.57078	1.08261	1.1916
27	6.6744	7.92231	12.1883	6.86614	7.14451	6.0221	1.8916	4.05087	5.06876	2.38904
28	33.77127	36.42345	43.57688	17.74377	5.07247	10.1257	4.87788	5.65815	9.38512	4.34097
29	118.4104	96.1604	61.08806	41.86992	29.68794	26.41247	2.33624	21.91133	71.63942	13.09448
30	0.744	0.74156	77.70677	1.32981	41.33262	21.34096	4.75237	0.91951	8.95235	0.61937
31	3.77329	2.69862	57.97974	32.51077	40.72458	54.41257	29.78026	17.23821	44.7197	11.13788
32	11.83184	15.71697	55.50956	41.74956	29.17609	25.97491	2.97936	19.4578	14.60386	13.26798
33	1.95214	0.56701	46.74683	3.37901	30.57662	0.96797	3.16383	4.33146	0.79088	3.41584

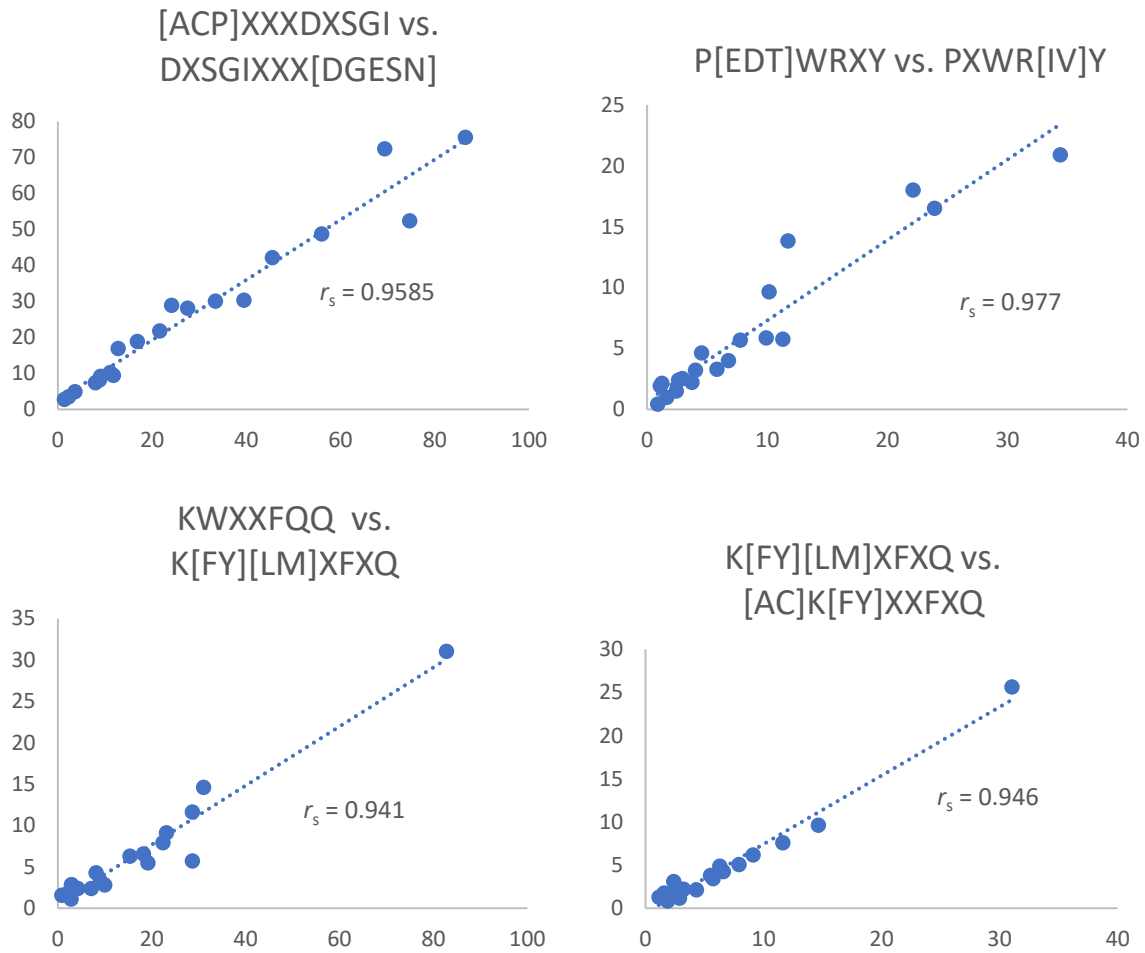


Figure S1. Correlation of IgG recognition of peptide motifs with overlapping amino acids. The fold enrichment in IgG with specificity for the different pairs of amino acid motifs (respective x and Y-axis) are shown along with the Spearman rank correlation (r_s) value. Each symbol represents a different COVID-19 mRNA vaccine recipient. Trend lines were fitted by linear regression. Spearman rank correlation (r_s) is displayed.

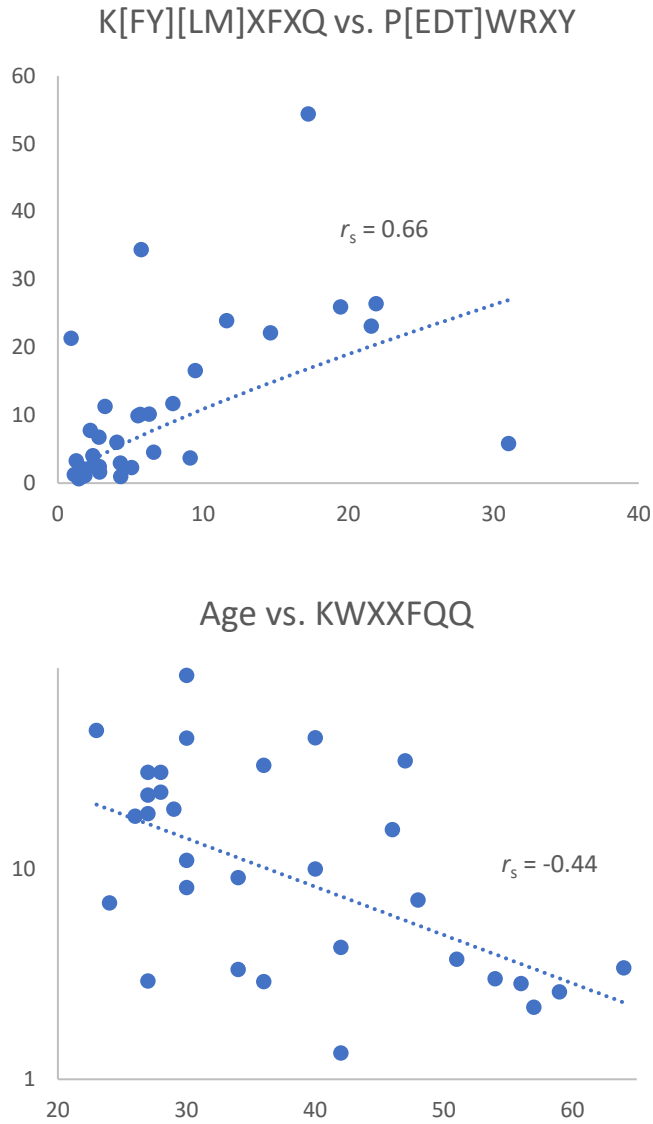
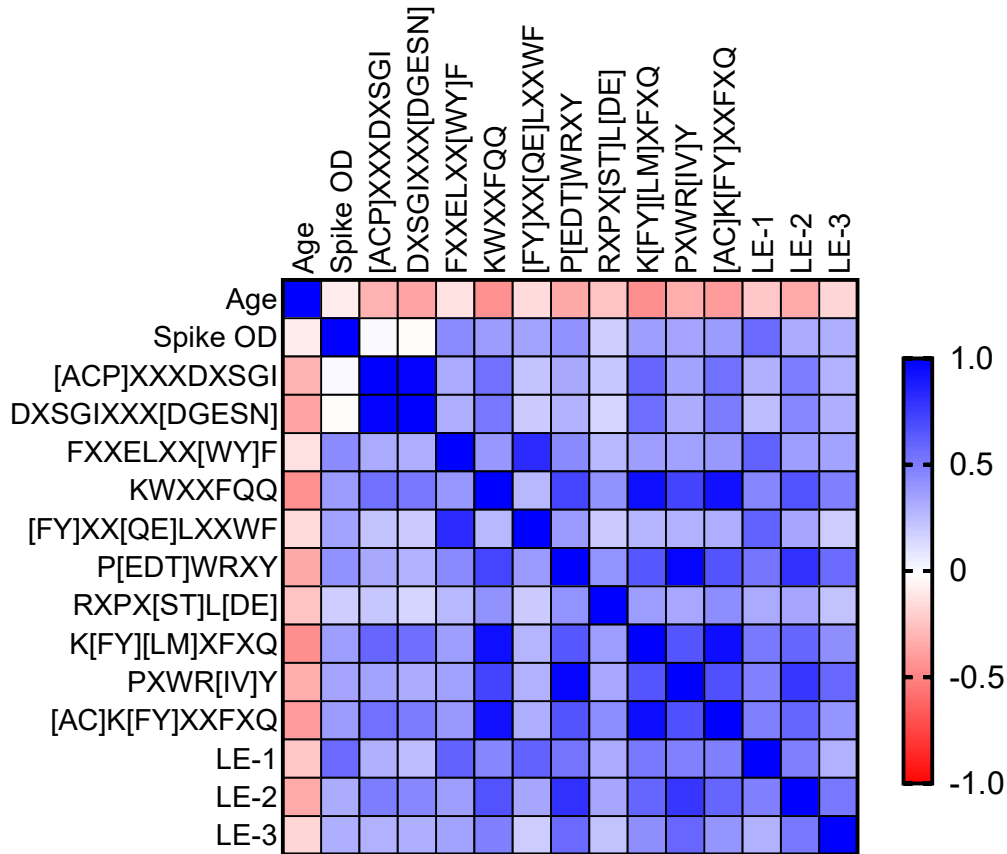


Figure S2. Correlation of amino acid motif recognition and age of vaccinated subjects. Panel A: The fold enrichment in IgG with specificity for the amino acid motifs K[FY][LM]XFXQ vs. P[EDT]WRXY is shown. Each symbol corresponds to a different vaccinated individual. Panel B: The fold enrichment in IgG with specificity for the amino acid motif KWXXFQQ (Y-axis) is plotted against age (X-axis). Trend lines were fitted by linear regression. Spearman rank correlation (r_s) is displayed.



Correlation Matrix 1. The correlation between enrichment in IgG recognition of amino acid motifs and linear epitopes. Spearman rank correlation values are displayed as a heat map by color according to key on right for different epitopes recognized by IgG from COVID-19 mRNA vaccine recipients.

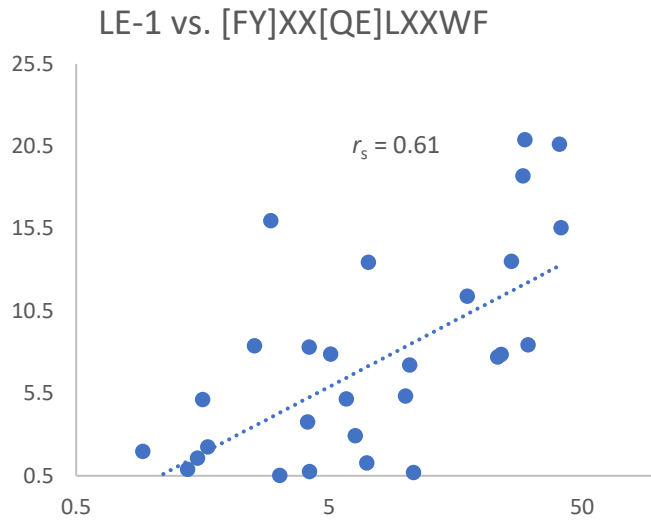


Figure S3. Correlation between IgG recognition of amino acid motif [FY]XX[QE]LXXWF and linear epitope LE-1. The PIWAS score (Y-axis) for IgG recognition of LE-1 for individual subjects was plotted vs. the fold enrichment in IgG (X-axis) with specificity for the amino acid motif [FY]XX[QE]LXXWF in subjects that received mRNA vaccine. Each symbol corresponds to a different vaccinated individual. Trend lines were fitted by linear regression. Spearman rank correlation (r_s) is displayed.

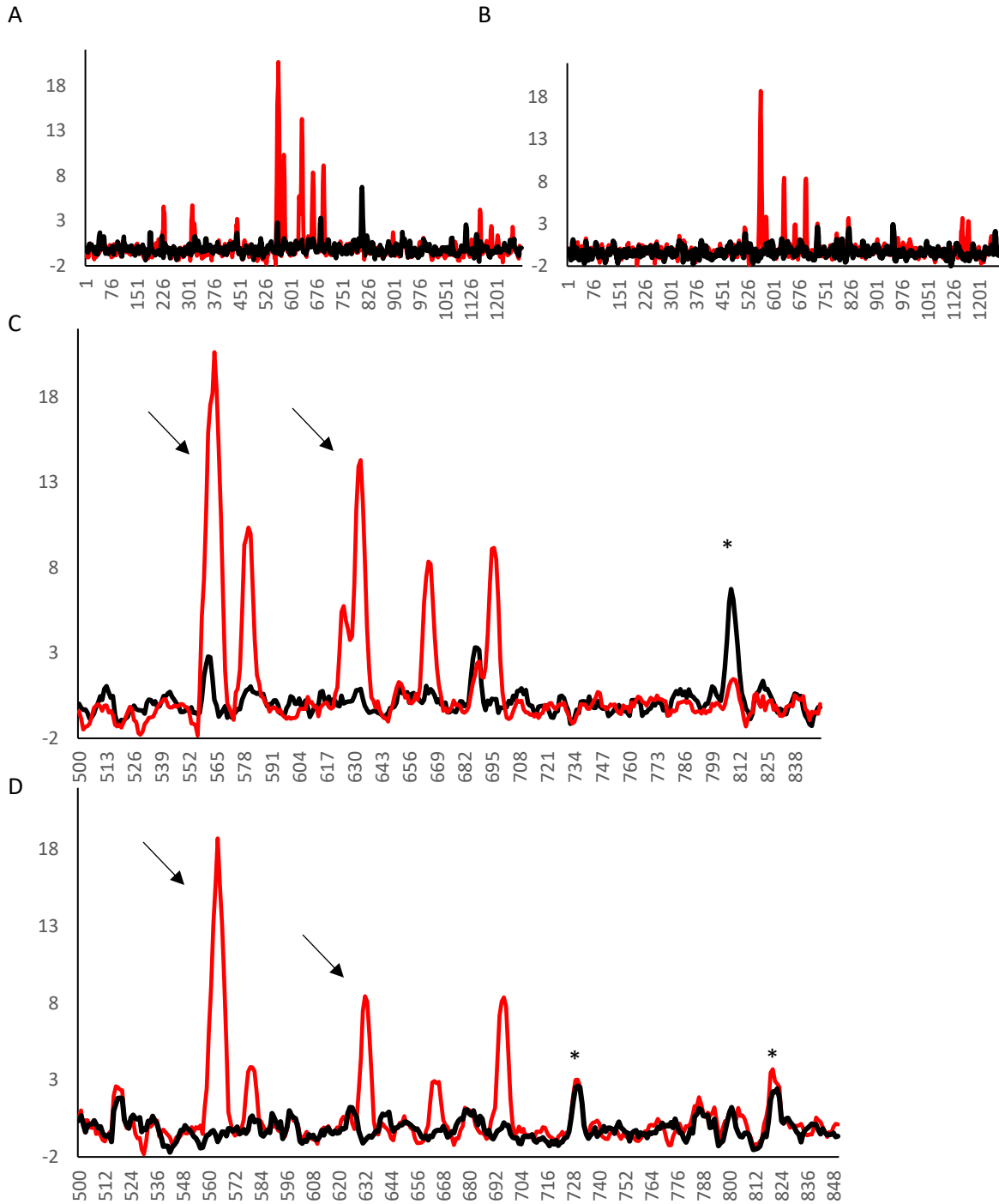


Figure S4. Examples of epitopes that drive IgG responses in mRNA vaccinated subjects with prior COVID-19. The PIWAS scores (Y-axis) are graphed for peptides surrounding different amino acids along the SARS-CoV-2 spike protein (X-axis) that are recognized by IgG after COVID-19/pre-vaccine (black) and after COVID-19 and vaccination (red). Panels A (subject 32) and B (subject 33) show entire spike protein and Panels C and D respectively focus on the region surrounding the S1/S2 junction. Note dominant epitope LE-1 and LE-2 (→) in both subjects and lack of substantial boost to highlighted naturally occurring epitopes (*)

List S1. Peptides with strongest identity to SARS-CoV-2 spike mRNA vaccine epitopes identified through BLAST® analysis (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>)

1) LE-1 Epitope: FLPFQQFGRDIA

>translation elongation factor Ts [Candidatus Omnitrophica bacterium]

Sequence ID: MBI1977219.1 Length: 198

Query 4 FQQFGRDIA 12

FQQFGRDIA

Sbjct 92 FQQFGRDIA 100

2) LE-2 Epitope: DQLTPTWRVY

>hypothetical protein [Thermoplasmata archaeon]

Query 1 DQLTPTWRV 9

DQLTPTWRV

Sbjct 39 DQLTPTWRV 47

3) LE-3 Epitope: FKEELDKYFKN

>MULTISPECIES: universal stress protein [Aquimarina]

Sequence ID: WP_074407094.1 Length: 283

Query 2 KEELDKYFKN 11

KEELDKYFKN

Sbjct 205 KEELDKYFKN 214

Table with columns for subject ID, amino acid position, and three different vaccine samples (Pre-vaccine samples from N=20 subjects, Post-vaccine samples from N=20 subjects, and Post-vaccine samples from N=5 prior COVID-19 post vaccination Moderna recipients). The table contains a dense grid of amino acid symbols (A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y) and dashes representing missing data.

Table with 10 columns of numerical data, likely representing amino acid sequences or epitope coordinates. The table is split into two main sections, each containing 1000 rows of data.

Table with 1000 rows and 1000 columns of numerical data. Each row contains a sequence of 1000 values, likely representing a specific data series or model output. The values are small, ranging from approximately -0.0001 to 0.0001.

Table with 10 columns: Index, Amino Acid, Methyl Epitope, Linear Epitope, and Index. The table lists various amino acids and their corresponding methyl and linear epitopes, with indices ranging from 664 to 9068.

Table with 100 columns and 1000 rows of numerical data. Each row contains 100 numerical values, likely representing data points for different variables or conditions.

Table with 100 columns and 1000 rows of numerical data. Each row contains 100 numerical values, likely representing data points for different variables or conditions.

Table with 100 columns and 1000 rows of numerical data. Each row contains 100 numerical values, likely representing data points for different variables or conditions.

Table with 100 columns and 1000 rows of numerical data. Each row contains 100 numerical values, likely representing data points for different variables or conditions.

Table with 10 columns: Gene, Allele, Frequency, Gene, Allele, Frequency, Gene, Allele, Frequency, Gene, Allele, Frequency. Contains genetic data for various HLA loci such as HLA-DQA1, HLA-DQB1, HLA-DPA1, HLA-DPB1, HLA-DQA2, HLA-DQB2, HLA-DPA2, HLA-DPB2, HLA-DQA3, HLA-DQB3, HLA-DPA3, HLA-DPB3, HLA-DQA4, HLA-DQB4, HLA-DPA4, HLA-DPB4, HLA-DQA5, HLA-DQB5, HLA-DPA5, HLA-DPB5, HLA-DQA6, HLA-DQB6, HLA-DPA6, HLA-DPB6, HLA-DQA7, HLA-DQB7, HLA-DPA7, HLA-DPB7, HLA-DQA8, HLA-DQB8, HLA-DPA8, HLA-DPB8, HLA-DQA9, HLA-DQB9, HLA-DPA9, HLA-DPB9, HLA-DQA10, HLA-DQB10, HLA-DPA10, HLA-DPB10.

Table with 100 columns and 100 rows of numerical data. The first column contains row indices from 1165 to 1273. The table contains a dense grid of floating-point numbers, likely representing a matrix or a set of data points. The values range from approximately -0.0001 to 0.0001.