

SUPPORTING INFORMATION

Whole-proteome peptide microarrays for profiling autoantibody repertoires within multiple sclerosis and narcolepsy

*Arash Zandian¹, Björn Forsström¹, Anna Häggmark-Månberg¹,
Jochen M. Schwenk¹, Mathias Uhlén¹, Peter Nilsson¹ and Burcu Ayoglu^{1,§,*}*

¹ **Address:** Affinity Proteomics, SciLifeLab, School of Biotechnology, KTH - Royal Institute of Technology, Stockholm, Sweden

§ **Current address:** Department of Medicine, Division of Immunology and Rheumatology, Stanford University School of Medicine, Stanford, California, USA

***Corresponding author:** burcu@kth.se

TABLE OF CONTENTS

Table S-1	S-1
Table S-2	S-2 – S-3
Table S-3	S-4
Figure S-1	S-5 – S-6
Figure S-2	S-7
Figure S-3	S-8
Figure S-4	S-9
Figure S-5	S-10
Figure S-6	S-11

Supplementary Tables

Table S-1. List of the peptide sequences which revealed the highest reactivity frequencies across all samples on the whole-proteome peptide microarrays shown in **Figure 2A**. Peptides which revealed antibody reactivity in nine or more samples are shown. There were two peptides for EGR2. Also note that the CCDS ID for one peptide has been withdrawn and therefore not shown here, but counted in **Figure 2A**. The column “position” indicates the region where the peptide corresponds to in the Uniprot protein sequence.

Gene Names	UniProt ID	ENSG ID	Peptide Sequence	Position	Reactive Samples (n)
EGR2	P11161	ENSG00000122877	QKERKSSAPSA	373-383	10
EGR2	P11161	ENSG00000122877	KSSAPSASVPA	427-437	10
RRP8	O43159	ENSG00000132275	KKCPKKASFASA	73-84	10
CHD7	Q9P2D1	ENSG00000171316	ESEGKGSTFAKA	1501-1512	9
DIP2A	Q14689	ENSG00000160305	PGGRPTTAPSA	187-197	9
DUSP8	Q13202	ENSG00000184545	AALGKQASFSGS	607-618	9
FBXO16	Q8IX29	ENSG00000214050	NSPEEKQSPLSA	187-198	9
GABRA3	P34903	ENSG00000011677	TISKGAAPSASS	415-426	9
GMPPA	Q96IJ6	ENSG00000144591	IHPTAKVAPSA	289-299	9
IFT88	Q13099	ENSG00000032742	TGYGSKTSLASS	67-78	9
IPO5	O00410	ENSG00000065150	ETYENIPGQSKI	49-60	9
KDM2B	Q8NHM5	ENSG00000089094	GRPKGKLGPPASA	583-594	9
MADD	Q8WXG6	ENSG00000110514	KGAREKATPFPS	871-882	9
MAST1	Q9Y2H9	ENSG00000105613	APAKCSAPSSA	1513-1523	9
PIP5K1B	O14986	ENSG00000107242	KTYKKTASSA	19-28	9
PNPLA7	Q6ZV29	ENSG00000130653	QGPSKKPASA	1243-1252	9
RBM20	Q5T481	ENSG00000203867	RLNNSKQGFIGA	367-378	9
SERTAD2	Q14140	ENSG00000179833	SGTASKMAPVSA	265-276	9
SOX8	P57073	ENSG00000005513	KSAPSASASPTTE	319-330	9
TBX21	Q9UL17	ENSG00000073861	SRFYPDLPQAK	367-378	9
TCOF1	Q13428	ENSG00000070814	PTPPGKTGPSA	913-923	9
TNFSF11	O14788	ENSG00000120659	SCRRIKQAFQGA	115-126	9
WDR82	Q6UXN9	ENSG00000164091	QFNPKFMTFASA	289-300	9

Table S-2. Extended list from Table 3, showing all peptide sequences (based on Figure 2B) that revealed differential serum/plasma reactivity frequencies among SPMS and narcolepsy sample collections. The column “position” indicates the region where the peptide corresponds to in the Uniprot protein sequence. “Narcolepsy”, “SPMS” and “CIS” stand for the number of reactive samples within each patient group out of a total of 7, 8, and 8 samples, respectively.

Differential peptide reactivity between narcolepsy vs. SPMS										
Gene	UniProt	ENSG	Peptide sequence	Position	Narcolepsy	SPMS	CIS	p-val#	p-valS	p-val§
BCL6	P41182	ENSG00000113916	EGLKPAAPSA	23-32	7	2	4	0.01	0.08	0.61
CNTN2	Q02246	ENSG00000184144	RDATKITLAPSS	506-517	7	2	4	0.01	0.08	0.61
DYSF	O75923	ENSG00000135636	FPDPYTELNTGK	536-547	5	0	2	0.01	0.13	0.47
NRXN1	Q9ULB1	ENSG00000179915	LHTGKSADYVNL	355-366	5	0	0	0.01	0.01	1.00
RXRG	P48443	ENSG00000143171	SMSPSAALSTGK	26-37	5	0	1	0.01	0.04	1.00
SCN5A	Q14524	ENSG00000183873	AMKKLGSKKPQK	1497-1508	5	0	0	0.01	0.01	1.00
ABL2	P42684	ENSG00000143322	TQEGGKKAALGA	118-129	6	1	5	0.01	0.57	0.12
RTN4	Q9NQC3	ENSG00000115310	MEAPLNSAVPSA	616-627	6	1	2	0.01	0.04	1.00
SOCS7	O14512	ENSG00000274211	SSPGRGGGGGGR	12-23	6	1	5	0.01	0.57	0.12
SOX11	P35716	ENSG00000176887	GGGAGGGAGGA	92-102	3	8	8	0.03	0.03	1.00
GOLGA6L10	A6NI86	ENSG00000278662	VALDSSSAISQ	96-107	0	5	3	0.03	0.20	0.62
MAP3K7	O43318	ENSG00000135341	SFMDIASTNTSN	65-76	0	5	2	0.03	0.47	0.31
PDE8B	O95263	ENSG00000113231	FNRRFMENSSII	234-245	0	5	3	0.03	0.20	0.62
PLCG2	P16885	ENSG00000197943	IIKHKKLGPRG	27-38	0	5	1	0.03	1.00	0.12
SETX	Q7Z333	ENSG00000107290	DEEMSNSTSVI	924-934	0	5	0	0.03	1.00	0.03
SYNJ2	O15056	ENSG00000078269	PYNVKQIKTTNA	30-41	0	5	2	0.03	0.47	0.31
ADRBK2	P35626	ENSG00000100077	EARKRAKNKQLG	12-23	4	0	1	0.03	0.12	1.00
BCL9L	Q86UU0	ENSG00000186174	PLGSNIPLHPNA	1168-1179	4	0	2	0.03	0.31	0.47
CCDC6	Q16204	ENSG00000108091	GGGGGGKSG	33-41	4	0	1	0.03	0.12	1.00
DNAJC2	Q99543	ENSG00000105821	KSTGGGGNGSKN	442-453	4	0	0	0.03	0.03	1.00
DYSF	O75923	ENSG00000135636	PDPYTELNTGKG	537-548	4	0	1	0.03	0.12	1.00
ENO1	P06733	ENSG00000074800	HINKTIAPALVS	68-79	4	0	0	0.03	0.03	1.00
ERP29	P30040	ENSG00000089248	EELQKSLNILTA	239-250	4	0	2	0.03	0.31	0.47
GRIN2A	Q12879	ENSG00000183454	FSPVSSKLSGKK	1329-1340	4	0	0	0.03	0.03	1.00
GRIN2B	Q13224	ENSG00000273079	PTNSKAQKKNRN	1285-1296	4	0	0	0.03	0.03	1.00
GRIN2B	Q13224	ENSG00000273079	KSSVPTAGHH	1353-1362	4	0	0	0.03	0.03	1.00
HIPK2	Q9H2X6	ENSG00000064393	VYSQSKNIPLSQ	38-49	4	0	2	0.03	0.31	0.47
NEK10	Q6ZWH5	ENSG00000163491	LPNKQKNAAKSN	85-96	4	0	3	0.03	0.62	0.20
NRG1	Q02297	ENSG00000157168	SSTSTSTTGTS	167-178	4	0	2	0.03	0.31	0.47
NRXN1	Q9ULB1	ENSG00000179915	GLMLHTGKSADY	352-363	4	0	1	0.03	0.12	1.00
P4HA1	P13674	ENSG00000122884	GNLKYFEYIMAK	65-76	4	0	0	0.03	0.03	1.00
PRKG2	Q13237	ENSG00000138669	REQLSKQTVAIA	9-20	4	0	2	0.03	0.31	0.47
RAD23B	P54727	ENSG00000119318	TSSTTTTVAQA	100-110	4	0	1	0.03	0.12	1.00
RARG	P13631	ENSG00000172819	DDSSQPGHPNA	13-24	4	0	2	0.03	0.31	0.47
RGS12	O14924	ENSG00000159788	VVEEMQSGGIFN	137-148	4	0	3	0.03	0.62	0.20
SLC9A3R1	O14745	ENSG00000109062	SPPKQDSTAPSS	81-92	4	0	1	0.03	0.12	1.00

SPTAN1	Q13813	ENSG00000197694	IQGVLDTGKKLS	1820-1831	4	0	1	0.03	0.12	1.00
SPTAN1	Q13813	ENSG00000197694	AIQGVLDTGKKL	1819-1830	4	0	2	0.03	0.31	0.47
SPTAN1	Q13813	ENSG00000197694	EPAIQGVLDTGK	1817-1828	4	0	1	0.03	0.12	1.00
SPTAN1	Q13813	ENSG00000197694	LDTGKKLSDDNT	1824-1835	4	0	0	0.03	0.03	1.00
UBE2E2	Q96LR5	ENSG00000182247	PKKKEGKISSKT	37-48	4	0	0	0.03	0.03	1.00
USP8	P40818	ENSG00000138592	PQIDRTKKPAVK	405-416	4	0	0	0.03	0.03	1.00
VIL1	P09327	ENSG00000127831	SQITAEVTSPOV	739-750	4	0	1	0.03	0.12	1.00
SLC4A5	Q9BY07	ENSG00000188687	GGAGSGGAGGTS	454-465	2	7	5	0.04	0.31	0.57
ATP10A	O60312	ENSG00000206190	ASMLSKHTAFSS	50-61	6	2	1	0.04	0.01	1.00
BPTF	Q12830	ENSG00000171634	TSAPGRGGRGGG	81-92	6	2	4	0.04	0.28	0.61
BRD2	P25440	ENSG00000204256	KKGAKLAALQGS	19-30	6	2	3	0.04	0.12	1.00
CCDC6	Q16204	ENSG00000108091	SSTSGGGGGG	28-37	6	2	5	0.04	0.57	0.31
EGR1	P18146	ENSG00000120738	GGGGGGGSN	70-78	6	2	3	0.04	0.12	1.00
FUS	P35637	ENSG00000089280	QYNSSSGGGG	75-84	6	2	6	0.04	1.00	0.13
HNRNPA1L2	Q32P51	ENSG00000139675	FGNDGSNFGGGG	247-258	6	2	2	0.04	0.04	1.00
NEFM	P07197	ENSG00000104722	AAEQNKPAIRSA	302-313	6	2	4	0.04	0.28	0.61
NFKB1	P19838	ENSG00000109320	FGSGGGGGGTG	389-399	6	2	4	0.04	0.28	0.61
RIPK1	Q13546	ENSG00000137275	NTEGKGTAAYSSA	126-137	6	2	4	0.04	0.28	0.61
RNF115	Q9Y4L5	ENSG00000265491	VTDDSSFLGGGG	49-60	6	2	4	0.04	0.28	0.61
STK19	P49842	ENSG00000204344	ANPSRGGGG	14-22	6	2	4	0.04	0.28	0.61
ADGRG6	Q86SQ4	ENSG00000112414	NSPSTTPPTVT	383-393	5	1	2	0.04	0.13	1.00
BNC1	Q01954	ENSG00000169594	QNPNLHKSLASS	978-989	5	1	4	0.04	0.61	0.28
ENO1	P06733	ENSG00000074800	DWGAWQKFTASA	300-311	5	1	3	0.04	0.31	0.57
FASTKD5	Q7L8L6	ENSG00000215251	GLEFSKTSSSKA	77-88	5	1	1	0.04	0.04	1.00
KIAA1462	Q9P266	ENSG00000165757	REDNPKGRQAAR	23-34	5	1	3	0.04	0.31	0.57
PRMT5	O14744	ENSG00000100462	VLGAGRGPLVNA	19-30	5	1	2	0.04	0.13	1.00
SIPA1	Q96FS4	ENSG00000213445	PAFGPALPAGGG	484-495	5	1	3	0.04	0.31	0.57
UBE2E2	Q96LR5	ENSG00000182247	KKKEGKISSKTA	38-49	5	1	4	0.04	0.61	0.28
USP32	Q8NFA0	ENSG00000170832	VLNGGKYSFGTA	8-19	5	1	1	0.04	0.04	1.00

p-value from Fisher's exact test comparing narcolepsy vs. SPMS

\$ p-value from Fisher's exact test comparing narcolepsy vs. CIS

§ p-value from Fisher's exact test comparing SPMS vs. CIS

Table S-3. Number of different single peptide epitopes, consecutive peptide epitopes, and differences in reactivity per protein among the different sample groups on the targeted design arrays.

	Single Peptide Epitopes	Consecutive Peptide Epitopes	Proteins
Narcolepsy vs. SPMS	65	5	41
Narcolepsy vs. CIS	45	6	36
Narcolepsy vs. SPMS+CIS	202	11	37
SPMS vs. CIS	9	2	15
Total (unique)	321 (244)	24 (16)	129 (88)

Supplementary Information

Potential zone variation is an important parameter to address when utilizing the planar peptide arrays. We eliminated the impact of any potential zone variation by taking into account the local spot background for the analysis of individual spots, as described in the Materials & Methods section. Below we provide an analysis demonstrating our strategy, where two different representative samples from the SPMS and CIS patient groups were analyzed on the opposite corners on the targeted 174k peptide microarrays.

Figure S-1 shows the data for the first sample belonging to an SPMS patient, which was analyzed on the top left corner of the targeted 174k peptide microarray.

Panel A illustrates the local MFI of the entire subarray, which was calculated by using a sliding window of 7x7 spots/position, i.e. 25 or 24 peptide-containing spots (because of the checkerboard pattern the peptides are synthesized in). The column position and row position of the subarray are shown in panel B and C. As can be seen in Panel A, there are some regions with higher local MFI (darker color). Yet, the distribution of the number of reactive peptide spots (passing the filtering criterion we describe in the Materials & Methods section) is even. Panel B shows the density distribution of reactive peptide spots (black line) per column position on the subarray, and the density seems to be rather stable in all column positions, which is expected since the spatial peptide positions are randomly assigned when the peptide microarrays are synthesized. The median local MFI calculated on the median per column position are shown as filled points (Panel B), where a slight fluctuation around column position 350 can be seen. However, the number of reactive peptide spots are fairly stable (black line).

Similarly, Panel C shows the density distribution of reactive peptide spots (black line) per row position on the subarray, where the local MFI decreases from the higher row positions (800) towards the lower row (0) positions, i.e. there is a gradient. Yet, the distribution of the reactive spots passing the filtering criterion are rather stable (black line) even though the median local MFI is varying (up to 25% from the top to bottom, ~500 MFI vs. 400 MFI). Finally, one can see that we find slightly more reactive spots in areas where the background is lower, which could be expected since lower signals can more easily be distinguished from noise. To support the conclusions which can be drawn by this analysis, we also included a **Figure S-2** below, representing the same analysis for a CIS sample which was screened on the bottom right corner of the same 174k peptide microarray slide. These two figures demonstrate that when we use the local background MFI, as well as spot morphology to assess the peptide spot reactivity, our downstream data analysis is not susceptible to differences in zone variation in and between subarrays.

Supplementary Figures

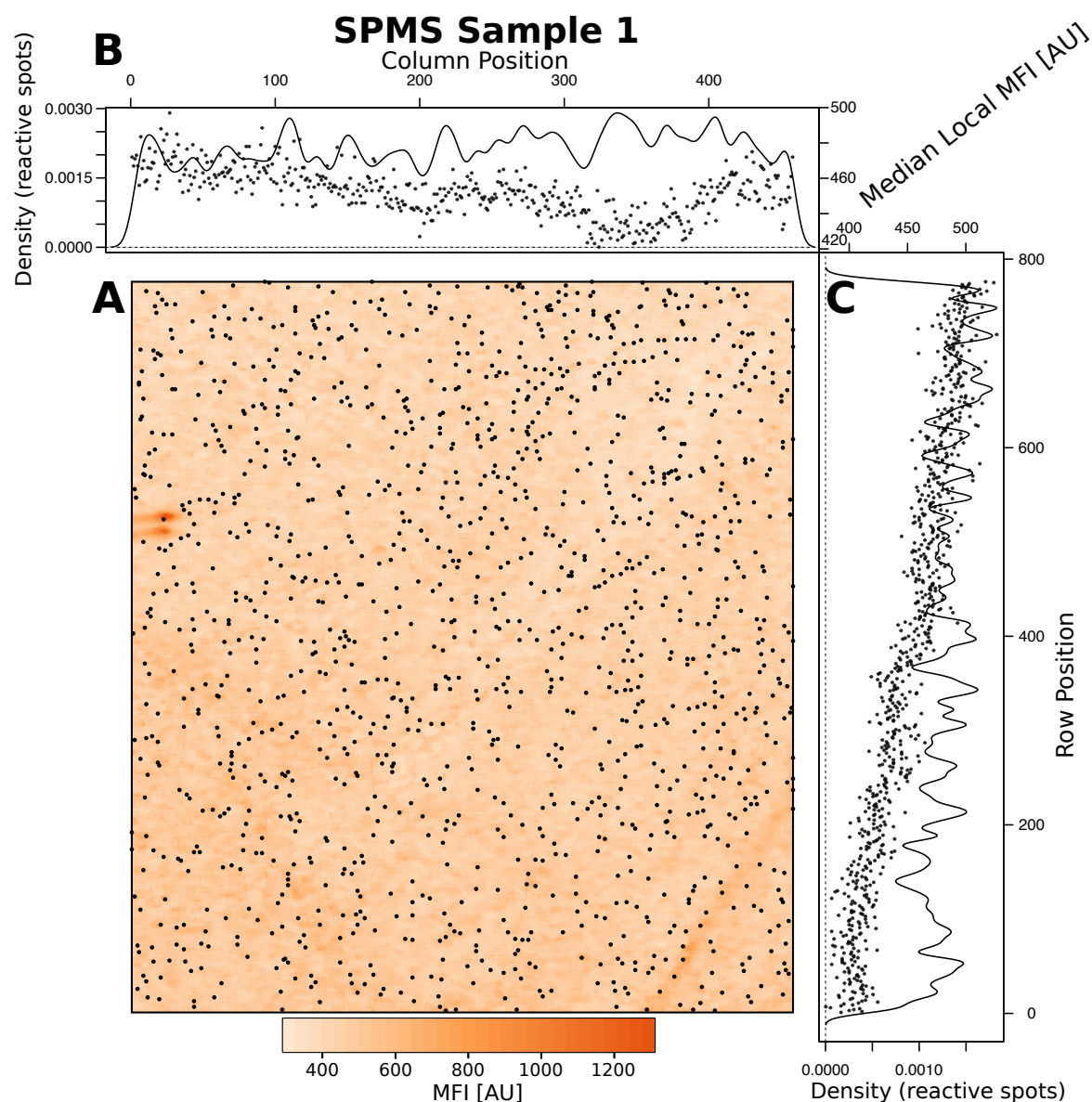


Figure S-1. Signal intensity gradient and peptide selection map for a subarray, where an SPMS sample was analyzed. An SPMS patient sample was analyzed on the 174k peptide microarrays, **A**) where the entire subarray was colored by the median signal intensity in a 7x7 spots sliding window with signals higher than the local background shown as filled black dots. The density of signal intensities above the local background, which have been considered as positive reactivity, are shown in density line graphs as a function of column or row position (**B** and **C**). The median local MFI per column position and row positions are also shown as filled black dots in **B** and **C**.

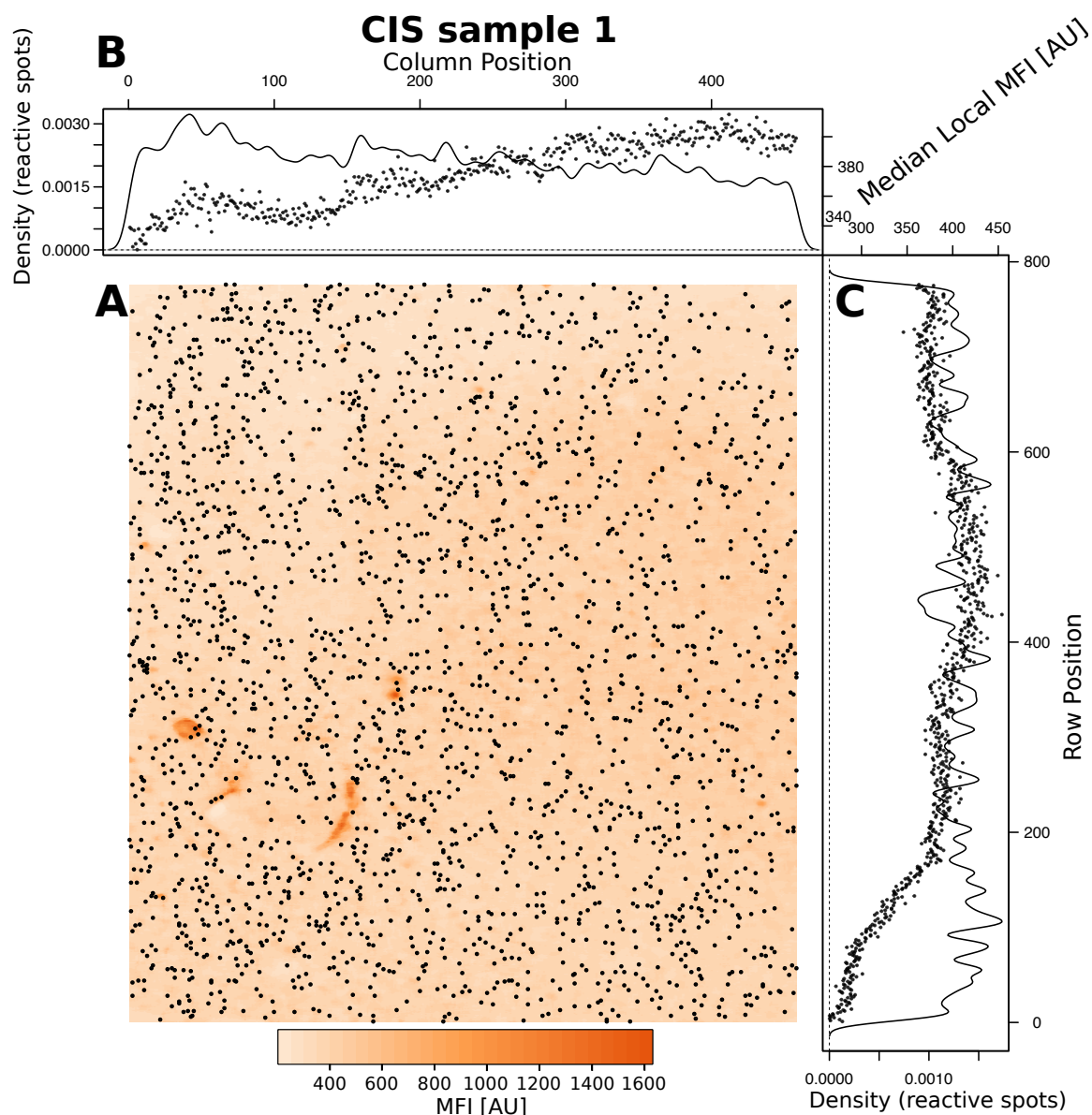


Figure S-2. Signal intensity gradient and peptide selection map for a subarray, where a CIS sample was analyzed. A CIS patient sample was analyzed on the 174k peptide microarrays, **A)** where the entire subarray was colored by the median signal intensity in a 7x7 spots sliding window with signals higher than the local background shown as filled black dots. The density of signal intensities above the local background, which have been considered as positive reactivity, are shown in density line graphs as a function of column or row position (**B** and **C**). The median local MFI per column position and row positions are also shown as filled black dots in **B** and **C**.

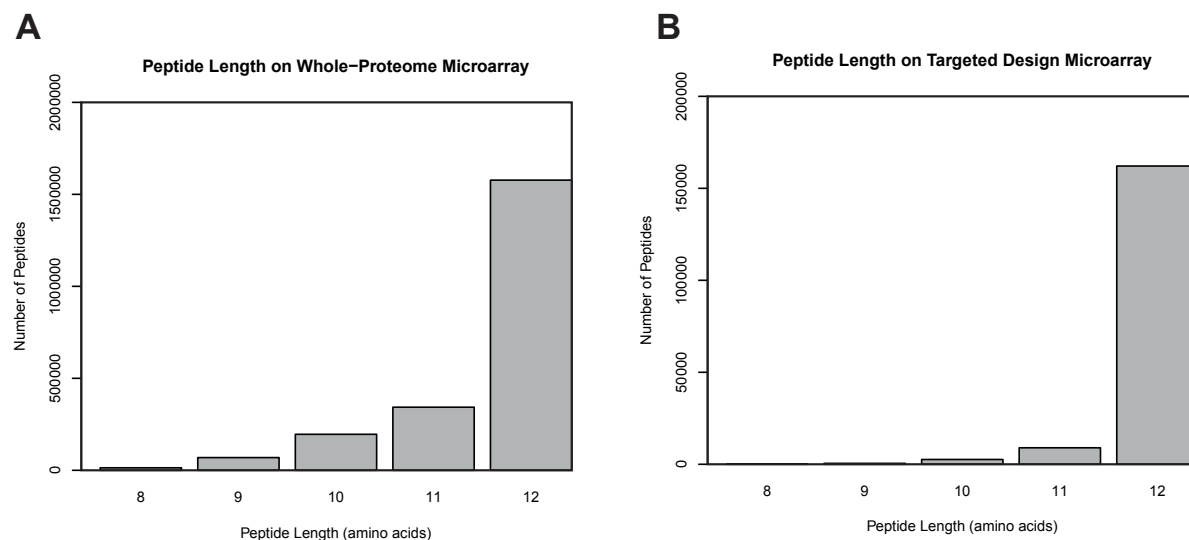


Figure S-3. Distribution of peptide length on the planar microarrays. The peptide length distribution of all the peptides synthesized on the whole-proteome peptide microarray (**A**) the targeted design microarray (**B**). Due to maximum synthesis cycles in the generation of the planar peptide microarrays, not all peptides could be synthesized as 12-mers.

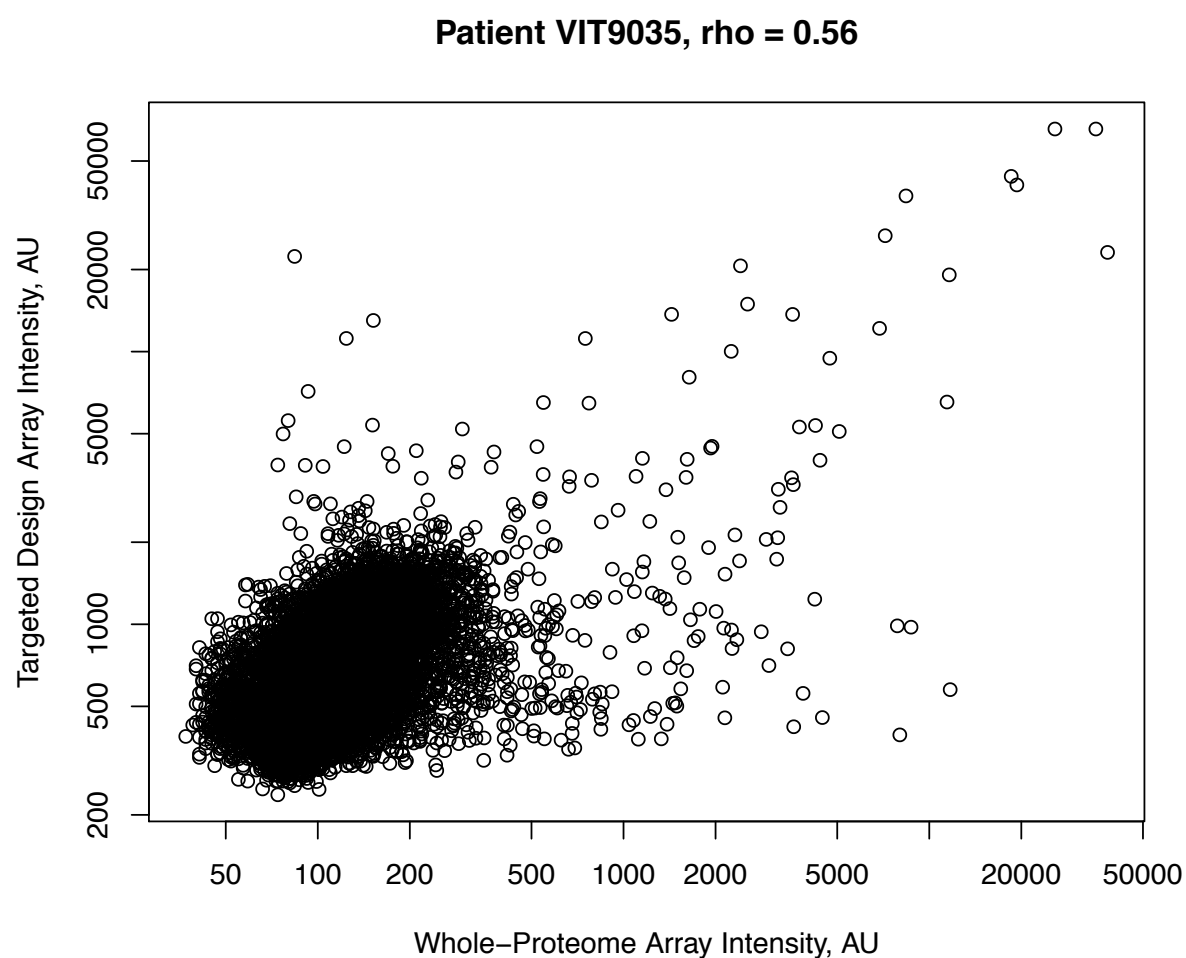


Figure S-4. Data correlation between whole-proteome and targeted design planar microarrays. Correlation of the antibody reactivity in an individual narcolepsy sample towards 23,072 peptides, which overlapped between the whole-proteome microarray and the array with targeted design. Each dot represents a peptide and the Spearman's rho is indicated.

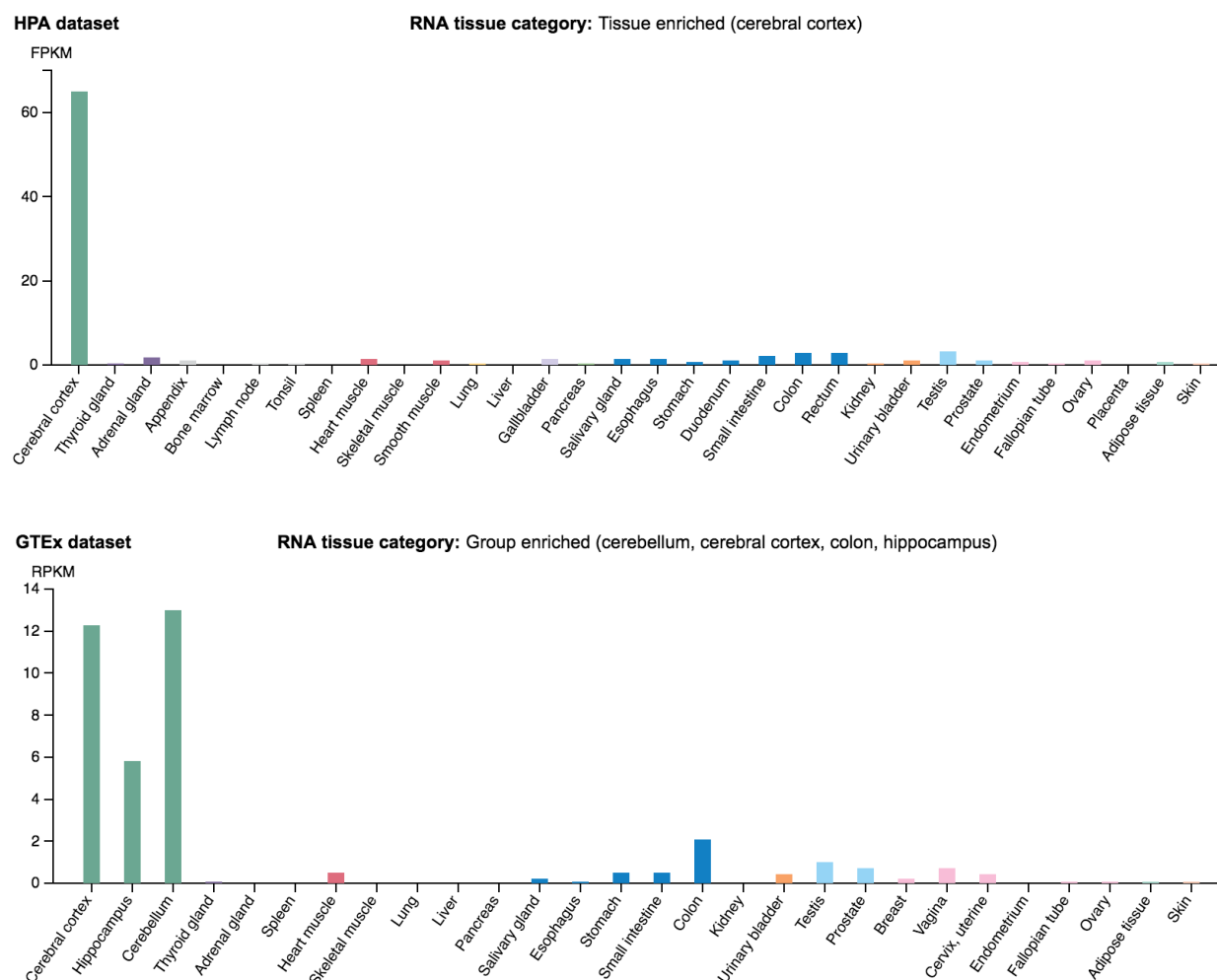


Figure S-5. Tissue-specific transcriptome data available for NRXN1 in two different publicly available data portals. Data available for RNA expression of NRXN1 in various human tissue types by the Human Protein Atlas portal (www.proteinatlas.org) and the GTEx Portal (www.gtexportal.org).

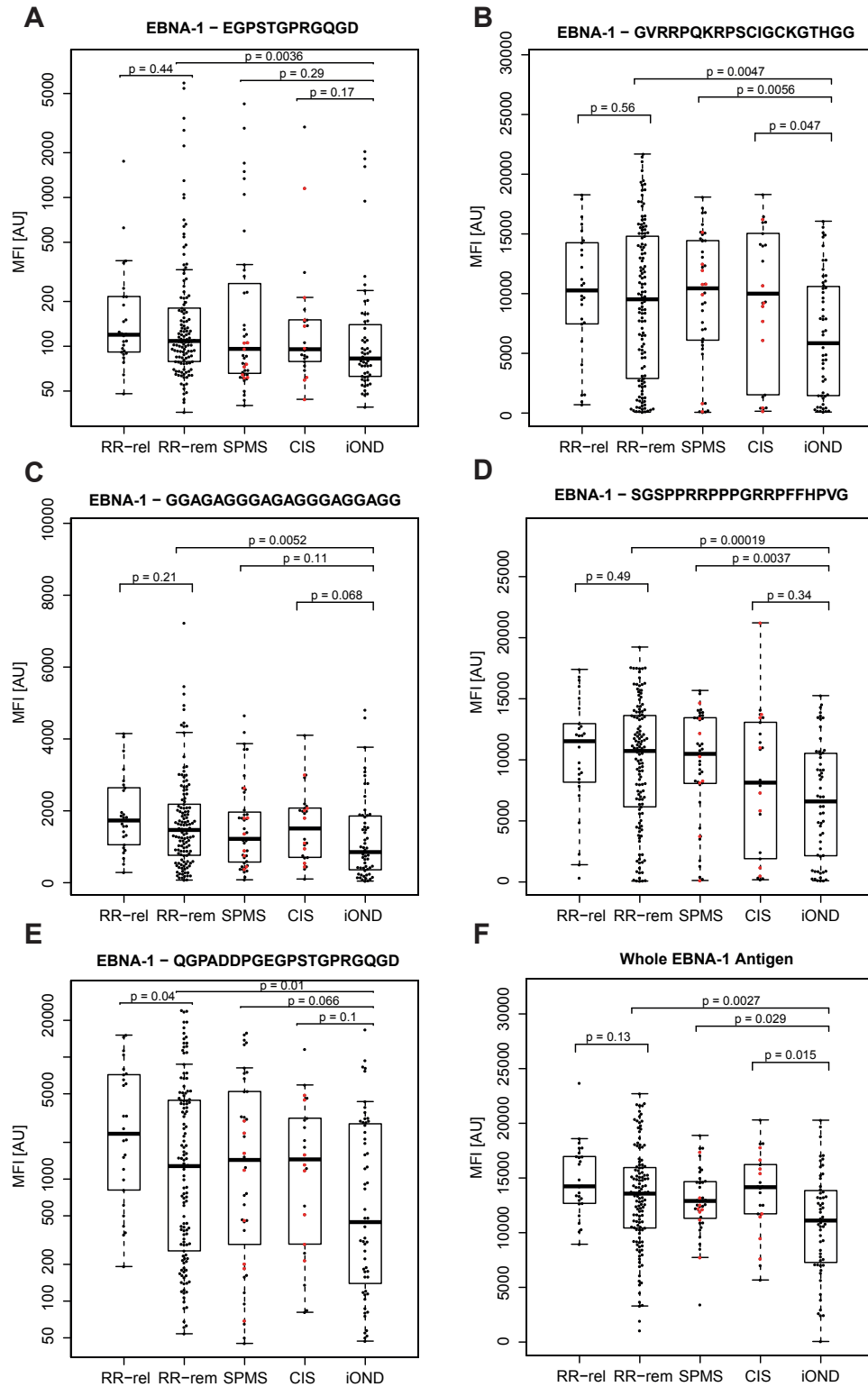


Figure S-6. Antibody reactivity towards six different EBNA-1 protein and peptides within MS-related sample set. The peptide EGPSTGPRGQGD (**A**) represents the region 444-455 of EBNA-1, the peptide GVRRPQKRPSICGCKGTHGG (**B**) represents the region 69-80 of EBNA1, the peptide GGAGAGGGAGAGGGAGG (**C**) represents the region 122-141 of EBNA1, the peptide SGSPRRPPPGRRPFFHPVG (**D**) represents the region 391-410 of EBNA1, the peptide QGPADDPGEGPSTGPRGQGD (**E**) represents the region 436-455 of EBNA1. In addition, antibody reactivity towards the full-length EBNA-1 protein was analyzed (**F**).