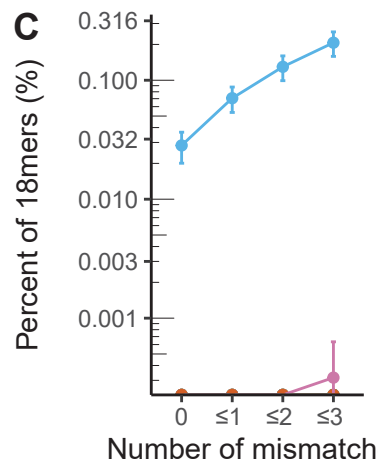
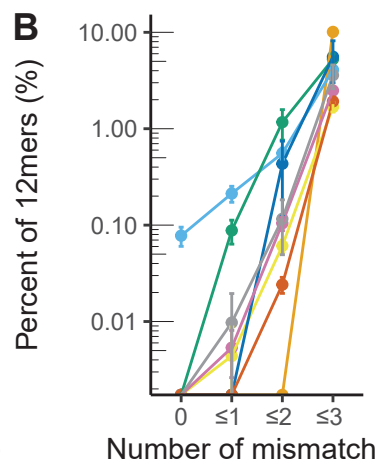
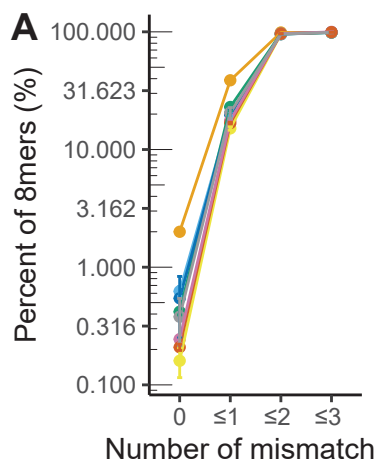


Figure S1 (Associated with Figure 1) – Evaluation of mimicry candidates by Baltimore classification. Percent of **A)** 8mer, **B)** 12mers, and **C)** 18mers with 0, ≤ 1 , ≤ 2 , or ≤ 3 mismatches by Baltimore genome classification.

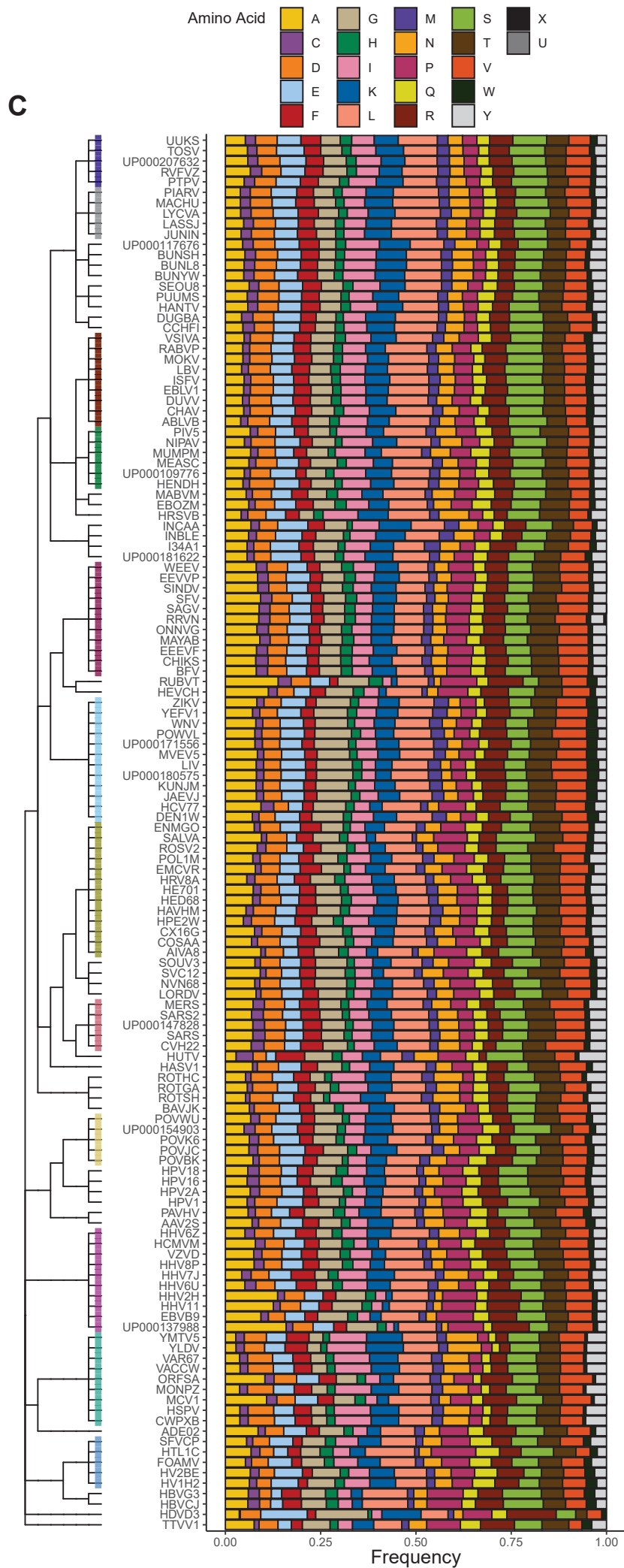
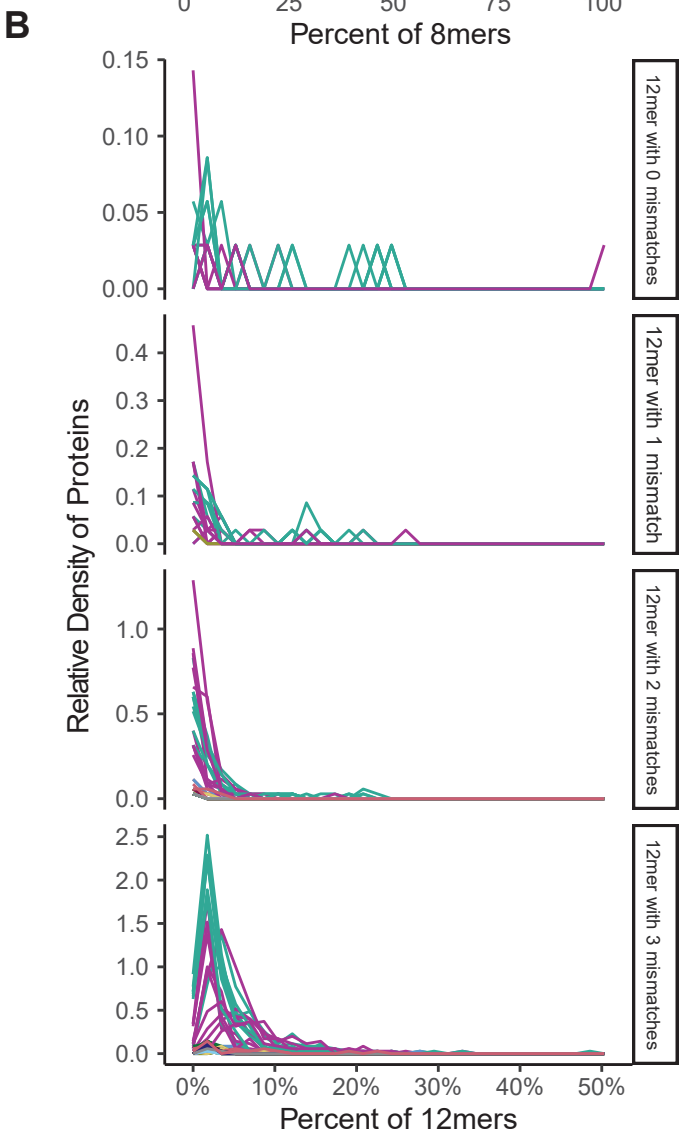
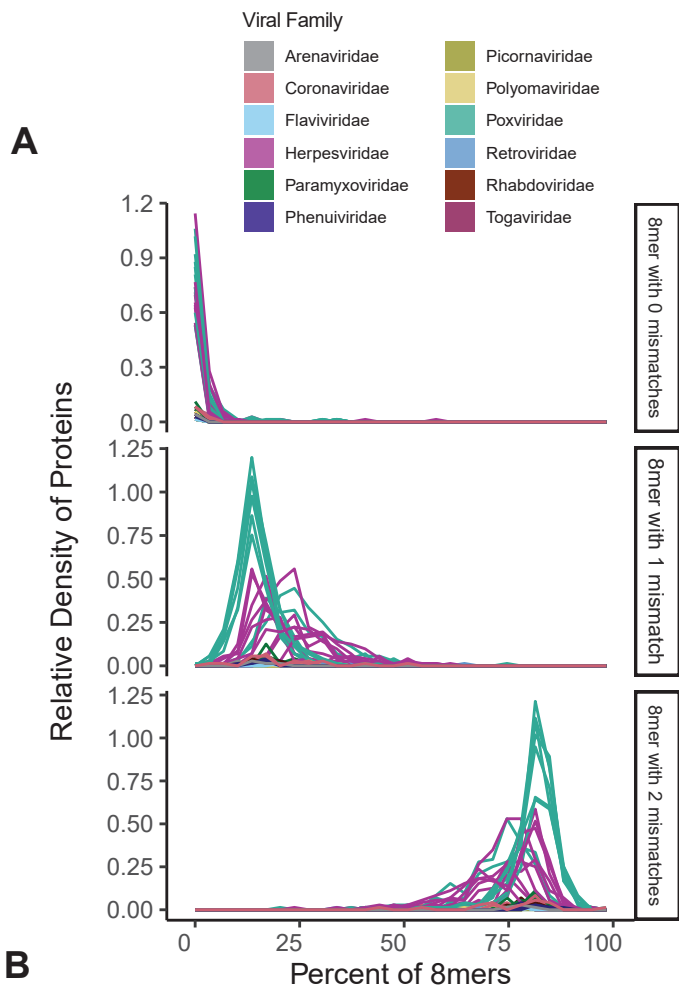


Figure S2 (Associated with Figure 1) – Molecular mimicry is dispersed through the viral proteome. For each virus (a line), the density of proteins that have varying percentages of either **A)** 8mers with 0, 1, or 2 mismatches or **B)** 12mers with 0, 1, 2, or 3 mismatches is plotted. Color indicates the family of the virus. As shown by the frequency plot peaking to the left, viruses overall display a pattern of low mimicry rates in any given protein. **C)** The relative proportion of amino acid usage for each virus organized by viral phylogeny. “U” indicates selenocysteine and “X” indicates an unknown amino acid.

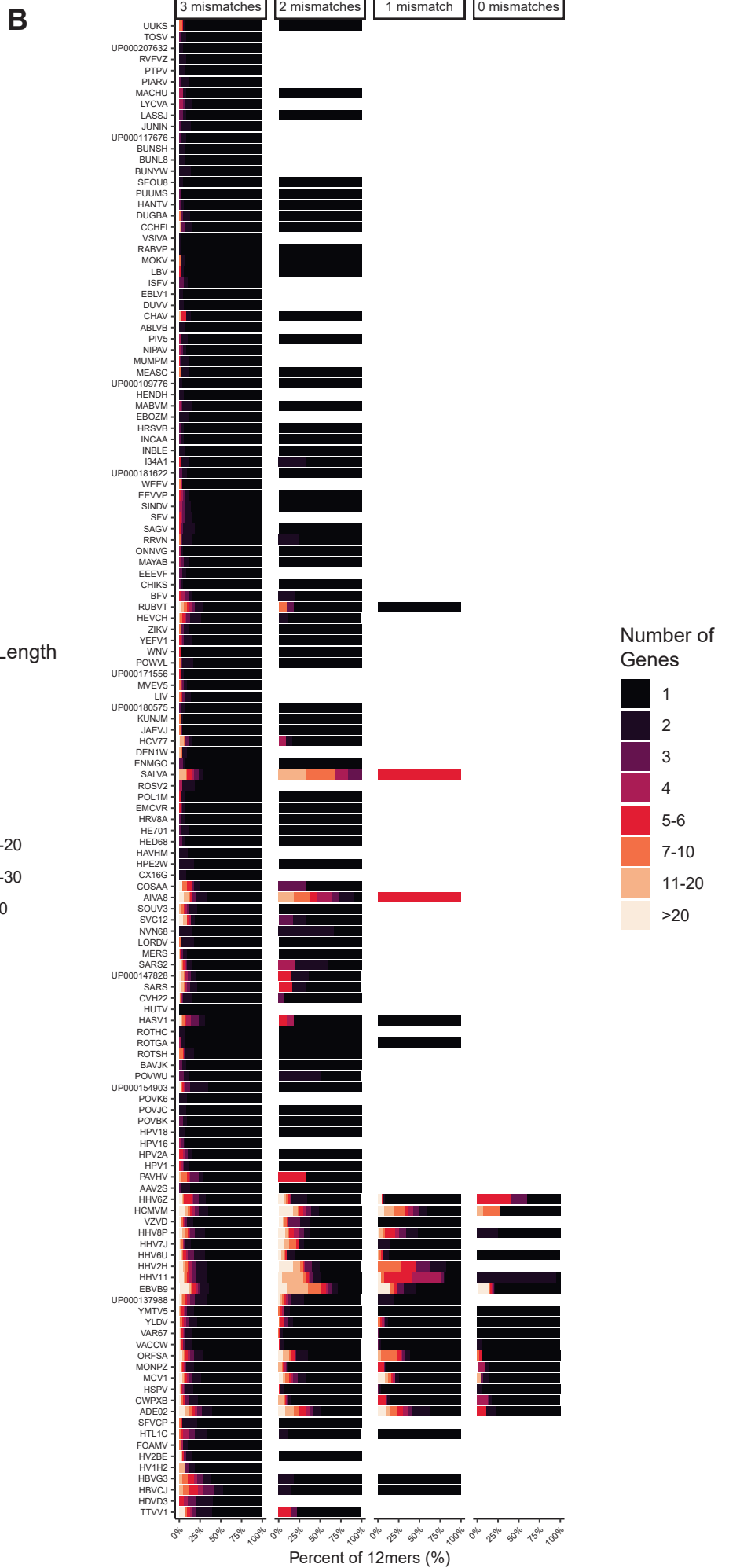
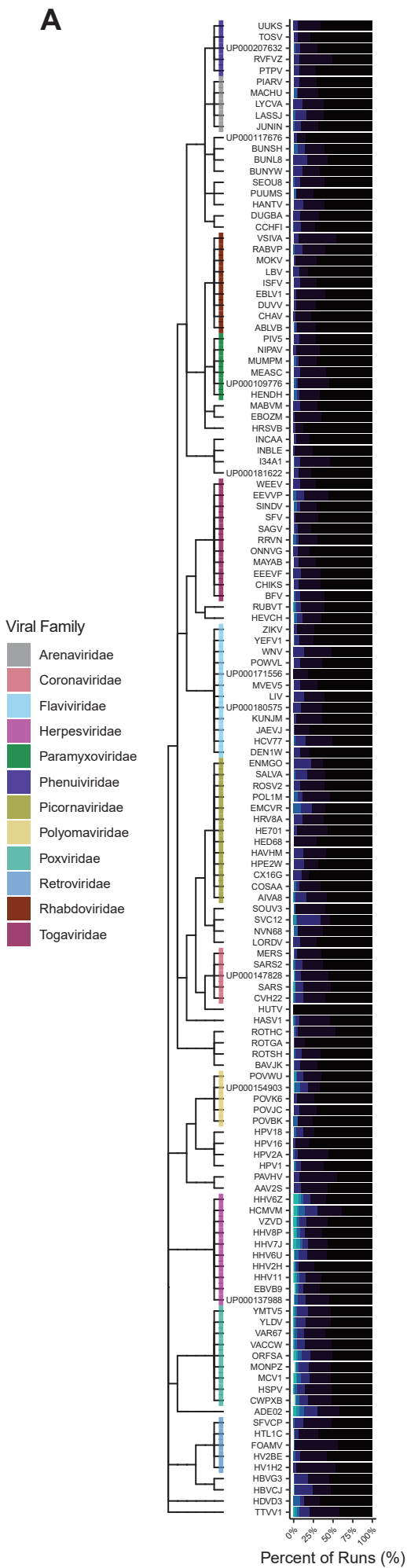


Figure S3 (Associated with Figure 3) – Phylogenetic analysis of run length and multi-mapping. A) Percent of k-mers at various lengths for all viruses in the cohort organized by viral phylogeny. **B)** Percent of viral 12mers and their corresponding number of mimicked human genes under mismatch conditions of 0, ≤ 1 , ≤ 2 , and ≤ 3 mismatches.

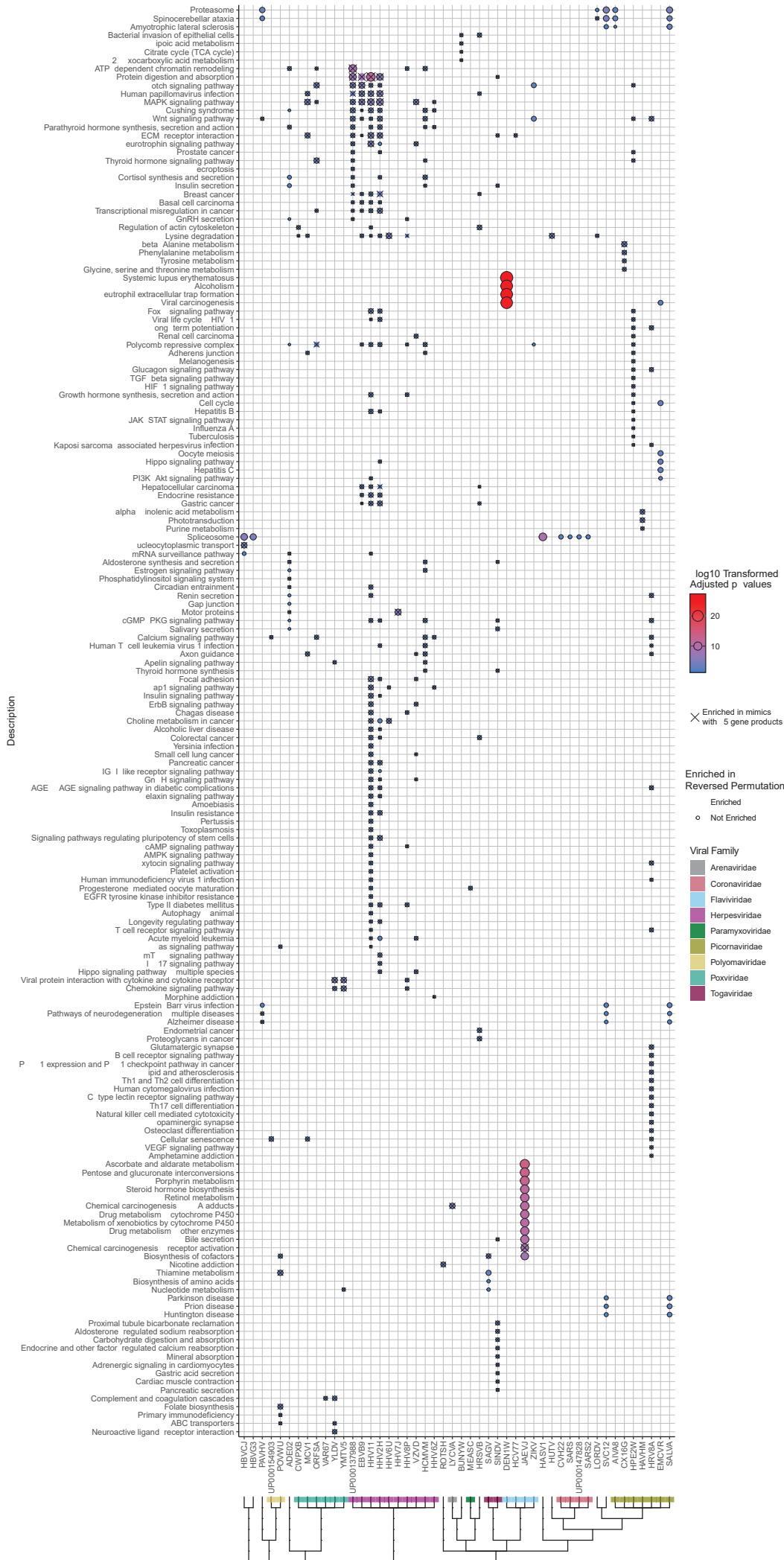


Figure S4 (Associated with Figure 4) – Complete enrichment results for all viruses.

Hypergeometric enrichment testing of KEGG pathways for human proteins that are mimicked by each virus. Significant enrichment is displayed as a dot and is outlined if that biological enrichment was not observed in the reverse proteome permutation (permutation 2). Viruses are organized by their phylogeny.

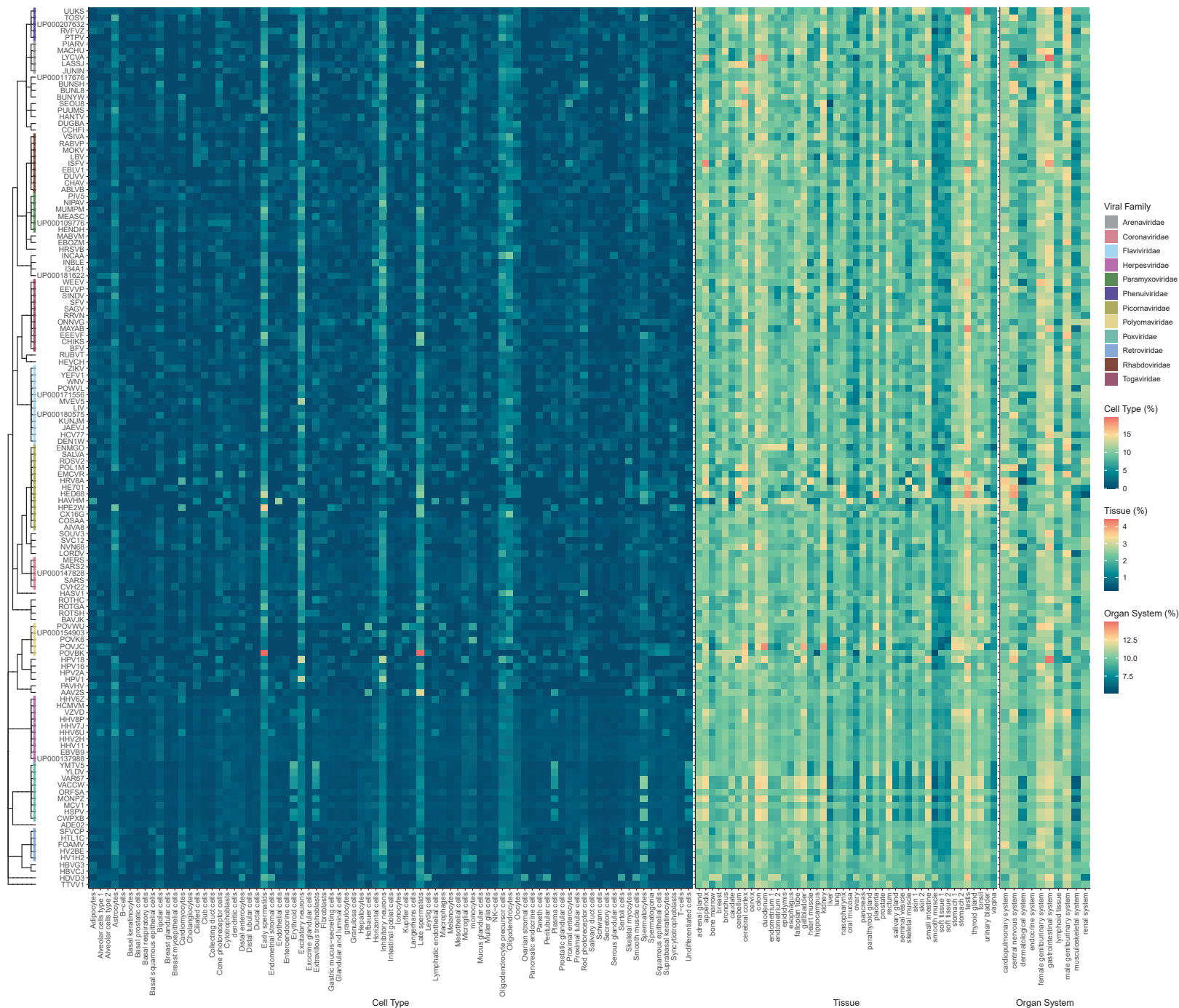


Figure S5 (Associated with Figure 4) – Cell type, tissue, and organ system percentages of 12mers. Heatmap of the percent of 12mer mimics with 3 or less mismatches that are known to be expressed in varying cell types, tissues, and organ systems as reported in the Human Protein Atlas. .

Viral Family

- Arenaviridae
- Coronaviridae
- Flaviviridae
- Herpesviridae
- Paramyxoviridae
- Phenuiviridae
- Picornaviridae
- Polyomaviridae
- Poxviridae
- Retroviridae
- Rhabdoviridae
- Togaviridae

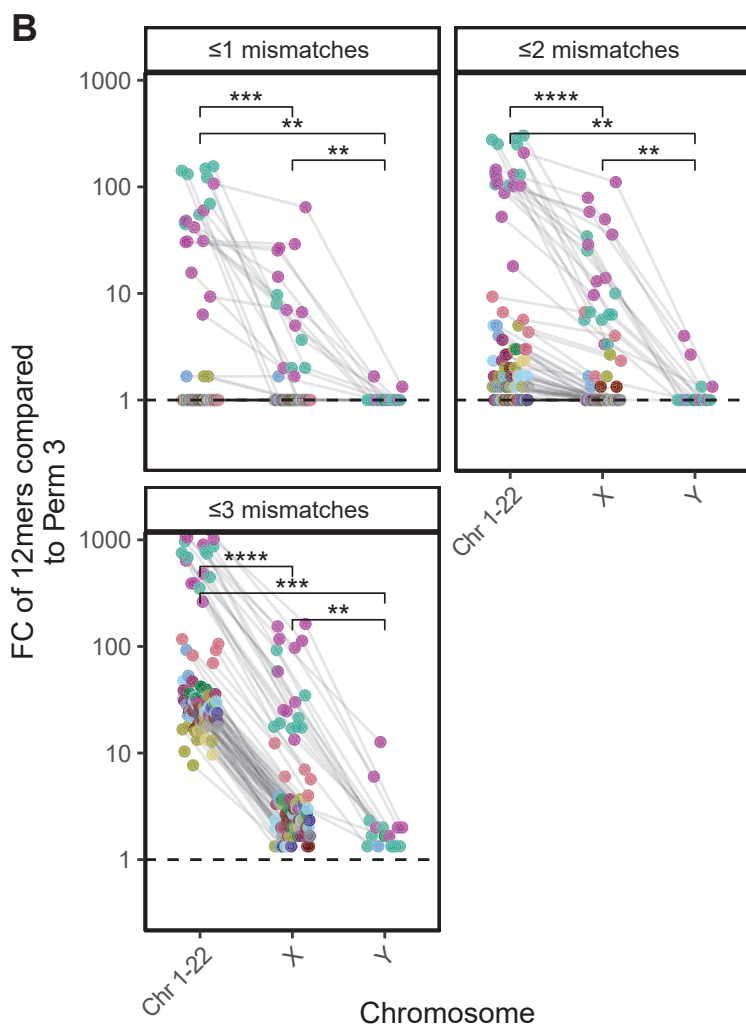
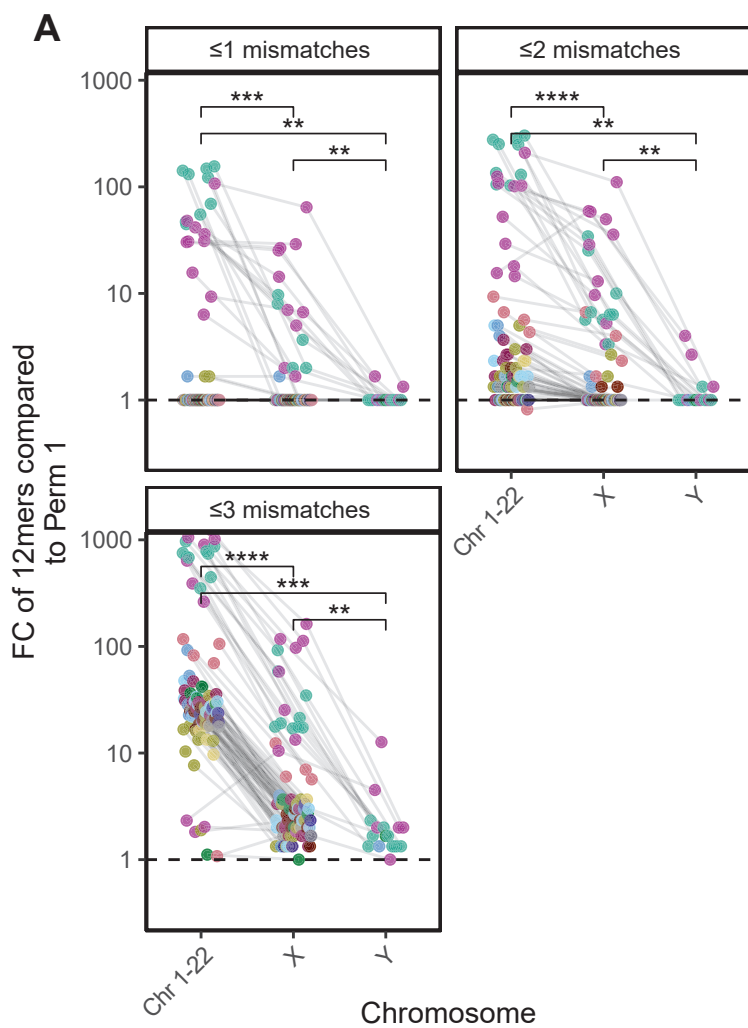


Figure S6 (Associated with Figure 4) – Chromosome enrichment compared to A) permutation 1 and B) permutation 3. Fold change of the percent of mimics whose human counterpart is encoded on either an autosome (Chromosomes 1-22), X, or Y chromosome over the rate in the **A)** randomly shuffled proteome (permutation 1) and **B)** the AA class shuffled proteome (permutation 3). Wilcoxon summed-rank test used for paired pairwise comparison. All p values adjusted for multiple hypothesis testing using Benjamini-Hochberg corrections (* p.adj $\leq$ 0.05, ** p.adj $\leq$ 0.01, *** p.adj $\leq$ 0.001, **** p.adj $\leq$ 0.0001).

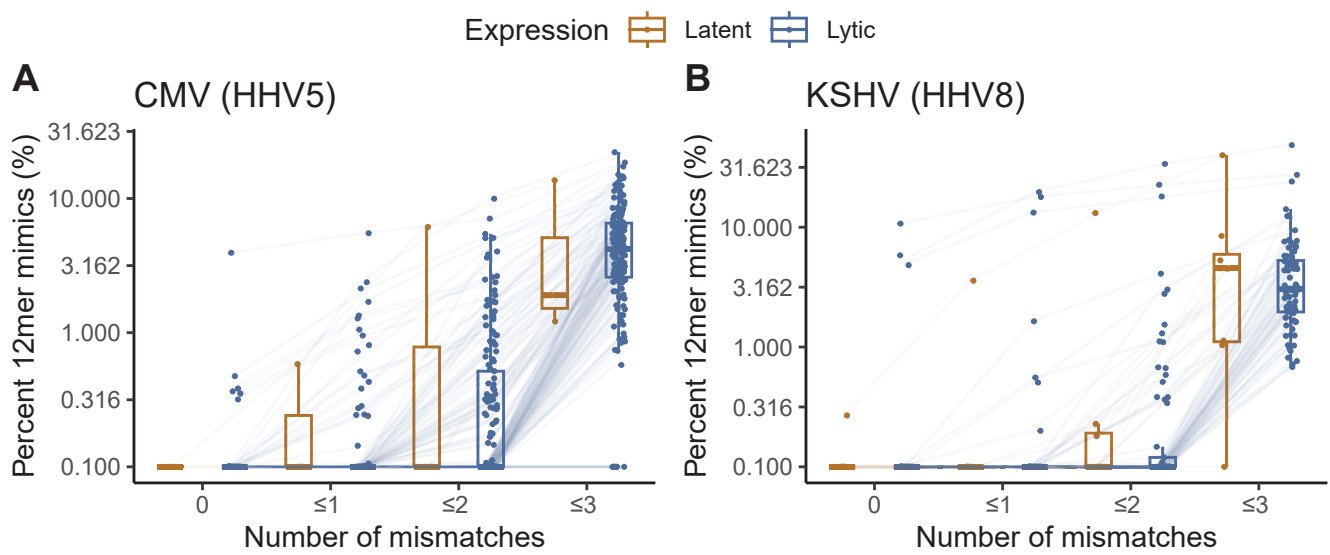


Figure S7 (Associated with Figure 5) – Percent of 12mers with 0, ≤ 1 , ≤ 2 , and ≤ 3 mismatches in latent and lytic **A)** CMV, and **B)** HHV8 proteins. Each point denotes a single protein connected across the mismatch levels with boxplots denoting the interquartile range. Wilcoxon summed-rank test used for paired pairwise comparison. All p values adjusted for multiple hypothesis testing using Benjamini-Hochberg corrections (* p.adj ≤ 0.05 , ** p.adj ≤ 0.01 , *** p.adj ≤ 0.001).

Post-Diagnosis

Pre-Diagnosis

IC Cluster

Other MS Associated Auto-Antibodies

TRIO_395474
LOC100652901_49861
SRSF4_682210
USP31_222213
MAP3K12_274862
MAP3K12_274437
SRSF7_331013
SRSF7_331012
CLASRP_441187
SRRM3_613856
KRT75_287686
SRRM3_613857
ZRANB2_281628
TRA2B_456817
NKTR_139361
RIMS2_627827
NKTR_139346
SRSF4_342020
SRSF4_342019
USP31_222214
SH3BP2_216602
RBMX1B_64485
SRSF4_682211
NKTR_139345
RIMS2_337979
DENND4C_307092
CLASRP_13378
C5orf60_596251
EXO1_90756
ZNF764_541112
SRSF7_101254
SRRM3_558302
MAP3K12_274863
KRT75_287685
NKTR_139362
CLASRP_13379
PPIG_130031
RIMS2_306096
TRA2B_456818
CHERP_554788
MAP3K12_274436
CHERP_554789
SRSF1_550431
SRSF1_550429

PHACTR4_682453
ABCA7_1900
NRXN2_428754
MYO7A_657421
H1FNT_366975
CAMK2N1_536454
ANKRD35_495219
TFAP2A_601685
OBSCN_88624
PITPNM2_537816
WNK1_461031
TFDP2_156590
NRAP_526548
STARD9_701718
CBX4_443152
TAF2_626658
ITIH6_513562
KMT2C_616997
STARD9_701717
ARHGAP31_155813
FRMPD1_312343
PXDNL_293885
ANK2_171146

FC over
Healthy Controls100
50
0

Figure S8 (Associated with Figure 6) – Additional MS associated auto-antibodies outside the IC cluster from Zamecnik et al.²⁶. Heatmap of fold change for auto-antibodies in MS patients before and after diagnosis over healthy control levels reveals that MS auto-antibodies outside the IC cluster identified in Zamecnik et al.²⁶ are not concordant.