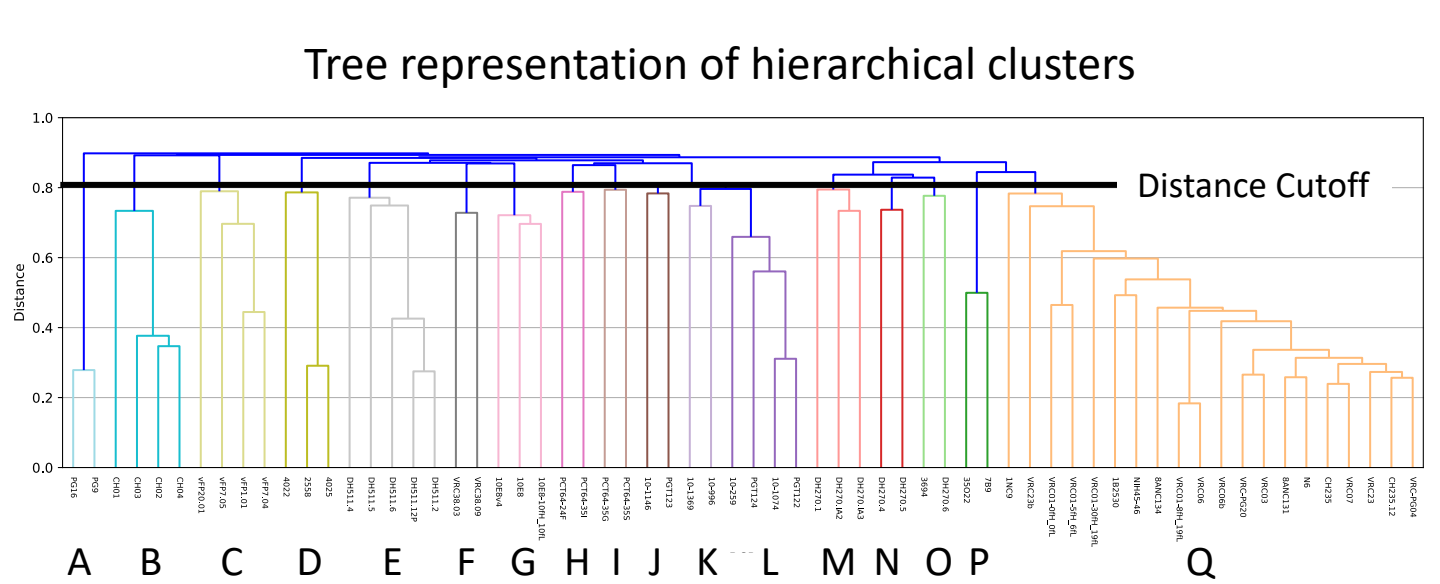
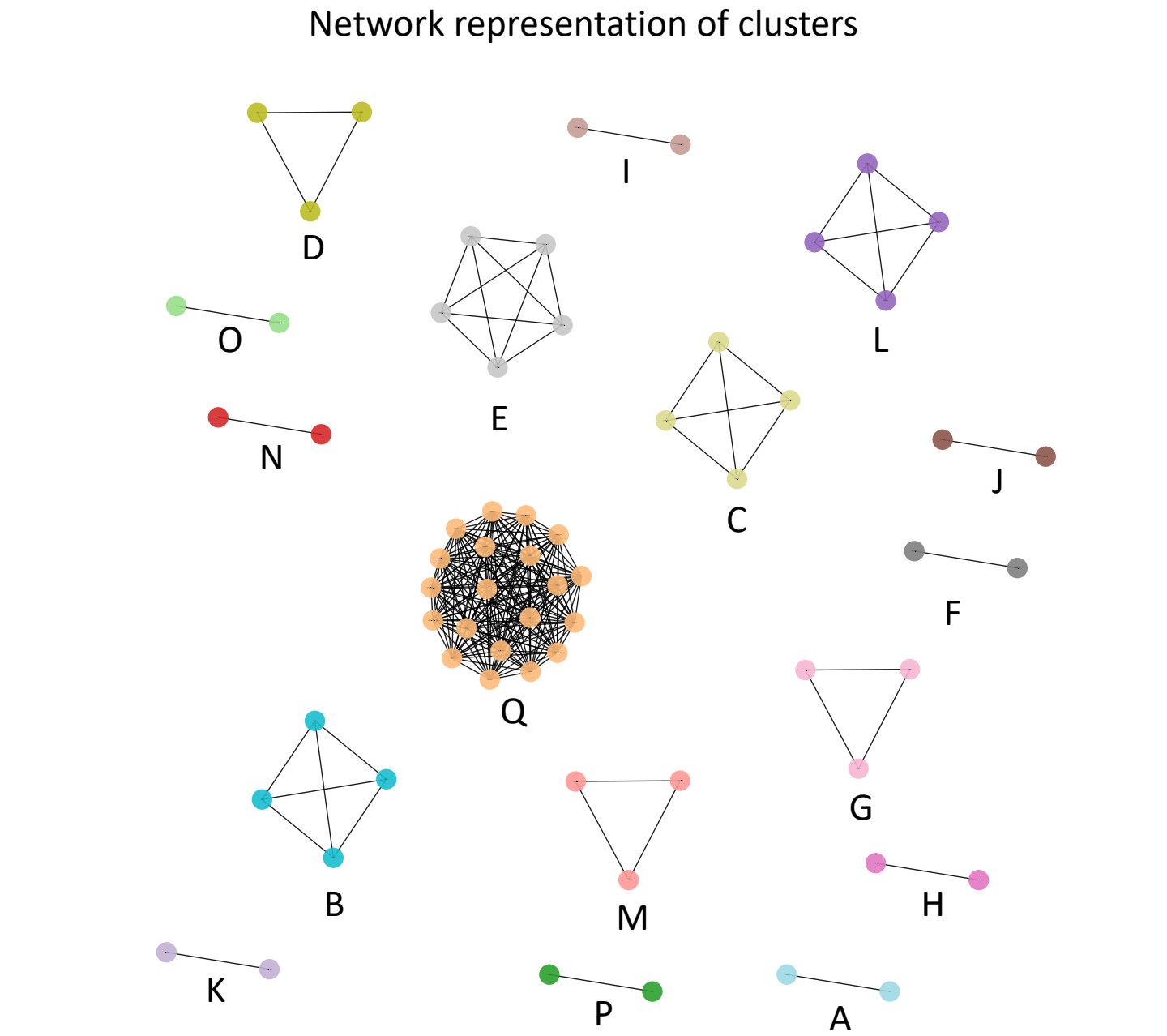
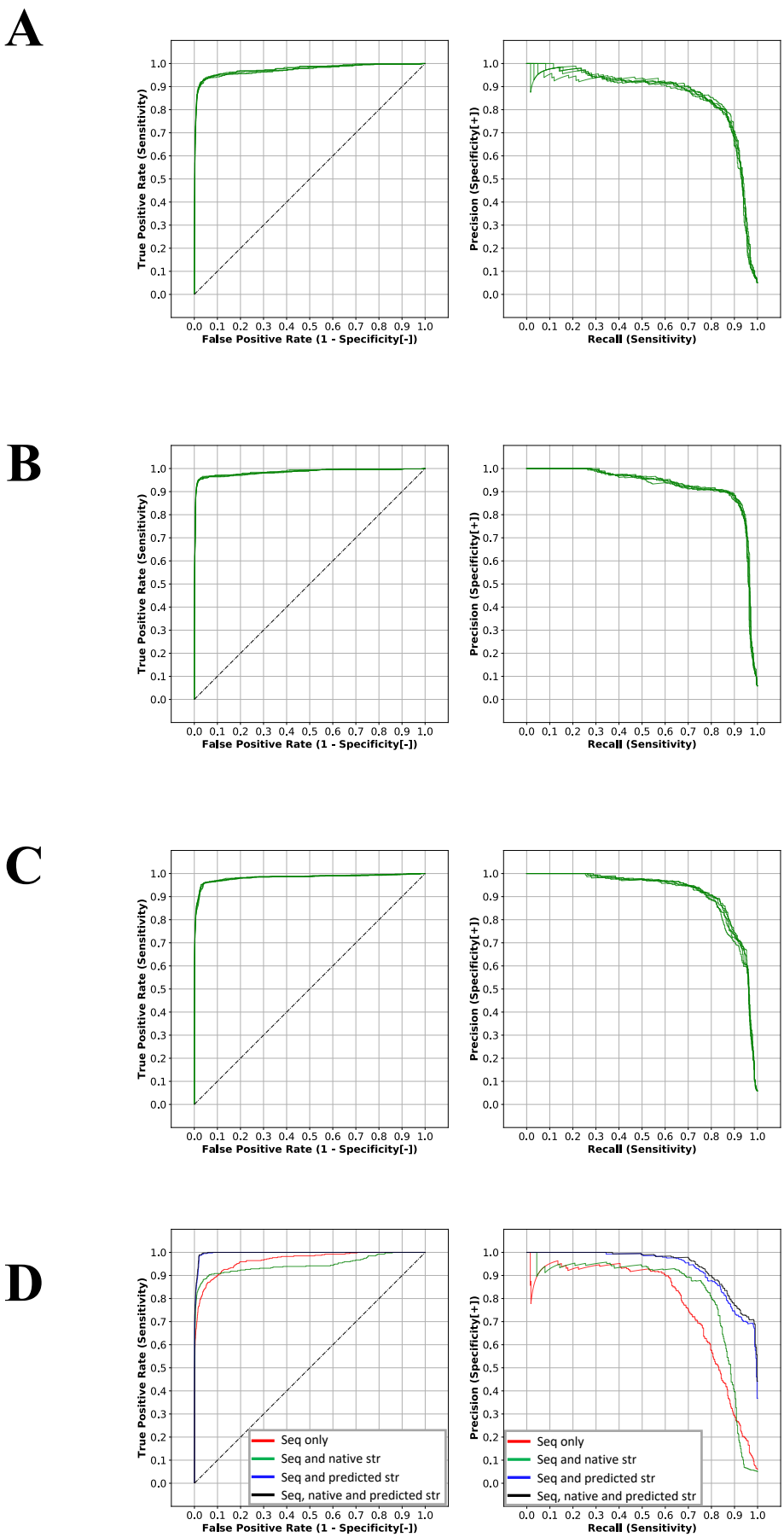
Clusters	Total number of PDBs	PDBs code
1	27	3se8HL,3se9HL,4j6rHL,4jb9HL,4jdtHL,4jpvHL,4jpwHL,4lsuHL,4olzHL,4rx4AD, 4s1qHL,4s1rHL,4s1sHL,4xmpHL,4xnyHL,4xnzHL,4xvtHL,4ydiHL,4ydlBC,4yflHL ,5cd5CD,5f96HL,5f9oHL,5f9wBC,5fecAB,5i9qBC,5te7HL
2	4	1deeFE,1xf5BA,1yntBA,4ioiBA
3	4	5u3jHL,5u3lHL,5u3nHL,5u3oHL
4	3	1ahwBA,1jpsHL,1uj3BA
5	3	1yy9DC,3b2uCD,5sx5HL
6	3	2b1hHL,3mlrHL,3mlwHL
7	3	3baeHL,3iflHL,3ifoAB
8	3	4tvpHL,5cezHL,5t3zHL
9	2	1a2yBA,1bvkBA
10	2	1eo8HL,1qfuHL
11	2	1frgHL,1ifhHL
12	2	1ggiHL,4jo1IM
13	2	1kc5HL,1kcsHL
14	2	1p2cBA,1yqvHL
15	2	1sy6HL,1xiwHG
16	2	1tziBA,4zffHL
17	2	2aepHL,2aeqHL
18	2	2eizBA,3d9aHL
19	2	2i9lHG,4u6hAB
20	2	2oslHL,3bkyHL
21	2	2xraHL,3ma9HL
22	2	3bn9DC,3so3CB
23	2	3effDC,3pjsBA
24	2	3eyfBA,4hhaBA
25	2	3gbnHL,4fqiHL
26	2	3l5wHL,3l5xHL
27	2	3u2sAB,4dqoHL
28	2	3ujiHL,3ujjHL
29	2	4al8HL,4bz1HL
30	2	4edwHL,4edxBA
31	2	4hkxAB,5ugyHL
32	2	4hlzGH,4py8IJ
33	2	4hs6HL,4hs8HL
34	2	4jpkHL,5iesHL
35	2	4jznIP,4z0xBA
36	2	4kvnHL,5k9oAB
37	2	4lqfHL,4lu5IM
38	2	4ojfHL,4onfHL
39	2	4r8wHL,5kaqQR
40	2	4ut6HL,5lcvHL
41	2	4utaHL,5h37KL
42	2	4y5vGH,4y5yAB
43	2	4ydvHL,5drzBA
44	2	4zs6CD,5do2HL
45	2	5d1qBA,5d1xBA
46	2	5esvHL,5eszAB
47	2	5f3bAB,5f3hAB
48	2	5fgbCA,5fgcEB
49	2	5jw3HL,5k9kHL
50	2	5kveHL,5kvfHL
51	2	5tljDC,5tlkHG
52	2	5vicHL,5vigAB

Suppl Table 1.

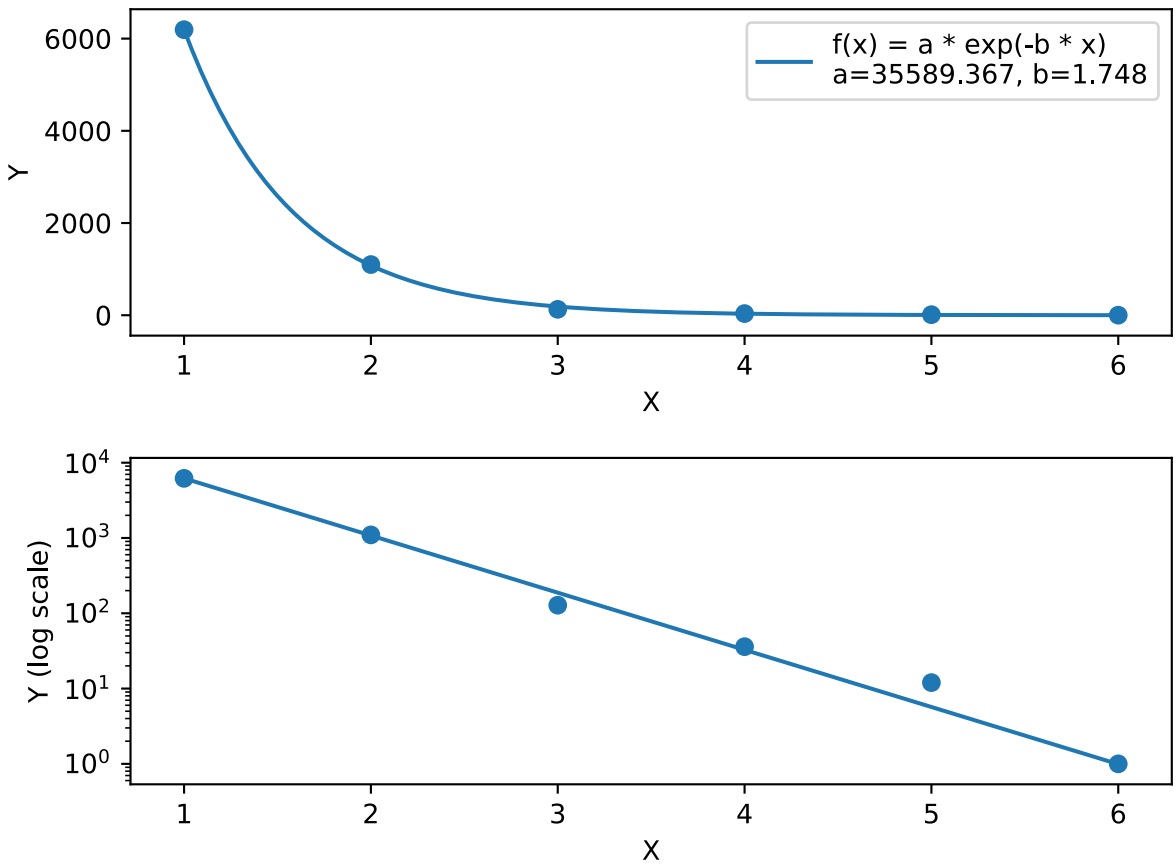




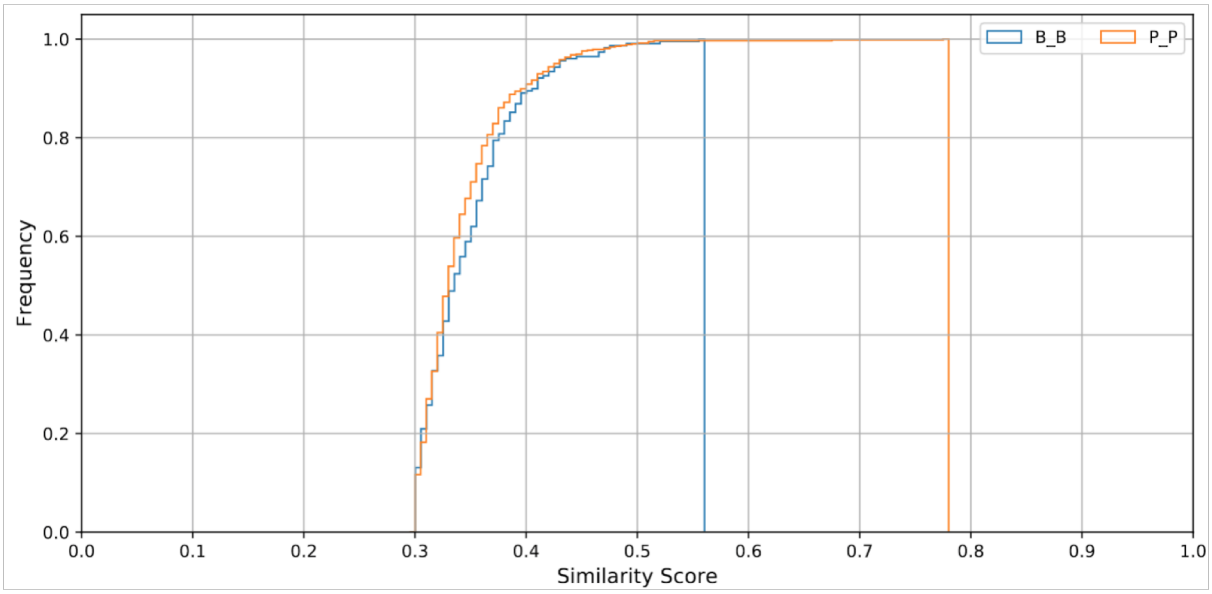
Suppl Fig. 1. Relationship between hierarchical clusters and networks. The Tree was sliced at a cutoff distance (0.796) and clusters were displayed as networks with edge-lengths proportional to their node distance.



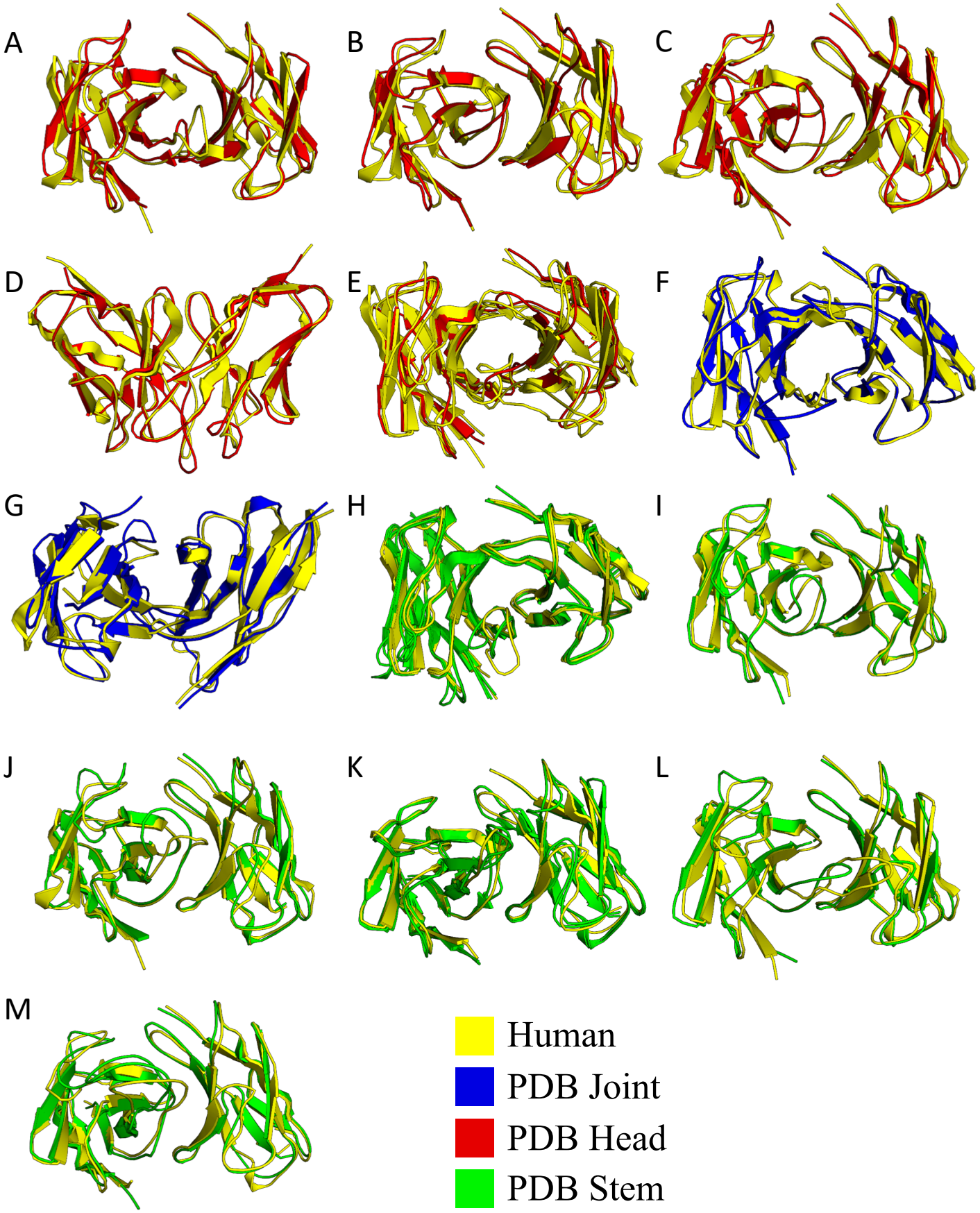
Suppl Fig. 2. Comparison of SVM models. The ROC and Precision-Recall curves for 5-fold cross-validation using sequence only (A), sequence and native structure (B) and sequence and predicted structures (C). The various SVMs, were also tested using Repertoire Builder models built with a template sequence identity blacklist of 80% (D).



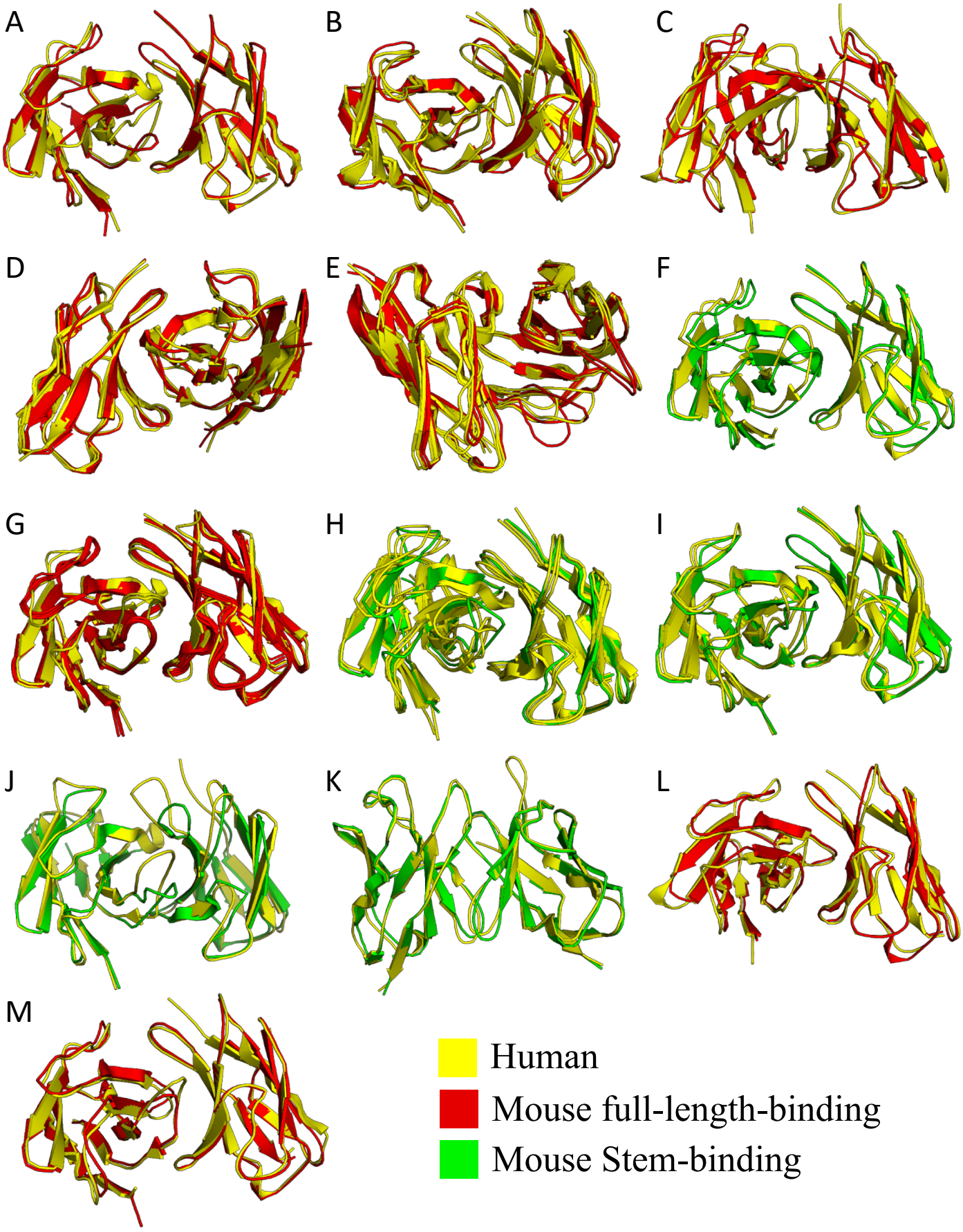
Suppl Fig. 3. Human BCR cluster size dependence. The best fit for distribution (Y) of cluster size (X) for the human BCR data was obtained by an exponential function.



Suppl Fig. 4. Cumulative distribution of human BCR similarity scores. The plot shows cumulative distributions for BCR similarity scores of pairs of plasmablasts (P_P) and pairs of B cells (B_B). Vertical lines indicate the maximum similarity observed.



Suppl Fig. 5. Visualization of structures of clusters containing human and known anti-HA PDB entries. The label of each illustration corresponds with the label of the cluster in Fig.5. The structures of PDB entries were characterized by their HA-binding modes as head (red), joint (blue) or stem (green). The structures of human BCRs were colored yellow.



Suppl Fig. 6. Visualization of structures of clusters composed of human and mouse BCRs obtained post flu vaccination. The label of each illustration is consistent with the cluster label in Fig.6. The structures of human, mouse stem-binding and mouse full-length-binding BCRs were showed as yellow, red and green, respectively.