Find A Gene Project

Sabrina Wu (A16731683)

[Q7] Generate a sequence identity based heatmap of your aligned sequences using R.

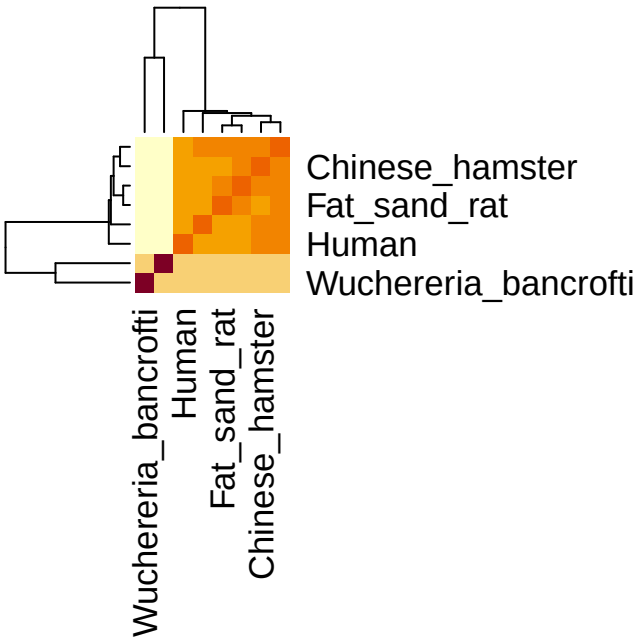
```
library(bio3d)
```

```
Warning: package 'bio3d' was built under R version 4.4.2
```

```
data <- read.fasta("bimm143projectseaview")
```

```
matrix <- seqidentity(data)
```

```
heatmap(matrix,margins = c(12,12))
```



[Q8] Using R/Bio3D (or an online blast server if you prefer), search the main protein structure database for the most similar atomic resolution structures to your aligned sequences.

```
consensus<- consensus(data)
```

```
blast <- blast.pdb(consensus$seq)
```

```
 Searching ... please wait (updates every 5 seconds) RID = MXJ00MNX013
 ..............................
 Reporting 125 hits
```

```
head(blast)
```

```
$hit.tbl
          queryid subjectids identity alignmentlength mismatches gapopens
1   Query_5967171     8VRW_B   68.846             260         37       25
2   Query_5967171     4AH2_B   74.619             197         15       21
3   Query_5967171     3PDO_B   74.619             197         15       21
4   Query_5967171     1AQD_B   74.619             197         15       21
5   Query_5967171     7YXB_B   73.232             198         18       21
6   Query_5967171     4X5X_B   74.112             197         16       21
7   Query_5967171     1ZGL_B   74.346             191         14       21
8   Query_5967171     7YX9_B   73.232             198         18       21
9   Query_5967171     2WBJ_B   72.222             198         20       21
10  Query_5967171     1YMM_B   72.589             197         19       21
11  Query_5967171     4I5B_B   74.869             191         13       21
12  Query_5967171     4FQX_B   74.479             192         14       21
13  Query_5967171     3L6F_B   74.869             191         13       21
14  Query_5967171     1FYT_B   74.869             191         13       21
15  Query_5967171     8CMB_B   73.958             192         15       21
16  Query_5967171     1FV1_B   74.074             189         14       21
17  Query_5967171   6QZC_BBB   75.132             189         12       21
18  Query_5967171     6CQJ_B   75.132             189         12       21
19  Query_5967171     1HXY_B   75.132             189         12       21
20  Query_5967171     6HBY_B   75.132             189         12       21
21  Query_5967171     8VSJ_B   75.132             189         12       21
22  Query_5967171     5V4N_C   75.132             189         12       21
23  Query_5967171     1D5X_B   73.822             191         15       21
24  Query_5967171     5JLZ_B   73.057             193         17       21
25  Query_5967171     4MCY_B   71.574             197         21       21
```

26	Query_5967171	6V0Y_B	71.574	197	21	21
27	Query_5967171	8EUQ_B	74.074	189	14	21
28	Query_5967171	8TRQ_B	73.822	191	15	21
29	Query_5967171	1D5M_B	73.298	191	16	21
30	Query_5967171	1DLH_B	75.401	187	11	21
31	Query_5967171	7NZE_BBB	73.684	190	15	21
32	Query_5967171	6BIJ_B	74.074	189	14	21
33	Query_5967171	8TRL_B	74.074	189	14	21
34	Query_5967171	3O6F_B	73.822	191	15	21
35	Query_5967171	6CPO_B	73.545	189	15	21
36	Query_5967171	6ATF_B	71.574	197	21	22
37	Query_5967171	6BIR_B	71.066	197	22	21
38	Query_5967171	4MD5_B	71.066	197	22	21
39	Query_5967171	7Z0Q_D	71.212	198	22	21
40	Query_5967171	4MDI_B	71.066	197	22	21
41	Query_5967171	6ATZ_B	74.332	187	13	22
42	Query_5967171	6CPL_B	73.016	189	16	21
43	Query_5967171	1BX2_B	73.016	189	16	21
44	Query_5967171	2FSE_B	72.727	187	16	20
45	Query_5967171	5V4M_C	73.016	189	16	21
46	Query_5967171	7N19_B	71.277	188	17	21
47	Query_5967171	3QIU_B	68.205	195	27	20
48	Query_5967171	4P2Q_B	67.000	200	31	20
49	Query_5967171	2Q6W_B	70.526	190	19	21
50	Query_5967171	1FNG_B	68.205	195	27	20
51	Query_5967171	1FNE_B	68.205	195	27	20
52	Query_5967171	3QIB_B	68.205	195	27	20
53	Query_5967171	1I3R_B	68.205	195	27	20
54	Query_5967171	1A6A_B	70.745	188	18	21
55	Query_5967171	3C5J_B	71.429	189	19	22
56	Query_5967171	4H25_B	72.193	187	17	22
57	Query_5967171	6BGA_B	67.005	197	30	20
58	Query_5967171	4P2O_B	67.005	197	30	20
59	Query_5967171	1IEB_B	68.229	192	26	20
60	Query_5967171	4H1L_B	71.658	187	18	22
61	Query_5967171	6PX6_B	59.227	233	53	22
62	Query_5967171	1KTD_B	67.027	185	26	20
63	Query_5967171	8VSP_B	56.223	233	60	21
64	Query_5967171	1R5V_B	66.848	184	26	20
65	Query_5967171	1KT2_B	67.033	182	25	20
66	Query_5967171	1S9V_B	64.130	184	33	19
67	Query_5967171	5KSU_B	63.784	185	34	19
68	Query_5967171	7ZAK_B	62.162	185	39	18

69	Query_5967171	7KEI_B	62.626	198	39	21
70	Query_5967171	6DIG_B	62.626	198	39	21
71	Query_5967171	1UVQ_B	64.324	185	33	20
72	Query_5967171	8JR4_B	60.847	189	39	18
73	Query_5967171	4D8P_B	64.286	182	32	19
74	Query_5967171	6XP6_B	62.842	183	35	19
75	Query_5967171	3WEX_B	59.474	190	46	18
76	Query_5967171	6U3M_B	62.842	183	35	19
77	Query_5967171	4OZF_B	62.842	183	35	19
78	Query_5967171	6MFF_C	62.842	183	35	19
79	Query_5967171	6MFG_E	62.500	184	36	19
80	Query_5967171	7QHP_B	59.091	198	47	20
81	Query_5967171	4P5M_B	61.878	181	38	18
82	Query_5967171	7SG1_B	64.045	178	31	19
83	Query_5967171	4P5K_B	61.878	181	38	18
84	Query_5967171	3LQZ_B	61.878	181	38	18
85	Query_5967171	7SG2_B	64.045	178	31	19
86	Query_5967171	7SG0_B	64.045	178	31	19
87	Query_5967171	4P4K_B	61.878	181	38	18
88	Query_5967171	3PL6_B	60.000	185	41	17
89	Query_5967171	4GRL_B	60.326	184	40	17
90	Query_5967171	4P57_B	61.326	181	39	18
91	Query_5967171	7T6I_B	61.582	177	37	18
92	Query_5967171	8W85_D	63.636	176	31	19
93	Query_5967171	7T2A_B	61.111	180	39	18
94	Query_5967171	7T2B_B	61.111	180	39	18
95	Query_5967171	8W86_D	63.636	176	31	19
96	Query_5967171	2P24_B	57.000	200	49	19
97	Query_5967171	8W83_D	63.636	176	31	19
98	Query_5967171	8W84_D	63.636	176	31	19
99	Query_5967171	4GG6_B	61.326	181	37	19
100	Query_5967171	4Z7U_B	59.783	184	41	19
101	Query_5967171	1JK8_B	61.236	178	36	19
102	Query_5967171	1F3J_B	61.582	177	36	19
103	Query_5967171	5KSA_B	59.783	184	41	19
104	Query_5967171	8VCX_B	61.236	178	36	19
105	Query_5967171	5KS9_B	59.783	184	41	19
106	Query_5967171	6XC9_C	59.783	184	41	19
107	Query_5967171	2NNA_B	61.236	178	36	19
108	Query_5967171	4P46_D	60.335	179	37	19
109	Query_5967171	6DFS_D	61.582	177	36	19
110	Query_5967171	6BLX_B	61.582	177	36	19
111	Query_5967171	4P23_D	59.218	179	39	18

112	Query_5967171	5UJT_B	61.017	177	36	19	
113	Query_5967171	6BLQ_B	61.582	177	36	19	
114	Query_5967171	8VD0_C	61.236	178	36	19	
115	Query_5967171	2PXY_D	59.218	179	41	18	
116	Query_5967171	6DFX_B	61.017	177	36	19	
117	Query_5967171	2IAD_B	60.674	178	36	20	
118	Query_5967171	7RDV_B	61.364	176	34	20	
119	Query_5967171	1IAO_B	60.112	178	37	20	
120	Query_5967171	4I0P_D	56.354	181	46	19	
121	Query_5967171	7PDY_B	47.568	185	64	16	
122	Query_5967171	7APZ_B	46.486	185	66	16	
123	Query_5967171	6T3Y_B	48.571	175	57	16	
124	Query_5967171	6KVM_B	47.312	186	65	17	
125	Query_5967171	6ZWA_B	49.714	175	55	17	

	q.start	q.end	s.start	s.end	evalue	bitscore	positives	mlog.evalue	pdb.id
1	2	217	4	263	1.21e-77	236.0	73.46	177.10843	8VRW_B
2	22	183	33	229	3.31e-66	205.0	78.17	150.77367	4AH2_B
3	22	183	3	199	4.52e-66	203.0	78.17	150.46210	3PDO_B
4	22	183	2	198	5.38e-66	203.0	78.17	150.28793	1AQD_B
5	22	184	1	198	2.04e-65	201.0	77.27	148.95508	7YXB_B
6	22	183	33	229	4.93e-65	202.0	77.66	148.07269	4X5X_B
7	22	177	2	192	4.29e-64	198.0	77.49	145.90916	1ZGL_B
8	22	184	20	217	5.87e-64	199.0	77.27	145.59559	7YX9_B
9	22	184	2	199	1.21e-63	197.0	77.27	144.87224	2WBJ_B
10	22	183	2	198	1.65e-63	197.0	77.16	144.56209	1YMM_B
11	22	177	1	191	3.33e-63	196.0	78.01	143.85989	4I5B_B
12	22	178	8	199	3.99e-63	196.0	78.12	143.67907	4FQX_B
13	22	177	3	193	4.72e-63	195.0	78.01	143.51105	3L6F_B
14	22	177	2	192	5.62e-63	195.0	78.01	143.33653	1FYT_B
15	19	175	3	194	1.50e-62	194.0	77.60	142.35481	8CMB_B
16	22	175	2	190	1.77e-62	194.0	77.25	142.18930	1FV1_B
17	22	175	3	191	1.99e-62	194.0	77.78	142.07214	6QZC_b
18	22	175	1	189	2.09e-62	194.0	77.78	142.02311	6CQJ_B
19	22	175	2	190	2.32e-62	194.0	77.78	141.91871	1HXY_B
20	22	175	3	191	2.35e-62	194.0	77.78	141.90586	6HBY_B
21	22	175	2	190	2.64e-62	193.0	77.78	141.78950	8VSJ_B
22	22	175	28	216	3.98e-62	194.0	77.78	141.37899	5V4N_C
23	22	177	2	192	4.23e-62	193.0	76.96	141.31807	1D5X_B
24	22	179	2	194	5.96e-62	193.0	75.65	140.97521	5JLZ_B
25	22	183	4	200	8.90e-62	192.0	74.62	140.57422	4MCY_B
26	22	183	2	198	9.01e-62	192.0	74.62	140.56194	6V0Y_B
27	22	175	29	217	1.19e-61	193.0	76.72	140.28374	8EUQ_B
28	22	177	2	192	1.22e-61	192.0	76.44	140.25884	8TRQ_B

29	22	177	2	192	1.34e-61	192.0	76.96	140.16502	1D5M_B
30	24	175	2	188	1.57e-61	191.0	78.07	140.00662	1DLH_B
31	22	176	2	191	1.73e-61	191.0	76.84	139.90957	7NZE_b
32	22	175	1	189	1.79e-61	191.0	76.72	139.87548	6BIJ_B
33	22	175	2	190	2.03e-61	191.0	76.72	139.74965	8TRL_B
34	22	177	31	221	2.46e-61	192.0	76.96	139.55753	3O6F_B
35	22	175	2	190	4.22e-61	190.0	77.25	139.01786	6CPO_B
36	22	183	4	200	4.27e-61	191.0	74.62	139.00608	6ATF_B
37	22	183	4	200	1.18e-60	189.0	74.11	137.98959	6BIR_B
38	22	183	4	200	2.03e-60	189.0	74.11	137.44707	4MD5_B
39	22	184	1	198	2.39e-60	189.0	75.25	137.28381	7Z0Q_D
40	22	183	4	200	2.84e-60	189.0	73.60	137.11130	4MDI_B
41	24	175	2	188	4.13e-60	188.0	77.01	136.73683	6ATZ_B
42	22	175	2	190	4.85e-60	187.0	75.13	136.57613	6CPL_B
43	24	177	2	190	7.92e-60	187.0	77.25	136.08571	1BX2_B
44	24	175	1	187	8.87e-60	187.0	76.47	135.97243	2FSE_B
45	22	175	28	216	1.20e-59	187.0	76.72	135.67020	5V4M_C
46	24	175	1	187	6.88e-57	179.0	74.47	129.31873	7N19_B
47	24	183	2	196	8.02e-57	180.0	73.85	129.16541	3QIU_B
48	24	188	8	207	9.91e-57	180.0	73.00	128.95381	4P2Q_B
49	22	175	2	190	1.22e-56	179.0	74.21	128.74591	2Q6W_B
50	24	183	30	224	1.32e-56	180.0	73.85	128.66713	1FNG_B
51	24	183	30	224	1.37e-56	180.0	73.85	128.62995	1FNE_B
52	24	183	8	202	1.82e-56	179.0	73.85	128.34593	3QIB_B
53	24	183	34	228	1.96e-56	180.0	73.85	128.27182	1I3R_B
54	25	176	1	187	2.26e-56	178.0	74.47	128.12940	1A6A_B
55	22	175	2	190	6.33e-56	177.0	74.60	127.09946	3C5J_B
56	24	175	2	188	7.31e-56	177.0	75.40	126.95552	4H25_B
57	27	188	32	228	8.77e-56	178.0	73.10	126.77343	6BGA_B
58	27	188	35	231	1.24e-55	178.0	73.10	126.42707	4P2O_B
59	27	183	36	227	2.97e-55	177.0	73.96	125.55362	1IEB_B
60	24	175	1	187	5.86e-55	174.0	74.87	124.87403	4H1L_B
61	2	192	7	239	3.91e-53	172.0	65.24	120.67347	6PX6_B
62	24	173	31	215	5.79e-50	163.0	72.43	113.37312	1KTD_B
63	2	192	7	239	8.81e-50	165.0	63.09	112.95337	8VSP_B
64	25	173	2	185	3.71e-49	160.0	72.28	111.51564	1R5V_B
65	27	173	32	213	2.71e-48	159.0	72.53	109.52714	1KT2_B
66	33	183	15	198	4.24e-46	152.0	68.48	104.47435	1S9V_B
67	33	184	15	199	4.44e-46	152.0	68.65	104.42826	5KSU_B
68	30	183	14	198	8.42e-46	154.0	69.73	103.78830	7ZAK_B
69	33	197	14	209	8.90e-46	152.0	67.17	103.73286	7KEI_B
70	33	197	19	214	1.54e-45	151.0	67.17	103.18455	6DIG_B
71	33	184	13	197	1.83e-45	150.0	69.19	103.01201	1UVQ_B

72	22	175	2	190	6.53e-45	149.0	66.67	101.73992	8JR4_B
73	33	181	48	229	1.70e-44	150.0	68.13	100.78312	4D8P_B
74	33	182	16	198	7.87e-44	147.0	67.21	99.25069	6XP6_B
75	27	185	29	218	7.94e-44	147.0	67.89	99.24183	3WEX_B
76	33	182	21	203	1.36e-43	146.0	67.21	98.70367	6U3M_B
77	33	182	28	210	1.55e-43	146.0	67.21	98.57290	4OZF_B
78	33	182	41	223	2.42e-43	146.0	67.21	98.12739	6MFF_C
79	33	183	41	224	4.72e-43	145.0	66.85	97.45935	6MFG_E
80	22	185	28	225	4.94e-43	145.0	65.15	97.41379	7QHP_B
81	27	176	32	212	5.23e-43	145.0	67.96	97.35675	4P5M_B
82	33	177	25	202	6.54e-43	144.0	67.98	97.13322	7SG1_B
83	27	176	32	212	7.24e-43	144.0	67.96	97.03154	4P5K_B
84	27	176	32	212	8.89e-43	144.0	67.96	96.82623	3LQZ_B
85	33	177	28	205	9.24e-43	144.0	67.98	96.78762	7SG2_B
86	33	177	30	207	1.10e-42	144.0	67.98	96.61326	7SG0_B
87	27	176	32	212	1.20e-42	144.0	67.96	96.52625	4P4K_B
88	33	184	17	201	1.62e-42	143.0	64.86	96.22615	3PL6_B
89	33	183	17	200	2.21e-42	143.0	64.67	95.91558	4GRL_B
90	27	176	32	212	4.04e-42	142.0	67.96	95.31233	4P57_B
91	30	175	12	188	6.82e-42	141.0	68.93	94.78871	7T6I_B
92	33	175	41	216	9.78e-42	142.0	67.61	94.42823	8W85_D
93	27	175	9	188	1.04e-41	141.0	67.78	94.36677	7T2A_B
94	27	175	11	190	1.08e-41	141.0	67.78	94.32903	7T2B_B
95	33	175	49	224	1.70e-41	142.0	67.61	93.87536	8W86_D
96	30	197	64	258	1.91e-41	142.0	64.50	93.75889	2P24_B
97	33	175	45	220	2.09e-41	141.0	67.61	93.66882	8W83_D
98	33	175	45	220	2.37e-41	141.0	67.61	93.54310	8W84_D
99	33	180	30	210	2.72e-41	140.0	66.85	93.40536	4GG6_B
100	33	183	28	211	4.60e-40	137.0	65.22	90.57735	4Z7U_B
101	33	177	13	190	5.29e-40	137.0	66.29	90.43759	1JK8_B
102	32	176	11	187	5.35e-40	136.0	66.67	90.42631	1F3J_B
103	33	183	40	223	6.19e-40	137.0	65.22	90.28047	5KSA_B
104	33	177	15	192	7.11e-40	136.0	66.29	90.14190	8VCX_B
105	33	183	45	228	8.42e-40	137.0	65.22	89.97279	5KS9_B
106	33	183	45	228	8.88e-40	137.0	65.22	89.91960	6XC9_C
107	33	177	30	207	1.09e-39	136.0	66.29	89.71464	2NNA_B
108	32	176	40	218	1.20e-39	136.0	67.04	89.61850	4P46_D
109	32	176	39	215	1.46e-39	136.0	66.67	89.42238	6DFS_D
110	32	176	39	215	1.96e-39	136.0	66.67	89.12787	6BLX_B
111	32	176	40	218	2.03e-39	136.0	65.92	89.09278	4P23_D
112	33	176	13	189	2.75e-39	135.0	66.10	88.78922	5UJT_B
113	32	176	39	215	2.76e-39	135.0	66.67	88.78559	6BLQ_B
114	33	177	36	213	3.57e-39	135.0	66.29	88.52825	8VD0_C

115	30	176	11	189	6.59e-39	134.0	65.92	87.91527	2PXY_D
116	33	176	39	215	7.37e-39	134.0	66.10	87.80340	6DFX_B
117	32	175	28	205	2.45e-38	133.0	66.85	86.60215	2IAD_B
118	32	173	11	186	3.96e-38	132.0	67.05	86.12199	7RDV_B
119	32	175	39	216	9.20e-38	132.0	66.85	85.27903	1IAO_B
120	30	177	9	189	1.20e-35	125.0	66.85	80.40816	4I0P_D
121	32	183	41	225	4.37e-27	104.0	58.38	60.69503	7PDY_B
122	32	183	41	225	2.95e-26	102.0	57.84	58.78541	7APZ_B
123	32	173	49	223	2.99e-25	99.8	58.29	56.46935	6T3Y_B
124	31	183	29	214	5.57e-25	98.6	58.06	55.84723	6KVM_B
125	32	173	44	218	8.84e-25	98.2	58.86	55.38534	6ZWA_B

	acc
1	8VRW_B
2	4AH2_B
3	3PDO_B
4	1AQD_B
5	7YXB_B
6	4X5X_B
7	1ZGL_B
8	7YX9_B
9	2WBJ_B
10	1YMM_B
11	4I5B_B
12	4FQX_B
13	3L6F_B
14	1FYT_B
15	8CMB_B
16	1FV1_B
17	6QZC_BBB
18	6CQJ_B
19	1HXY_B
20	6HBY_B
21	8VSJ_B
22	5V4N_C
23	1D5X_B
24	5JLZ_B
25	4MCY_B
26	6V0Y_B
27	8EUQ_B
28	8TRQ_B
29	1D5M_B
30	1DLH_B
31	7NZE_BBB

32	6BIJ_B
33	8TRL_B
34	3O6F_B
35	6CPO_B
36	6ATF_B
37	6BIR_B
38	4MD5_B
39	7Z0Q_D
40	4MDI_B
41	6ATZ_B
42	6CPL_B
43	1BX2_B
44	2FSE_B
45	5V4M_C
46	7N19_B
47	3QIU_B
48	4P2Q_B
49	2Q6W_B
50	1FNG_B
51	1FNE_B
52	3QIB_B
53	1I3R_B
54	1A6A_B
55	3C5J_B
56	4H25_B
57	6BGA_B
58	4P2O_B
59	1IEB_B
60	4H1L_B
61	6PX6_B
62	1KTD_B
63	8VSP_B
64	1R5V_B
65	1KT2_B
66	1S9V_B
67	5KSU_B
68	7ZAK_B
69	7KEI_B
70	6DIG_B
71	1UVQ_B
72	8JR4_B
73	4D8P_B
74	6XP6_B

75	3WEX_B
76	6U3M_B
77	4OZF_B
78	6MFF_C
79	6MFG_E
80	7QHP_B
81	4P5M_B
82	7SG1_B
83	4P5K_B
84	3LQZ_B
85	7SG2_B
86	7SG0_B
87	4P4K_B
88	3PL6_B
89	4GRL_B
90	4P57_B
91	7T6I_B
92	8W85_D
93	7T2A_B
94	7T2B_B
95	8W86_D
96	2P24_B
97	8W83_D
98	8W84_D
99	4GG6_B
100	4Z7U_B
101	1JK8_B
102	1F3J_B
103	5KSA_B
104	8VCX_B
105	5KS9_B
106	6XC9_C
107	2NNA_B
108	4P46_D
109	6DFS_D
110	6BLX_B
111	4P23_D
112	5UJT_B
113	6BLQ_B
114	8VD0_C
115	2PXY_D
116	6DFX_B
117	2IAD_B

```
118   7RDV_B
119   1IAO_B
120   4IOP_D
121   7PDY_B
122   7APZ_B
123   6T3Y_B
124   6KVM_B
125   6ZWA_B

$raw
          queryid subjectids identity alignmentlength mismatches gapopens
1   Query_5967171     8VRW_B   68.846             260         37       25
2   Query_5967171     4AH2_B   74.619             197         15       21
3   Query_5967171     3PDO_B   74.619             197         15       21
4   Query_5967171     1AQD_B   74.619             197         15       21
5   Query_5967171     7YXB_B   73.232             198         18       21
6   Query_5967171     4X5X_B   74.112             197         16       21
7   Query_5967171     1ZGL_B   74.346             191         14       21
8   Query_5967171     7YX9_B   73.232             198         18       21
9   Query_5967171     2WBJ_B   72.222             198         20       21
10  Query_5967171     1YMM_B   72.589             197         19       21
11  Query_5967171     4I5B_B   74.869             191         13       21
12  Query_5967171     4FQX_B   74.479             192         14       21
13  Query_5967171     3L6F_B   74.869             191         13       21
14  Query_5967171     1FYT_B   74.869             191         13       21
15  Query_5967171     8CMB_B   73.958             192         15       21
16  Query_5967171     1FV1_B   74.074             189         14       21
17  Query_5967171   6QZC_BBB   75.132             189         12       21
18  Query_5967171     6CQJ_B   75.132             189         12       21
19  Query_5967171     1HXY_B   75.132             189         12       21
20  Query_5967171     6HBY_B   75.132             189         12       21
21  Query_5967171     8VSJ_B   75.132             189         12       21
22  Query_5967171     5V4N_C   75.132             189         12       21
23  Query_5967171     1D5X_B   73.822             191         15       21
24  Query_5967171     5JLZ_B   73.057             193         17       21
25  Query_5967171     4MCY_B   71.574             197         21       21
26  Query_5967171     6V0Y_B   71.574             197         21       21
27  Query_5967171     8EUQ_B   74.074             189         14       21
28  Query_5967171     8TRQ_B   73.822             191         15       21
29  Query_5967171     1D5M_B   73.298             191         16       21
30  Query_5967171     1DLH_B   75.401             187         11       21
31  Query_5967171   7NZE_BBB   73.684             190         15       21
32  Query_5967171     6BIJ_B   74.074             189         14       21
```

33	Query_5967171	8TRL_B	74.074	189	14	21
34	Query_5967171	3O6F_B	73.822	191	15	21
35	Query_5967171	6CPO_B	73.545	189	15	21
36	Query_5967171	6ATF_B	71.574	197	21	22
37	Query_5967171	6BIR_B	71.066	197	22	21
38	Query_5967171	4MD5_B	71.066	197	22	21
39	Query_5967171	7Z0Q_D	71.212	198	22	21
40	Query_5967171	4MDI_B	71.066	197	22	21
41	Query_5967171	6ATZ_B	74.332	187	13	22
42	Query_5967171	6CPL_B	73.016	189	16	21
43	Query_5967171	1BX2_B	73.016	189	16	21
44	Query_5967171	2FSE_B	72.727	187	16	20
45	Query_5967171	5V4M_C	73.016	189	16	21
46	Query_5967171	7N19_B	71.277	188	17	21
47	Query_5967171	3QIU_B	68.205	195	27	20
48	Query_5967171	4P2Q_B	67.000	200	31	20
49	Query_5967171	2Q6W_B	70.526	190	19	21
50	Query_5967171	1FNG_B	68.205	195	27	20
51	Query_5967171	1FNE_B	68.205	195	27	20
52	Query_5967171	3QIB_B	68.205	195	27	20
53	Query_5967171	1I3R_B	68.205	195	27	20
54	Query_5967171	1A6A_B	70.745	188	18	21
55	Query_5967171	3C5J_B	71.429	189	19	22
56	Query_5967171	4H25_B	72.193	187	17	22
57	Query_5967171	6BGA_B	67.005	197	30	20
58	Query_5967171	4P2O_B	67.005	197	30	20
59	Query_5967171	1IEB_B	68.229	192	26	20
60	Query_5967171	4H1L_B	71.658	187	18	22
61	Query_5967171	6PX6_B	59.227	233	53	22
62	Query_5967171	1KTD_B	67.027	185	26	20
63	Query_5967171	8VSP_B	56.223	233	60	21
64	Query_5967171	1R5V_B	66.848	184	26	20
65	Query_5967171	1KT2_B	67.033	182	25	20
66	Query_5967171	1S9V_B	64.130	184	33	19
67	Query_5967171	5KSU_B	63.784	185	34	19
68	Query_5967171	7ZAK_B	62.162	185	39	18
69	Query_5967171	7KEI_B	62.626	198	39	21
70	Query_5967171	6DIG_B	62.626	198	39	21
71	Query_5967171	1UVQ_B	64.324	185	33	20
72	Query_5967171	8JR4_B	60.847	189	39	18
73	Query_5967171	4D8P_B	64.286	182	32	19
74	Query_5967171	6XP6_B	62.842	183	35	19
75	Query_5967171	3WEX_B	59.474	190	46	18

76	Query_5967171	6U3M_B	62.842	183	35	19
77	Query_5967171	4OZF_B	62.842	183	35	19
78	Query_5967171	6MFF_C	62.842	183	35	19
79	Query_5967171	6MFG_E	62.500	184	36	19
80	Query_5967171	7QHP_B	59.091	198	47	20
81	Query_5967171	4P5M_B	61.878	181	38	18
82	Query_5967171	7SG1_B	64.045	178	31	19
83	Query_5967171	4P5K_B	61.878	181	38	18
84	Query_5967171	3LQZ_B	61.878	181	38	18
85	Query_5967171	7SG2_B	64.045	178	31	19
86	Query_5967171	7SG0_B	64.045	178	31	19
87	Query_5967171	4P4K_B	61.878	181	38	18
88	Query_5967171	3PL6_B	60.000	185	41	17
89	Query_5967171	4GRL_B	60.326	184	40	17
90	Query_5967171	4P57_B	61.326	181	39	18
91	Query_5967171	7T6I_B	61.582	177	37	18
92	Query_5967171	8W85_D	63.636	176	31	19
93	Query_5967171	7T2A_B	61.111	180	39	18
94	Query_5967171	7T2B_B	61.111	180	39	18
95	Query_5967171	8W86_D	63.636	176	31	19
96	Query_5967171	2P24_B	57.000	200	49	19
97	Query_5967171	8W83_D	63.636	176	31	19
98	Query_5967171	8W84_D	63.636	176	31	19
99	Query_5967171	4GG6_B	61.326	181	37	19
100	Query_5967171	4Z7U_B	59.783	184	41	19
101	Query_5967171	1JK8_B	61.236	178	36	19
102	Query_5967171	1F3J_B	61.582	177	36	19
103	Query_5967171	5KSA_B	59.783	184	41	19
104	Query_5967171	8VCX_B	61.236	178	36	19
105	Query_5967171	5KS9_B	59.783	184	41	19
106	Query_5967171	6XC9_C	59.783	184	41	19
107	Query_5967171	2NNA_B	61.236	178	36	19
108	Query_5967171	4P46_D	60.335	179	37	19
109	Query_5967171	6DFS_D	61.582	177	36	19
110	Query_5967171	6BLX_B	61.582	177	36	19
111	Query_5967171	4P23_D	59.218	179	39	18
112	Query_5967171	5UJT_B	61.017	177	36	19
113	Query_5967171	6BLQ_B	61.582	177	36	19
114	Query_5967171	8VD0_C	61.236	178	36	19
115	Query_5967171	2PXY_D	59.218	179	41	18
116	Query_5967171	6DFX_B	61.017	177	36	19
117	Query_5967171	2IAD_B	60.674	178	36	20
118	Query_5967171	7RDV_B	61.364	176	34	20

```
119 Query_5967171     1IAO_B   60.112             178         37       20
120 Query_5967171     4I0P_D   56.354             181         46       19
121 Query_5967171     7PDY_B   47.568             185         64       16
122 Query_5967171     7APZ_B   46.486             185         66       16
123 Query_5967171     6T3Y_B   48.571             175         57       16
124 Query_5967171     6KVM_B   47.312             186         65       17
125 Query_5967171     6ZWA_B   49.714             175         55       17
    q.start q.end s.start s.end   evalue bitscore positives
1         2   217       4   263 1.21e-77    236.0     73.46
2        22   183      33   229 3.31e-66    205.0     78.17
3        22   183       3   199 4.52e-66    203.0     78.17
4        22   183       2   198 5.38e-66    203.0     78.17
5        22   184       1   198 2.04e-65    201.0     77.27
6        22   183      33   229 4.93e-65    202.0     77.66
7        22   177       2   192 4.29e-64    198.0     77.49
8        22   184      20   217 5.87e-64    199.0     77.27
9        22   184       2   199 1.21e-63    197.0     77.27
10       22   183       2   198 1.65e-63    197.0     77.16
11       22   177       1   191 3.33e-63    196.0     78.01
12       22   178       8   199 3.99e-63    196.0     78.12
13       22   177       3   193 4.72e-63    195.0     78.01
14       22   177       2   192 5.62e-63    195.0     78.01
15       19   175       3   194 1.50e-62    194.0     77.60
16       22   175       2   190 1.77e-62    194.0     77.25
17       22   175       3   191 1.99e-62    194.0     77.78
18       22   175       1   189 2.09e-62    194.0     77.78
19       22   175       2   190 2.32e-62    194.0     77.78
20       22   175       3   191 2.35e-62    194.0     77.78
21       22   175       2   190 2.64e-62    193.0     77.78
22       22   175      28   216 3.98e-62    194.0     77.78
23       22   177       2   192 4.23e-62    193.0     76.96
24       22   179       2   194 5.96e-62    193.0     75.65
25       22   183       4   200 8.90e-62    192.0     74.62
26       22   183       2   198 9.01e-62    192.0     74.62
27       22   175      29   217 1.19e-61    193.0     76.72
28       22   177       2   192 1.22e-61    192.0     76.44
29       22   177       2   192 1.34e-61    192.0     76.96
30       24   175       2   188 1.57e-61    191.0     78.07
31       22   176       2   191 1.73e-61    191.0     76.84
32       22   175       1   189 1.79e-61    191.0     76.72
33       22   175       2   190 2.03e-61    191.0     76.72
34       22   177      31   221 2.46e-61    192.0     76.96
35       22   175       2   190 4.22e-61    190.0     77.25
```

36	22	183	4	200	4.27e-61	191.0	74.62
37	22	183	4	200	1.18e-60	189.0	74.11
38	22	183	4	200	2.03e-60	189.0	74.11
39	22	184	1	198	2.39e-60	189.0	75.25
40	22	183	4	200	2.84e-60	189.0	73.60
41	24	175	2	188	4.13e-60	188.0	77.01
42	22	175	2	190	4.85e-60	187.0	75.13
43	24	177	2	190	7.92e-60	187.0	77.25
44	24	175	1	187	8.87e-60	187.0	76.47
45	22	175	28	216	1.20e-59	187.0	76.72
46	24	175	1	187	6.88e-57	179.0	74.47
47	24	183	2	196	8.02e-57	180.0	73.85
48	24	188	8	207	9.91e-57	180.0	73.00
49	22	175	2	190	1.22e-56	179.0	74.21
50	24	183	30	224	1.32e-56	180.0	73.85
51	24	183	30	224	1.37e-56	180.0	73.85
52	24	183	8	202	1.82e-56	179.0	73.85
53	24	183	34	228	1.96e-56	180.0	73.85
54	25	176	1	187	2.26e-56	178.0	74.47
55	22	175	2	190	6.33e-56	177.0	74.60
56	24	175	2	188	7.31e-56	177.0	75.40
57	27	188	32	228	8.77e-56	178.0	73.10
58	27	188	35	231	1.24e-55	178.0	73.10
59	27	183	36	227	2.97e-55	177.0	73.96
60	24	175	1	187	5.86e-55	174.0	74.87
61	2	192	7	239	3.91e-53	172.0	65.24
62	24	173	31	215	5.79e-50	163.0	72.43
63	2	192	7	239	8.81e-50	165.0	63.09
64	25	173	2	185	3.71e-49	160.0	72.28
65	27	173	32	213	2.71e-48	159.0	72.53
66	33	183	15	198	4.24e-46	152.0	68.48
67	33	184	15	199	4.44e-46	152.0	68.65
68	30	183	14	198	8.42e-46	154.0	69.73
69	33	197	14	209	8.90e-46	152.0	67.17
70	33	197	19	214	1.54e-45	151.0	67.17
71	33	184	13	197	1.83e-45	150.0	69.19
72	22	175	2	190	6.53e-45	149.0	66.67
73	33	181	48	229	1.70e-44	150.0	68.13
74	33	182	16	198	7.87e-44	147.0	67.21
75	27	185	29	218	7.94e-44	147.0	67.89
76	33	182	21	203	1.36e-43	146.0	67.21
77	33	182	28	210	1.55e-43	146.0	67.21
78	33	182	41	223	2.42e-43	146.0	67.21

79	33	183	41	224	4.72e-43	145.0	66.85
80	22	185	28	225	4.94e-43	145.0	65.15
81	27	176	32	212	5.23e-43	145.0	67.96
82	33	177	25	202	6.54e-43	144.0	67.98
83	27	176	32	212	7.24e-43	144.0	67.96
84	27	176	32	212	8.89e-43	144.0	67.96
85	33	177	28	205	9.24e-43	144.0	67.98
86	33	177	30	207	1.10e-42	144.0	67.98
87	27	176	32	212	1.20e-42	144.0	67.96
88	33	184	17	201	1.62e-42	143.0	64.86
89	33	183	17	200	2.21e-42	143.0	64.67
90	27	176	32	212	4.04e-42	142.0	67.96
91	30	175	12	188	6.82e-42	141.0	68.93
92	33	175	41	216	9.78e-42	142.0	67.61
93	27	175	9	188	1.04e-41	141.0	67.78
94	27	175	11	190	1.08e-41	141.0	67.78
95	33	175	49	224	1.70e-41	142.0	67.61
96	30	197	64	258	1.91e-41	142.0	64.50
97	33	175	45	220	2.09e-41	141.0	67.61
98	33	175	45	220	2.37e-41	141.0	67.61
99	33	180	30	210	2.72e-41	140.0	66.85
100	33	183	28	211	4.60e-40	137.0	65.22
101	33	177	13	190	5.29e-40	137.0	66.29
102	32	176	11	187	5.35e-40	136.0	66.67
103	33	183	40	223	6.19e-40	137.0	65.22
104	33	177	15	192	7.11e-40	136.0	66.29
105	33	183	45	228	8.42e-40	137.0	65.22
106	33	183	45	228	8.88e-40	137.0	65.22
107	33	177	30	207	1.09e-39	136.0	66.29
108	32	176	40	218	1.20e-39	136.0	67.04
109	32	176	39	215	1.46e-39	136.0	66.67
110	32	176	39	215	1.96e-39	136.0	66.67
111	32	176	40	218	2.03e-39	136.0	65.92
112	33	176	13	189	2.75e-39	135.0	66.10
113	32	176	39	215	2.76e-39	135.0	66.67
114	33	177	36	213	3.57e-39	135.0	66.29
115	30	176	11	189	6.59e-39	134.0	65.92
116	33	176	39	215	7.37e-39	134.0	66.10
117	32	175	28	205	2.45e-38	133.0	66.85
118	32	173	11	186	3.96e-38	132.0	67.05
119	32	175	39	216	9.20e-38	132.0	66.85
120	30	177	9	189	1.20e-35	125.0	66.85
121	32	183	41	225	4.37e-27	104.0	58.38

```
122      32   183      41   225 2.95e-26    102.0     57.84
123      32   173      49   223 2.99e-25     99.8     58.29
124      31   183      29   214 5.57e-25     98.6     58.06
125      32   173      44   218 8.84e-25     98.2     58.86

$url

"https://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Get&FORMAT_OBJECT=Alignment&ALIGNMENT_VIEW=Tabu
```

```
top_hits <- head(blast$hit.tbl,n=3)
top_hits
```

```
        queryid subjectids identity alignmentlength mismatches gapopens q.start
1 Query_5967171     8VRW_B   68.846             260         37       25       2
2 Query_5967171     4AH2_B   74.619             197         15       21      22
3 Query_5967171     3PDO_B   74.619             197         15       21      22
  q.end s.start s.end   evalue bitscore positives mlog.evalue pdb.id    acc
1   217       4   263 1.21e-77      236     73.46    177.1084 8VRW_B 8VRW_B
2   183      33   229 3.31e-66      205     78.17    150.7737 4AH2_B 4AH2_B
3   183       3   199 4.52e-66      203     78.17    150.4621 3PDO_B 3PDO_B
```

```
annotations <- pdb.annotate(top_hits$pdb.id)
```

```
results <- data.frame(
  ID = top_hits$pdb.id,
  Technique = annotations$experimentalTechnique,
  Resolution = annotations$resolution,
  Source = annotations$source,
  Evalue = top_hits$evalue,
  Identity = top_hits$identity
)
results
```

```
      ID Technique Resolution       Source   Evalue Identity
1 8VRW_B        EM       3.03 Homo sapiens 1.21e-77   68.846
2 4AH2_B     X-ray       2.36 Homo sapiens 3.31e-66   74.619
3 3PDO_B     X-ray       1.95 Homo sapiens 4.52e-66   74.619
```

Q9

```
results_dir <- "project_10863_0/"
aln_file <- list.files(path=results_dir,
                       pattern=".a3m$",
                        full.names = TRUE)
aln_file
```

```
[1] "project_10863_0/project_10863_0.a3m"
```

```
aln <- read.fasta(aln_file[1], to.upper = TRUE)
```

```
[1] " ** Duplicated sequence id's: 101 **"
```

```
sim <- conserv(aln)
```

```
con <- consensus(aln, cutoff = 0.5)
con$seq
```

```
  [1] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
 [19] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
 [37] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
 [55] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
 [73] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "S" "-" "-" "G"
 [91] "-" "-" "-" "-" "-" "-" "-" "-" "G" "-" "-" "-" "-" "-" "-" "-" "N" "-"
[109] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[127] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[145] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[163] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[181] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[199] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[217] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[235] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[253] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[271] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[289] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[307] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[325] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[343] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[361] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[379] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
[397] "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-" "-"
```

```
[415] "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_"
[433] "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_"
[451] "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_" "_"
```