

TABLE SI. Classifier Performance for Within-class Tests¹

Methods	Class ²	Errors	FP	FN	Accu- racy	Sensi- tivity	Speci- ficity	FP rate	MCC
SAM	A	2	1	1	1.00	1.00	1.00	0.00	0.99
SVM_Fisher	A	2.5	1.5	1	1.00	1.00	0.99	0.01	0.99
SVM_pairwise	A	2.5	1.5	1	0.99	1.00	0.99	0.01	0.99
SVM_AA(rbf)	A	12.5	8	4.5	0.97	0.98	0.96	0.04	0.94
SVM_AA(pol)	A	18.5	14	4.5	0.95	0.98	0.93	0.07	0.91
SVM_AA(sig)	A	28.5	25	3.5	0.93	0.98	0.88	0.12	0.87
SVM_AA(lin)	A	30	25	5	0.93	0.98	0.88	0.12	0.86
DT	A	33.5	11.5	22	0.92	0.89	0.95	0.06	0.84
SAM	N	7	0	7	0.98	0.91	1	0	0.94
SVM_Fisher	N	0.5	0.5	0	1.00	1	1.00	0.00	1.00
SVM_pairwise	N	8.5	4.5	4	0.97	0.95	0.98	0.02	0.93
SVM_AA(rbf)	N	15	12	3	0.95	0.96	0.94	0.06	0.88
SVM_AA(pol)	N	40	17	23	0.86	0.72	0.92	0.08	0.65
SVM_AA(sig)	N	55.5	6	49.5	0.81	0.39	0.97	0.03	0.48
SVM_AA(lin)	N	56	6.5	49.5	0.81	0.39	0.97	0.0	0.48
DT	N	29.5	14	15.5	0.90	0.81	0.93	0.07	0.75

¹The classification results were not obtained at the MEP, but obtained as the default classifier outputs. For SAM, the e-value threshold of 0.05 was used. Values shown are the average from two independent tests.

²The dataset used to train each classifier. A: Class A, N: non-Class A.

TABLE SII. Classifier Performance for Between-class Tests¹

Methods	Class ²	Errors	FP	FN	Accu- racy	Sensi- tivity	Speci- ficity	FP rate	MCC
SAM	AN	144	0	144	0.61	0.11	1	0	0.26
SVM_Fisher	AN	114	0	114	0.69	0.30	1	0	0.44
SVM_pairwise	AN	108	1	107	0.71	0.34	1.00	0.00	0.46
SVM_AA(rbf)	AN	58	9	49	0.84	0.70	0.96	0.04	0.69
SVM_AA(pol)	AN	65	17	48	0.83	0.70	0.92	0.08	0.65
SVM_AA(sig)	AN	85	31	54	0.77	0.67	0.85	0.15	0.53
SVM_AA(lin)	AN	84	31	53	0.77	0.67	0.85	0.15	0.54
DT	AN	89	14	75	0.76	0.54	0.93	0.07	0.52
SAM	NA	188	2	186	0.54	0.07	0.99	0.01	0.16
SVM_Fisher	NA	156	0	156	0.62	0.22	1	0	0.36
SVM_pairwise	NA	46	10	36	0.89	0.82	0.95	0.05	0.78
SVM_AA(rbf)	NA	43	17	26	0.90	0.87	0.92	0.08	0.79
SVM_AA(pol)	NA	42	42	0	0.90	1.00	0.80	0.20	0.81
SVM_AA(sig)	NA	46	45	1	0.89	1.00	0.79	0.21	0.80
SVM_AA(lin)	NA	46	45	1	0.89	1.00	0.79	0.21	0.80
DT	NA	89	20	69	0.78	0.66	0.90	0.10	0.58

¹The classification results were not obtained at the MEP, but obtained as the default classifier outputs. For SAM, the e-value threshold of 0.05 was used.

²AN: trained on a Class A dataset and tested on a non-Class A dataset, NA: trained on a non-Class A dataset and tested on a Class A dataset.

TABLE SIII. The list of accession numbers of the sequences used in the study¹

Training dataset for Class A: 200 entries										
P08908	P08913	P47899	P08588	P79148	P61168	P61169	P43140	Q25190	P28221	Q91081
Q9ERZ4	P11614	P10980	O77408	P53452	Q9TST5	P54833	P32304	P04274	Q9N2B2	P18599
P28335	Q13639	P50406	P47898	Q923Y6	Q923Y4	Q923Y0	P56439	O77590	Q13725	Q28519
P51686	P32303	O62747	O54689	P33533	P31391	P34993	P49660	O97666	P34997	Q9WV08
Q8BGE9	Q9NSD7	P33766	P30551	P21452	P30553	Q9EP86	P56479	Q924H0	Q9Y5X5	O02835
Q9EQD2	P34992	Q9Y5Y4	P97295	Q90252	Q61125	Q28756	P30992	Q00991	P20789	P32307
Q969V1	Q8K458	O75388	Q864H7	P56446	Q865E8	Q865E9	Q864J1	Q864K1	Q864J0	Q864J6
Q864G2	Q864K8	Q9TUK4	Q864K6	Q864G4	Q864G7	P56450	P70596	P14763	P23945	P32212
P20395	P30730	Q90215	Q9YGZ3	Q28886	P56514	P22328	Q9YGZ2	P15409	P56516	P51488
P35403	P79812	P79901	P51471	P79911	P51474	Q9W6A5	O35478	O18910	Q9R024	P60015
P22331	P87368	Q9WUK7	Q26495	P28680	P24603	O18312	O35214	Q9Z2B3	Q8VGS3	Q8VFD3
Q8VGO3	Q15620	P23266	Q8NH18	Q9GZK3	Q8NGZ4	Q8NGQ6	Q8VEW2	Q9TU99	Q8VFD0	Q8VGI6
Q8NGT7	Q96KK4	Q8NGP8	P37069	Q8NGP6	Q8NGG5	Q8NH51	P37072	Q8NGI8	Q8NGN6	Q8NGS4
Q8NGN0	Q8NGN3	Q9TUA8	Q8NGI7	Q8IXE1	Q8NGC5	Q8NGB6	P30954	P23267	Q8NGX9	Q8NH02
Q96R69	Q8NGA8	Q8NH21	Q8NGW6	Q8NG94	Q8NHC5	Q8NGH9	Q8NGF0	Q8NH67	Q8NGH8	P34979
P43253	Q60612	P28190	P30543	P29274	P79928	P35383	P20272	Q98895	P25105	P30968
Q9MZI6	Q8CH60	Q9TTI8	O42329	P34981	Q92847	P16751	P21453	P46628	Q9JL06	Q9NPC1
P56412	Q9BDS7	Q9H244	P79945	Q64121	Q9BPV8	O35881	P35410	P46095	Q9NS66	Q9JH3
Q9Z2H4	O14842									
Test dataset for Class A: 200 entries										
Q01337	Q24563	Q25322	P23944	P56496	P46636	Q9WU25	P19328	P17200	P18841	P08172
Q9QYN8	P50130	P42289	P25962	P52703	Q61616	P10608	P42291	P28223	O08890	P50128
P16395	P28566	P31387	Q96RI9	Q969N4	Q8HZ64	O97883	P60574	O55193	P56482	P51679
Q9GLN9	P51676	P29755	Q9Z0D9	P35344	Q28003	P41143	P28646	P30872	O08707	O19024
O18983	O00574	Q9BDS6	Q28642	Q04683	P32302	P50052	O08726	P35351	P79242	P79190
Q05394	P97468	Q98982	Q8TDU9	Q04573	O62709	P21555	Q63645	P31392	P47751	P30731
P56494	Q9WU02	P79217	P70536	O55197	Q9UKP6	Q8JZL2	Q90328	Q920E0	P48044	Q9JKL1
P47798	O62729	P56448	Q864I7	Q864H2	Q29154	Q864K4	Q864G1	Q864I2	Q63634	Q864G9
Q920E1	O15354	P33033	P32244	O60883	Q9GLX0	Q95LF5	O02721	P32308	P51470	Q9YGZ8
Q90214	Q9YGZ0	P79914	P79798	P79903	O35476	P22332	P28684	Q9W6A9	O42490	Q17292
Q9QXZ9	Q25158	O16005	P90680	O18486	Q8VFK1	Q8VFL5	Q8NGG7	Q8NGG3	Q8NGV5	Q8NGG4
Q96RD0	Q8NGR5	Q8NGL1	Q8VG05	Q8VG02	Q8N0Y5	Q8NGS5	P23269	Q8VG44	Q9H210	O76100
O14581	Q8NHA8	Q8NGG0	Q8NGS6	O60431	P23270	Q8NH50	Q9Y4A9	Q8NGP3	Q8NGX2	Q8NGA6
Q8NG83	Q8NH00	Q8NGZ3	Q8NGR1	Q8NGB9	Q8NH70	P47883	Q8NGY2	Q8NGE0	Q8NH57	Q8NGH7
Q8NGK3	Q8TCB6	Q8NGI0	Q9H341	Q8NGF1	Q8NH53	Q8NGJ9	Q8NGK2	Q8NGK0	Q8NH63	P43114
P35408	P43119	P34978	P49652	P11616	P11617	P41232	Q9JJS7	Q60614	O13076	P28647
Q61618	Q9ERK9	Q9H1C0	Q99MT6	P30969	Q95MG6	Q95254	O43193	P49285	P16849	O95136
Q9CPV9	P50132	Q15391	Q924T8	Q9EQQ4	P04201	Q924T9	P35412	Q9Z1P4	O00270	P60895
O15552	Q99679									
Training dataset for Non-Class A: 81 entries										
P50133	P25107	O97817	O88923	P41586	Q29627	P35353	P97751	Q28992	O42603	P47872
Q90308	O97831	O94910	O46502	P41588	P49190	P30082	Q02643	P34999	P43220	O35659
Q60755	Q63118	P43218	O08893	Q16602	P43219	P25117	O95838	Q9HBW9	Q923X1	Q9ESC1
Q61549	Q9UHX3	Q80ZF8	O60242	O14514	Q9HCU4	Q9NYQ6	Q9VRN2	Q91ZI0	P83119	O88278
Q9NYQ7	Q9VSE7	Q9W0V7	Q9V5N8	Q14833	P31423	O00222	O15303	P31422	Q9QYS2	P48442
Q9QY96	P31424	P35384	P23385	P97772	Q13255	P31421	Q9WV18	Q9UBS5	Q9Z0U4	O88871
Q24738	Q9JIL6	Q9NQ84	Q8BHL4	Q9NZD1	Q9NZH0	Q8NFF5	P56502	Q92275	P31302	P31303
P31397	P06842	P35352	Q9TX43							
Test dataset for Non-Class A: 81 entries										
Q03431	P25961	O95490	P41593	P35347	P70205	P32215	P30083	P34998	O97827	Q9HAR2
O62772	Q90812	O42602	P32241	Q91085	P47866	Q13324	Q60748	P23811	O88917	P41587
P35000	Q91V95	P47871	P70555	Q61606	P32301	Q02644	Q9R1W5	P48546	P32082	P32214
P30988	P79222	Q9Z0W0	Q91ZE5	P48960	Q16983	P35464	Q9BY15	Q8SQA4	Q14246	Q9Z0M6
O60241	Q9R0M0	Q95NQ0	Q9W0R5	Q95NT6	Q9GT50	P83120	O97148	Q9V818	Q9VXD9	O35161

Q9V817	Q9VS77	Q9VGG8	Q9W0R6	P70579	P47743	P35400	Q14831	P35349	Q14832	P41180
P41594	Q14416	P91685	Q09630	O75899	P22815	Q8K3J9	Q923Z0	P78741	Q99063	P06783
P12384	Q00619	P13773	P34907							

Training dataset for Non-GPCR: 210 entries

P29309	P30462	P04074	Q00897	P33097	Q9U9K0	P35611	P46257	O52196	Q9L7P2	P38448
Q46244	P0A4W0	Q8DZA4	Q9KBQ0	P68729	Q8YGI8	P73997	Q97AR9	P0AB93	O70228	Q9RZN7
P07678	P00283	P26163	P0C0X6	Q9R8E3	O88181	Q9BH13	O23066	Q06110	O65162	P29193
Q05581	P38973	P45033	Q8L373	Q9ZN50	P37852	O23730	P23292	Q8PNI4	P10000	Q36421
Q9AFI1	P59091	P28659	P49488	Q8WBP9	Q9G250	O99256	O20672	P14832	P53355	Q9CIL5
Q88Z64	P0C1F9	P52476	P29507	P16027	Q9W0S9	P0ACB0	Q89B02	P28339	Q8XEC6	P19720
P43136	P33128	P33159	Q7N9B2	Q9KUZ6	P61336	O14134	O35795	Q9Q8Y2	Q9Z191	Q61188
Q88YZ4	Q97B98	Q89AZ6	P07343	Q8R2I0	Q974T4	O27341	O42248	Q03330	P13020	Q890I8
Q8DEF3	P74260	P27442	P42881	P29007	Q7WES2	Q9DCM2	P30107	P23751	P04242	P19016
P68168	Q88B91	Q938W0	Q8P9P1	Q970Y9	P08715	P54871	O29748	Q22633	Q20502	P08833
Q9V3D4	P34167	P23108	P47965	P21183	P27744	Q46893	P04104	Q03215	Q8G3Z8	P19638
Q8Z3G6	P07333	Q10276	P17801	Q8PXV5	P02758	O35704	Q7N1X3	Q51876	P63435	Q834R0
Q8ZGD2	O87077	Q8NNJ0	O51253	Q8P991	P0A153	P16527	O84089	Q9UUZ3	Q49127	Q9LDD8
Q17298	Q975N4	O59488	P16056	O43037	P50757	P20085	Q7VQJ5	P62478	P59963	Q99102
O83325	P01106	P68278	P36161	Q8ZPU5	Q8XIZ1	P18092	Q62645	P26050	P45017	O21519
Q32516	Q95H60	P51267	P29111	Q10067	Q9CEY4	P22045	Q9PUH3	Q9HBI1	P07633	Q97Z84
Q16654	P17946	Q03958	P79208	Q8TK72	Q8XU39	Q08087	P20779	Q9XBG0	Q01970	Q02809
O84817	Q99797	Q98D24	P33554	Q12417	O13958	Q87RV0	P92983	P34124	P09489	Q9BBT1
Q90ZE4	P37178	P02945	P29563	Q57101	O93741	P33742	Q48315	Q48334	O93743	P42196
P42197										

Test dataset for Non-GPCR: 210 entries

Q41418	Q9V2L2	Q67477	P06686	Q9Y2Q0	Q9JT07	Q87GM3	O30273	Q88NC8	P59786	P0A4G0
Q9Y4X0	Q00746	P28039	Q9X1A4	Q9HII9	P63557	Q8E4F3	Q9A0E4	Q05123	Q8U4E6	Q92105
P16000	P42007	O42197	Q9Z5D9	P44422	Q9NSY1	P10845	P51199	P28481	P87072	Q9Y6Q1
Q9A4D6	Q05600	P22002	P30932	P07182	Q00864	Q9Y2V2	P17972	Q96MX0	Q87AA7	P80440
O17403	Q9BR10	Q7M4E9	P08550	O48010	Q9XNV7	Q35621	O47976	Q15438	P19092	Q27883
Q7MWA2	Q8CNH7	P34420	Q8CNW8	P50168	Q949R4	O60884	Q98QY7	Q87EY0	P21482	O75461
O35921	Q9JHL3	Q9UTI6	Q8F983	Q9ZDT7	Q7NC21	P0A6N1	Q61701	P22140	P28919	Q92800
Q08354	P12314	P49936	O30391	Q91X85	P78033	Q9HDU5	P87383	P01861	Q82WQ4	Q83PJ7
Q9PMT4	P09827	P14221	Q8NQV3	Q24570	P24393	P57612	P32235	P05653	P01906	P10058
P02135	P27686	Q8DD85	Q9RRJ5	P58877	Q989E7	P11599	P53772	Q86YZ3	Q8K981	P01786
P08454	Q41741	O94083	P03970	P16297	Q9WTQ8	Q9KP34	Q8G257	P16878	Q55988	Q08415
Q9XB52	P60319	P53233	P06855	Q927E8	P01629	P82791	P27782	P41027	P48573	Q9EVG9
Q9KRA7	Q8K1S1	Q9P8W5	P11150	Q9AC16	P01701	P00701	Q969L2	O32462	P48190	P21949
P61869	P18798	P78003	Q9RA51	Q39659	Q8Y5Z7	Q8ZGW5	P62132	Q8E9P0	Q929Y0	Q97DD9
Q8YM74	O66652	P02154	P12979	P65499	Q60430	P05855	P68739	P60324	Q29498	Q99748
Q9G2W9	P43196	P70581	Q58134	O54340	Q8EJ78	Q7VJY0	P00615	Q13219	P32959	Q92824
P18248	Q9BGI3	P92203	O60760	Q7VZB8	P04917	Q01951	P11393	Q86YI8	P00914	Q59121
P60487	P54679	Q9PL46	P06442	O88554	P67987	P60673	P0C1E4	P18917	P23653	P49481
Q9V0N9	Q8XMP8	P69051	O93740	P94854	P16102	P15647	O93742	Q48314	P25964	P33743
P71411										

Dataset 1 for AR: 63 entries

O42384	Q61224	P79250	P28566	Q25190	Q17239	P30966	Q25321	P22270	Q90WY6	Q01337
Q8JG70	Q90WY4	Q60475	P53453	P21917	Q93126	P32304	P20905	Q4LBB6	O70528	Q24563
O42574	P43141	P54833	P47800	P35406	P42290	Q5QD12	Q5QD07	Q5QD17	Q5QD15	Q9WU25
P97717	P28223	Q02152	P31389	P08912	P32211	P08172	Q9U7D5	Q9H3N8	P50406	Q18007
Q9W6A9	P51491	P87365	P51471	O13227	P51475	O42490	P06002	Q94741	O18312	Q25157
Q26495	P91657	P08255	P24603	Q5YKK9	O15974	Q6U736	P23820			

Dataset 2 for AR: 63 entries

O42385	Q6XXX9	P79748	O08890	P28285	P28286	Q25414	P31387	Q93127	Q17232	Q91081
P35405	P08913	Q90WY5	P20288	P35462	Q8IS44	O77408	Q91559	Q4LBB9	Q9VCZ3	Q76MS7
P07700	Q9TT96	Q9XT57	Q8UUY8	P42291	Q8BLD9	P41596	Q5QD04	O14804	Q5QD16	Q96RJ0
Q91175	P25100	P28335	Q9N2B1	P11229	P41984	P30544	P16395	P58406	Q16950	Q09388
O01670	P87368	Q9W6A8	P28682	Q9W6A5	P22330	O13018	O42266	P08099	O01668	P35360

Q25158 P90680 P04950 O61303 O15973 O14718 Q9H1Y3 P47803

Dataset 1 for PE: 69 entries

P56451 P56445 P34974 P47751 O54799 O62709 Q99JG2 P43142 P51678 Q9BDS8 P51676
Q95NC9 Q9Z0D9 P56484 O00590 Q9R0M1 Q5MD61 O97571 P25024 P25930 P51686 P35350
P25104 Q3BCUO Q61125 Q9NSD7 P35414 Q16581 O88680 O08790 P21462 Q49SQ2 P97468
P30992 P30993 P55086 P47749 O00254 Q96RI0 O88853 Q91V45 Q99705 P30872 P30935
P30874 P79350 P35377 Q61212 Q9GZQ6 P58307 Q5DUB3 P29371 P30975 Q9HB89 O88319
P32239 Q63931 Q90352 Q90252 Q9WU02 Q75W84 Q9GLX0 P35374 Q9Y5Y4 Q4EW11 P49146
Q868T3 P83861 Q8SPN2

Dataset 2 for PE: 70 entries

P32245 P33033 P55167 P32247 P21729 P21450 P32940 Q9QY42 O15218 Q9Z2I3 Q2Y2P0
P41597 Q8MJW8 P51685 O08707 P46094 P46092 Q04683 P35343 P70612 P32248 P51684
Q9BDS6 P32303 Q8HZN9 O70526 Q8TDU9 P79960 O09047 P25090 P25089 O75388 O88416
P79175 O70129 Q9P296 P30558 O08675 O88634 O88854 P56479 Q6BD04 Q8SQ54 P49660
O08858 P32300 P33534 O02813 Q61041 Q924H0 Q9TUP7 Q98982 P05363 O55040 Q58CW4
P70310 P70031 Q9WTV8 P56449 Q90334 O77808 Q95LF4 Q9QUI6 P30974 Q8ITC7 Q9TQ9
O62729 Q9UKP6 P25931 Q8BMP4

Dataset 1 for OL: 154 entries

P23265 Q96RA2 O76099 Q8NG99 P34987 O43749 Q15619 Q8NGR9 O60431 Q8NGA1 Q8NGS2
Q8NH94 Q8NGR5 Q9P1Q5 Q8VFK2 Q8VFK1 Q8NGG2 Q8VFK7 Q8VGS1 Q8NH51 Q8NGG5 Q8VFL5
Q8NGP8 Q8NGP4 Q95155 Q96R09 Q8NGF7 Q8N162 O95221 Q8VEZ0 Q8VG43 Q8NGK9 Q8NGL4
Q60882 Q96RC9 Q15617 Q60895 Q60893 Q8WZ84 Q8NGJ0 Q8NGI8 Q8N127 Q8NH87 Q8NGQ5
Q8NGQ6 Q8NGE7 Q8NGV7 Q8NHB7 Q8NGR6 Q9NQ1 Q8NGS9 Q8NGS4 Q8NGT2 Q8NGH3 Q9UGF6
Q96R47 Q60883 Q9GZK3 Q8NGZ5 Q8NGZ4 Q9GZK4 Q8N628 Q7Z3T1 Q8NGU4 Q8NGZ0 Q96R27
Q96R30 Q8NH16 Q8NG84 Q8NH02 Q8NGX1 Q8NH01 O43869 Q8NGE0 Q9H207 P58181 Q8NGE3
Q8NH19 O60404 Q8NHC4 Q5JRS4 Q8NGX6 Q8NGX5 Q8NGI7 Q8NGC9 Q8NG94 Q8NGX0 Q60889
P34986 Q8NGX9 Q8NGQ2 Q8NGZ6 Q8NGM8 Q8NH40 Q8NGY3 Q8NGY2 Q8NGY6 Q8NGR1 Q60891
Q9TU89 Q60888 Q8NGC3 Q8NH81 Q8NGN2 Q9UGF7 Q8NH83 Q8NGL6 Q8NGM1 Q8NH72 Q8NGP0
Q8NH73 Q8NH49 Q6IEV9 Q8NGB4 Q8NGB9 Q8NGA8 Q8NH41 Q8NGD5 Q8NGD2 Q8NGC6 Q8NGI6
Q15615 Q8NGD0 Q8NH05 Q96RD1 Q9NZP2 Q8NGT5 Q8NHC5 Q9UGF5 Q8NHC7 Q8NH64 Q8NGK0
Q8NH63 Q8NH61 Q9H344 Q9H339 Q9Y5P1 Q9H2C8 Q8NGF3 Q9H255 Q9H2C5 Q8NGI0 Q8NH56
Q8NH57 Q8NGK5 Q96RD2 Q8NGF0 Q8NGJ3 Q8NGH9 Q96RD3 Q9H346 Q8NGK6 Q8NGI1 Q8NGH5

Dataset 2 for OL: 155 entries

O76100 P23268 Q8NG98 Q8NGA0 Q8NG95 P34982 Q8NHA8 Q8NGS0 P47890 Q8NH92 Q60879
P23269 Q8NGR8 Q8NGR3 Q15612 Q8VF76 Q8VFL9 Q8VF13 Q8VGR8 Q8VGR9 Q8NH50 Q8NGP2
Q96RB7 Q8NGP6 Q8VGS3 Q8NGR4 Q96R08 Q8VFX2 Q8NOY5 Q8NGC0 Q8VGI5 Q8VF66 Q8NGL3
Q8NGL2 Q96RD0 P34985 Q60884 P34983 Q8NGM9 Q9GZM6 Q8NGI9 Q95154 P37067 Q8NGQ1
Q8NGE9 Q60894 Q8NH90 Q9NZP5 Q8NH09 Q8NGS7 Q8NGS6 Q8NGV5 Q8NGT1 Q9H210 Q13607
Q8NG92 O95047 P59922 O76000 Q5TZ20 P23275 Q8NGZ6 Q9Y3N9 Q5JQS5 Q8NHA4 Q96R28
Q8NG77 Q8NG97 Q8N349 Q9H205 Q8NGZ9 Q8NGX2 Q8NH04 Q8NHC8 Q9H209 Q8NGE5 Q96KK4
Q60885 Q9Y4A9 Q8NGY0 P30954 Q8NGY1 Q8NGX3 Q8NGQ4 Q8NGF6 Q8NGC7 Q8NGC1 P23270
Q6IFH4 Q95007 Q8NGX8 P23267 Q8NGC5 Q8NGN1 Q8NH79 Q8NGW6 Q8NGY5 Q9GZK7 Q8NGZ3
P47883 Q8NGN7 Q60887 Q8NGC4 Q8NGN3 P58182 Q8NH70 Q6IF82 Q60878 Q8NGB2 Q96R67
Q8NGL7 Q60881 Q8NGF9 Q8NGL9 Q8NGB8 Q6IEY1 Q8NGD3 Q8NH42 Q8NGD4 Q96R72 Q8NH43
Q8NGJ1 Q8NGN0 Q8NGD1 Q8NGC2 Q9NZP0 Q8NGE2 Q8N148 Q8NGZ2 Q8NHC6 Q8NGJ7 Q8NGK1
Q8NH59 Q8NGJ5 Q9H343 Q9H341 Q9H340 Q9Y5P0 Q8NGJ9 Q8TCB6 Q9UKL2 Q8NH53 Q8NGI2
Q8NGH7 Q8NGF1 Q8NGK4 Q8NGJ2 Q8NGK2 Q8NGJ4 Q8NH55 Q8NH60 Q6IF63 Q8NGI3 Q8NH76
Q8NGJ8

Dataset N1: 158 entries

O14514 P30988 Q9R1W5 Q13324 P35464 Q29627 P32241 Q90308 Q03431 P43218 O35659
Q8SQA4 Q9Z0M6 Q9BY15 Q14246 Q923X1 Q97817 Q9NYQ6 Q91ZIO Q97148 Q9V818 Q9VS77
P83119 Q99063 Q13467 P58421 P18537 Q9DEB5 Q61088 Q8WMU5 P91682 Q9NZH0 Q9NZD1
Q9WV18 Q6PCP7 P34907 P35352 P51810 Q5U9X3 Q8K4Z6 Q7RTX1 Q8TE23 Q925D8 P41594
Q14832 Q14833 Q68ED2 Q09630 P06842 Q646D7 Q9NYW3 Q9NYW1 Q7TQA7 P59531 Q646D9
Q645Z5 Q5Y500 Q645Z9 Q7TQA8 Q7TQA6 Q646G7 Q646C8 Q7TQB8 Q646D3 Q646D1 Q7TQB0
Q5Y4Z0 Q7TQA9 P59532 Q7Z5H5 Q9BXE9 Q9FKY5 O22757 Q9SXB6 O49873 O49914 O80580

Q94KB2	O23693	Q19992	O17240	P91209	O01608	Q19509	O17823	O17822	Q19505	Q19473
Q19061	Q6JAN0	Q9VVF3	Q9V589	Q9VAW0	P81914	P81909	Q9W1P7	Q9V9I2	P81918	Q9VCS8
P81919	Q9VLE5	Q8WTE6	Q9VHQ7	Q9VHQ2	Q9VT90	P82985	P82982	P81924	P83292	Q9VM08
Q8IMN5	P58957	P58959	Q8IMQ6	P58953	P84180	P58954	Q9VJF2	P58985	P83297	Q9W0M2
P83295	Q8IN23	Q9W1U5	Q9VYZ1	Q9W2U9	P81911	P81922	P81912	O46077	O77438	O17800
Q9V568	Q9VHE6	Q8WTE7	Q9W0R6	P31397	Q9VKJ7	Q19975	P58960	Q8IRL8	Q9VSH2	Q9VT92
Q9VFN2	Q9VGG8	Q9W1N6	Q9VZW8	Q9VNB5	Q8SYV9	Q8IQ72	Q9VEU0	Q9VD76	Q8IPU5	Q9V969
Q9W2B1	Q9V4Q0	Q9VZL7	P91536							

Dataset N2: 158 entries

O60242	O60241	P25117	P35347	Q16983	Q02643	P41588	Q91085	P47872	Q91V95	Q61606
O95838	P48960	Q9UHX3	Q91ZE5	Q61549	O97831	O97827	Q9HCU4	Q9W0R5	Q9VRN2	P83118
Q9VSE7	Q92275	P31302	Q9H461	Q9VVX3	O70421	O00144	Q9NPG1	P56726	P22815	Q9NQ84
Q8NFX5	O75899	Q8NFX8	Q9TX43	P13773	Q9QY96	Q9PW88	Q99PG6	Q925I4	Q7RTX0	P97772
Q14416	P91685	P47743	O15303	P12384	Q9JKT4	P59530	Q646E5	Q645Y5	Q5Y4Z2	P59528
Q7TQA4	Q7M720	Q646E4	Q646A0	Q697L6	Q9JKT2	Q645Z1	Q646D4	Q7TQA5	Q9JKT3	P59529
Q7TQB9	Q646D0	Q646C7	Q8WN92	Q8NFX6	Q7YRP1	O22752	O22815	O80961	P93766	Q94KB9
O49621	Q9FI00	P92001	Q21767	P92015	P91210	O01609	O17821	Q19507	Q19508	Q19474
Q09572	Q8WXG9	Q9V3Q2	P81913	P81921	Q9I816	P81915	Q9V9I4	Q9VAZ3	Q9VHS4	Q9VCS9
Q9VUK5	Q9V3N2	Q9V6H2	P81917	Q9VHQ6	Q9VT08	Q9VU27	P82983	Q9VNB3	Q9VQE7	Q9VM09
Q8IMN6	P58956	P58958	Q8IMZ5	P84181	P58951	P58952	P58955	Q8INZ2	Q9W497	P83296
P83293	Q9VZJ6	Q9VD74	Q8IN58	P81916	P83294	Q9W3I5	Q17760	P81923	P78741	P82986
Q9NBW1	Q9V6A9	Q9VXL0	Q9W0V7	Q9W5G6	P06783	Q9VTN0	Q00619	Q9V8Y7	Q9VKA5	Q9V4K2
Q9VNK9	Q9VB30	P58950	Q9W2B2	Q9VT20	O16975	Q8I1F0	Q9W594	Q9W367	Q9VYZ2	Q8INM9
P58962	Q9W1N5	Q8INZ1	Q9VPT1							

¹ See Tables 2 and 4 of the main text for the description of each dataset. The GPCR sequence data were originally obtained from GPCRDB July 2004 release. Sequences are available from: <http://bioinfolab.unl.edu/emlab/gpcr/index.html>