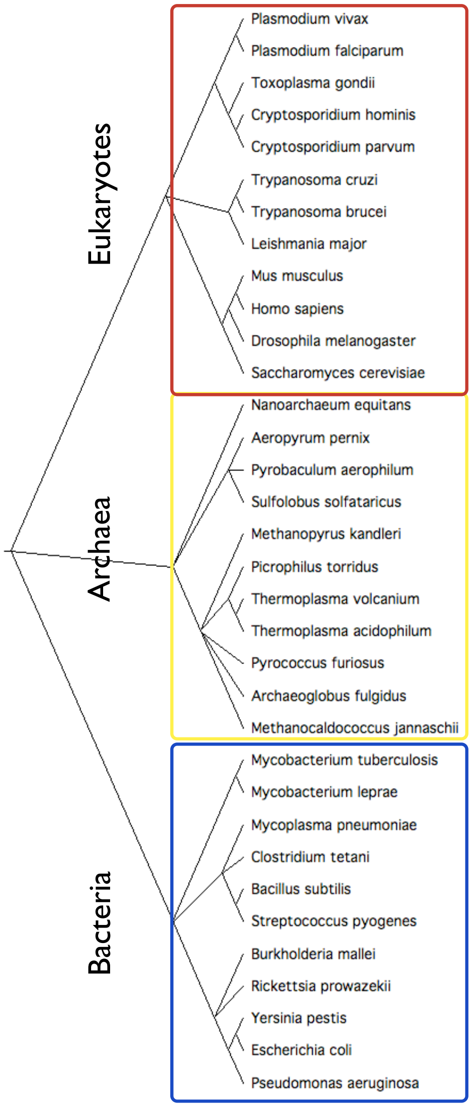
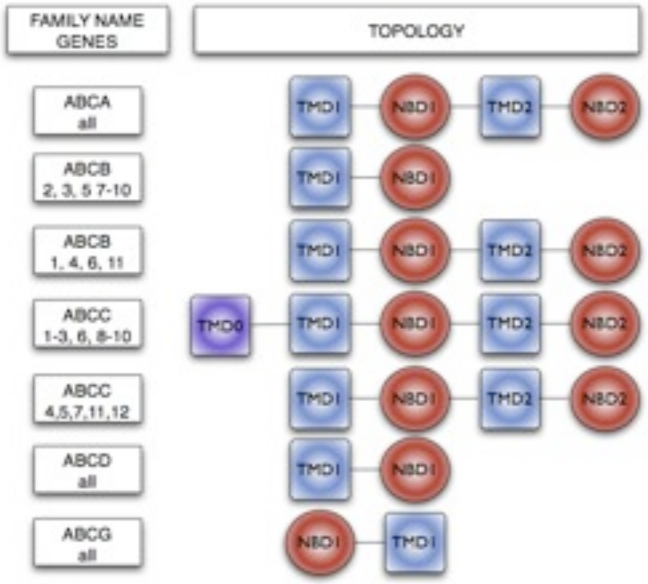
A survey of integral membrane proteins:

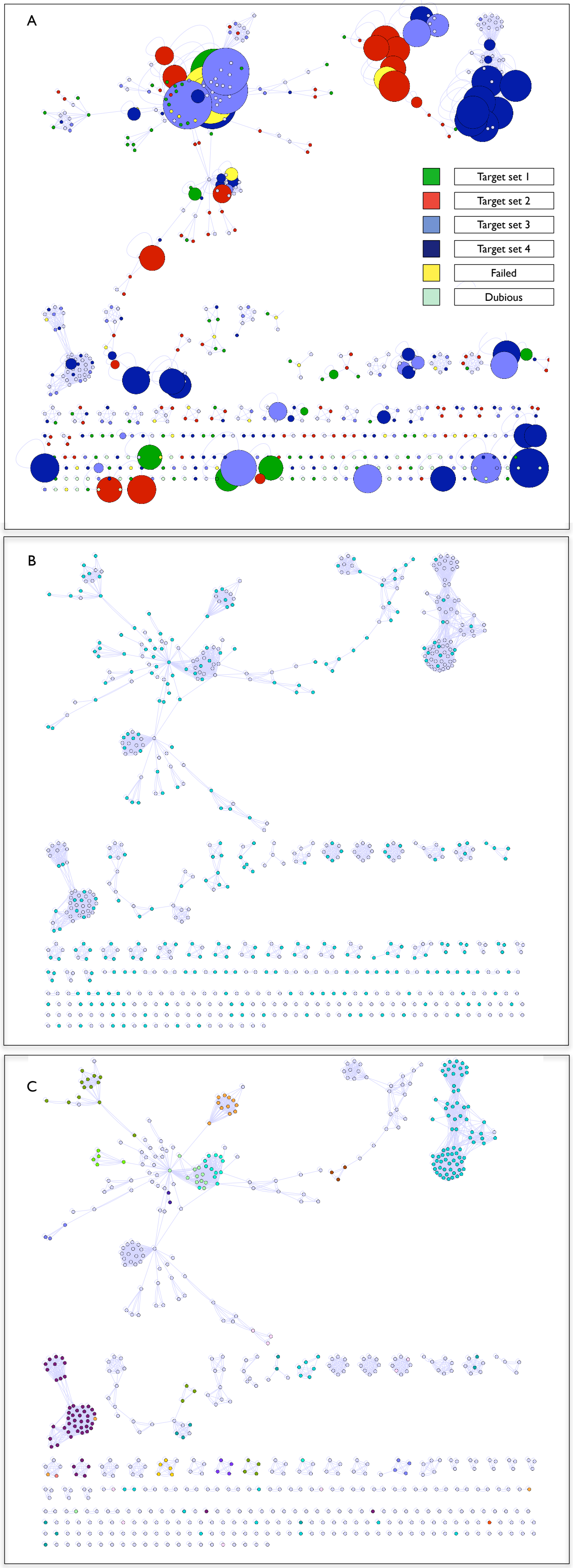
Supplementary figures and tables



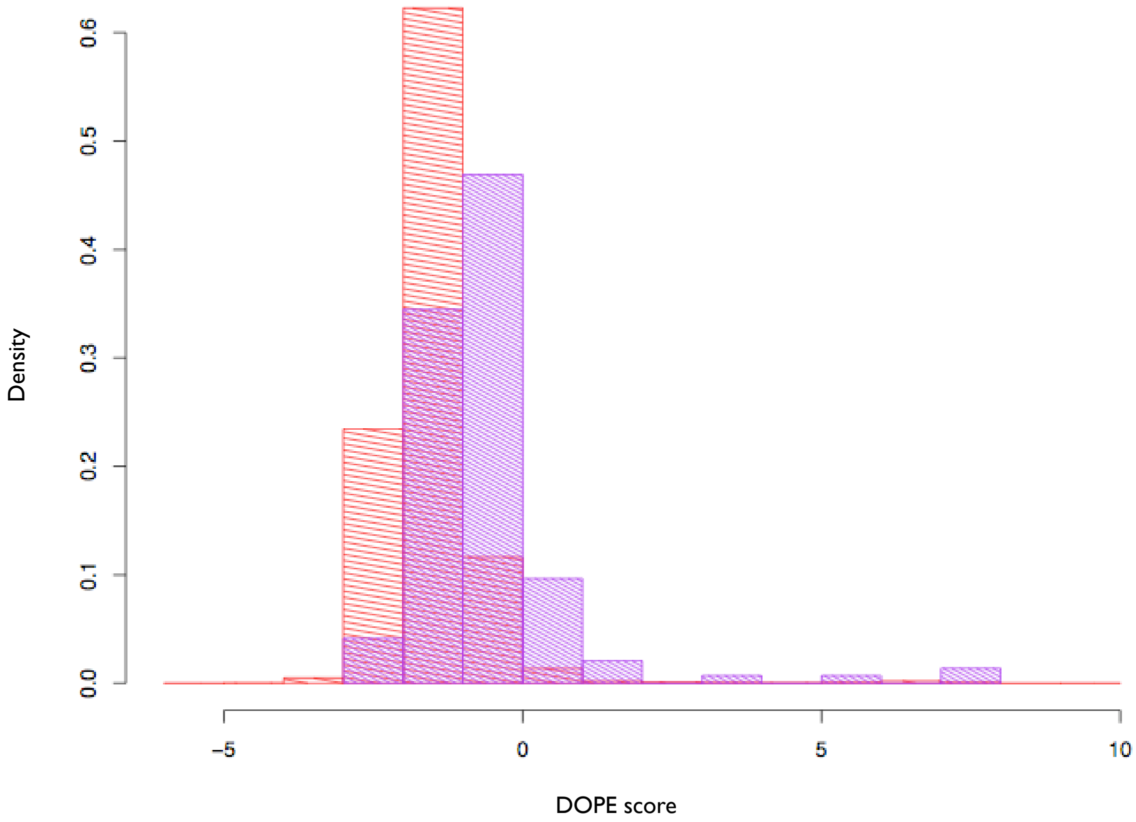
Supplementary Figure 1: 34 selected organisms.



Supplementary Figure 2: Human ABC transporter domain structure. The human ABC transporter superfamily consists of 48 proteins with varying numbers of transmembrane domains (TMD, purple and cyan) and globular nucleotide binding domains (NBD, red). There are both full transporters, such as ABCB1 (P-glycoprotein) which have four domains, two TMDs and two TMHs, and half transporters such as ABCD1 (ALD) which have one TMD and one NBD per polypeptide chain. ABCG family members have the NBD N-terminal to the TMD, in contrast to the other families.


A
🟩 Target set 1
🟥 Target set 2
🟦 Target set 3
🟦 Target set 4
🟨 Failed
🟩 Dubious
B
C

Supplementary Figure 3: Membrane protein target selection in S. cerevisiae. A) There are 621 predicted IM proteins in yeast, four sets of 96 targets, colored green, red, light blue, and dark blue were selected from this set for a structural genomics project in yeast. Each of the IM proteins is represented as a node, and nodes are connected if their sequence profiles are significantly related and are colored by their target set. Dubious targets (yellow) were sequences that the Saccharomyces Genome Database annotated as “Dubious,” which SGD defines as an open reading frame that is unlikely to encode an expressed protein. Failed targets (mint green) were sequences that failed any step of the experimental pipeline, for example, sequences that could not be cloned. Node size indicates the number of sequences homologous to the target. B) Yeast IM proteins with human homologs are shown in cyan. C) Pfam-defined clans mapped to Yeast IM proteins.



Supplementary Figure 4: ZDOPE statistical potential scores for native protein structures solved using X-ray crystallography. The distribution of Z-scores for integral alpha-helical membrane proteins (n=145, blue) is shifted higher than the scores for globular proteins (n=36786, red). The average Z-score for a membrane protein is -0.63, while the score for the globular proteins is -1.56. We use a Z-score cutoff of 0 when assessing models of membrane proteins.

Supplementary Table 1: 598 Pfam families predicted by TMHMM to have three or more transmembrane helices

Pfam_ID	**Pfam_description**
PF00001.12	7tm_1
PF00002.15	7tm_2
PF00003.12	7tm_3
PF00029.9	Connexin
PF00032.8	Cytochrom_B_C
PF00033.9	Cytochrom_B_N
PF00060.17	Lig_chan
PF00083.14	Sugar_tr
PF00115.10	COX1
PF00119.11	ATP-synt_A
PF00124.9	Photo_RC
PF00146.11	NADHdh
PF00209.9	SNF
PF00223.9	PsaA_PsaB
PF00230.10	MIP
PF00324.11	AA_permease
PF00335.10	Tetraspannin
PF00344.11	SecY
PF00361.10	Oxidored_q1
PF00375.9	SDF
PF00420.14	Oxidored_q2
PF00421.9	PSII
PF00474.8	SSF
PF00487.14	FA_desaturase
PF00499.10	Oxidored_q3
PF00510.9	COX3
PF00520.21	Ion_trans
PF00528.12	BPD_transp_1
PF00654.10	Voltage_CLC
PF00664.13	ABC_membrane
PF00689.11	Cation_ATPase_C
PF00771.10	FHIPEP
PF00810.9	ER_lumen_recept
PF00813.10	FliP
PF00822.11	PMP22_Claudin
PF00854.12	PTR2
PF00873.9	ACR_tran
PF00876.9	Innexin
PF00892.11	DUF6
PF00893.10	Multi_Drug_Res
PF00902.8	TatC
PF00909.11	Ammonium_transp
PF00950.8	ABC-3
PF00951.9	Arteri_Gl

PF00953.11	Glycos_transf_4
PF00955.11	HCO3_cotransp
PF00999.11	Na_H_Exchanger
PF01001.9	HCV_NS4b
PF01002.9	Flavi_NS2B
PF01005.9	Flavi_NS2A
PF01007.10	IRK
PF01027.11	UPF0005
PF01032.9	FecCD
PF01036.9	Bac_rhodopsin
PF01040.9	UbiA
PF01045.8	EIAV_GP45
PF01061.14	ABC2_membrane
PF01062.11	Bestrophin
PF01080.8	Presenilin
PF01098.10	FTSW_RODA_SPOVE
PF01124.9	MAPEG
PF01127.12	Sdh_cyt
PF01148.10	CTP_transf_1
PF01151.9	ELO
PF01184.10	Grp1_Fun34_YaaH
PF01219.9	DAGK_prokar
PF01222.8	ERG4_ERG24
PF01226.8	Form_Nir_trans
PF01235.8	Na_Ala_symp
PF01252.9	Peptidase_A8
PF01275.9	Myelin_PLP
PF01284.13	MARVEL
PF01292.10	Ni_hydr_CYTB
PF01306.10	LacY_symp
PF01311.11	Bac_export_1
PF01312.10	Bac_export_2
PF01349.8	Flavi_NS4B
PF01350.8	Flavi_NS4A
PF01384.11	PHO4
PF01461.8	7tm_4
PF01478.8	Peptidase_A24
PF01490.8	Aa_trans
PF01496.9	V_ATPase_I
PF01528.7	Herpes_glycop
PF01534.7	Frizzled
PF01538.8	HCV_NS2
PF01545.11	Cation_efflux
PF01554.9	MatE
PF01566.9	Nramp
PF01569.12	PAP2
PF01578.11	Cytochrom_C_asm
PF01589.7	Alpha_E1_glycop
PF01594.7	UPF0118
PF01595.10	DUF21

PF01598.8	Sterol_desat
PF01602.10	Adaptin_N
PF01621.8	Fusion_gly_K
PF01635.9	Corona_M
PF01654.8	Bac_Ubq_Cox
PF01694.12	Rhomboid
PF01699.14	Na_Ca_ex
PF01733.8	Nucleoside_tran
PF01748.15	Serpentine_recp
PF01757.12	Acyl_transf_3
PF01769.7	MgtE
PF01770.9	Folate_carrier
PF01790.9	LGT
PF01794.9	Ferric_reduct
PF01810.9	LysE
PF01889.8	DUF63
PF01891.7	CbiM
PF01901.7	DUF70
PF01914.8	MarC
PF01925.9	DUF81
PF01940.7	DUF92
PF01943.8	Polysacc_synt
PF01944.7	DUF95
PF01956.7	DUF106
PF01970.7	DUF112
PF01988.9	DUF125
PF02028.8	BCCT
PF02040.6	ArsB
PF02076.6	STE3
PF02077.6	SURF4
PF02096.11	60KD_IMP
PF02101.6	Ocular_alb
PF02109.6	DAD
PF02117.7	Sra
PF02118.11	Srg
PF02163.12	Peptidase_M50
PF02175.7	Srb
PF02233.7	PNTB
PF02293.6	AmiS_UreI
PF02300.8	Fumarate_red_C
PF02308.7	MgtC
PF02313.8	Fumarate_red_D
PF02322.6	Cyto_ox_2
PF02340.6	PRRSV_Env
PF02355.7	SecD_SecF
PF02361.7	CbiQ
PF02364.6	Glucan_synthase
PF02366.9	PMT
PF02378.9	PTS_EIIC
PF02405.7	DUF140

PF02411.6	MerT
PF02417.6	Chromate_transp
PF02447.7	GntP_permease
PF02453.8	Reticulon
PF02460.9	Patched
PF02461.6	AMO
PF02487.8	CLN3
PF02508.5	Rnf-Nqr
PF02516.5	STT3
PF02537.6	CRCB
PF02544.7	Steroid_dh
PF02554.5	CstA
PF02592.5	DUF165
PF02600.7	DsbB
PF02605.5	PsaL
PF02628.6	COX15-CtaA
PF02632.5	BioY
PF02652.5	Lactate_perm
PF02653.6	BPD_transp_2
PF02654.6	CobS
PF02656.6	DUF202
PF02660.6	DUF205
PF02665.5	Nitrate_red_gam
PF02667.5	SCFA_trans
PF02673.9	BacA
PF02674.6	Colicin_V
PF02681.5	DUF212
PF02683.6	DsbD
PF02687.11	FtsX
PF02688.11	DUF215
PF02690.6	Na_Pi_cotrans
PF02694.6	UPF0060
PF02705.7	K_trans
PF02714.6	DUF221
PF02932.6	Neur_chan_memb
PF02949.11	7tm_6
PF02967.5	Fum_red_TM
PF02990.7	EMP70
PF03006.10	HlyIII
PF03010.5	GP4
PF03023.5	MVIN
PF03025.5	Papilloma_E5
PF03030.7	H_PPase
PF03034.6	PSS
PF03040.5	CemA
PF03053.5	Corona_NS3b
PF03062.9	MBOAT
PF03073.6	TspO_MBR
PF03076.5	GP3
PF03092.7	BT1

PF03094.6	Mlo
PF03116.5	NQR2_RnfD_RnfE
PF03124.5	EXS
PF03125.8	Sre
PF03137.10	OATP
PF03151.7	TPT
PF03155.6	Alg6_Alg8
PF03169.6	OPT
PF03186.4	CobD_Cbib
PF03188.6	Cytochrom_B561
PF03189.4	DUF270
PF03203.5	MerC
PF03208.9	PRA1
PF03209.6	PUCC
PF03219.5	TLC
PF03222.4	Trp_Tyr_perm
PF03239.4	FTR1
PF03248.4	Rer1
PF03253.5	UT
PF03268.5	DUF267
PF03303.4	WTF
PF03311.4	Cornichon
PF03323.4	GerA
PF03334.5	PhaG_MnhG_YufB
PF03348.6	Serinc
PF03379.4	CcmB
PF03383.6	Serpentine_r_xa
PF03390.6	Na_citrate
PF03402.4	V1R
PF03419.4	Peptidase_U4
PF03530.4	SK_channel
PF03547.8	Mem_trans
PF03552.5	Cellulose_synt
PF03553.4	Na_H_antiporter
PF03591.5	AzlC
PF03594.4	BenE
PF03595.8	C4dic_mal_tran
PF03596.4	Cad
PF03600.6	CitMHS
PF03601.5	Cons_hypoth698
PF03605.4	DcuA_DcuB
PF03606.6	DcuC
PF03608.4	EII-GUT
PF03609.5	EII-Sor
PF03611.5	EIIC-GAT
PF03613.5	EIID-AGA
PF03616.4	Glt_symporter
PF03619.7	DUF300
PF03626.4	COX4_pro
PF03631.5	Ribonuclease_BN

PF03647.4	TMEM14
PF03649.4	UPF0014
PF03653.4	UPF0093
PF03661.4	UPF0121
PF03694.4	Erg28
PF03699.4	UPF0182
PF03706.4	UPF0104
PF03739.4	YjgP_YjgQ
PF03741.6	TerC
PF03772.6	Competence
PF03773.4	DUF318
PF03788.5	LrgA
PF03798.6	LAG1
PF03806.4	ABG_transport
PF03812.4	KdgT
PF03814.6	KdpA
PF03817.4	MadL
PF03820.8	Mtc
PF03821.6	Mtp
PF03824.6	NicO
PF03825.7	Nuc_H_symport
PF03842.4	Silic_transp
PF03845.4	Spore_permease
PF03857.4	Colicin_im
PF03862.4	SpoVA
PF03866.4	HAP
PF03878.6	YIF1
PF03899.6	ATP_synt_I
PF03901.7	Glyco_transf_22
PF03916.5	NrfD
PF03956.4	DUF340
PF03977.4	OAD_beta
PF03982.3	DAGAT
PF04018.4	DUF368
PF04020.4	DUF360
PF04039.4	MnhB
PF04061.5	ORMDL
PF04080.3	Per1
PF04087.4	DUF389
PF04093.3	MreD
PF04103.5	CD20
PF04109.6	APG9
PF04114.4	Gaa1
PF04120.3	Iron_permease
PF04123.3	DUF373
PF04133.4	Vps55
PF04138.4	GtrA
PF04142.6	Nuc_sug_transp
PF04144.4	SCAMP
PF04145.6	Ctr

PF04148.4	DUF396
PF04160.3	Borrelia_orfX
PF04161.4	Arv1
PF04163.3	Tht1
PF04165.3	DUF401
PF04171.3	DUF405
PF04172.6	LrgB
PF04173.4	DoxD
PF04178.3	Got1
PF04188.3	Mannosyl_trans2
PF04191.3	PEMT
PF04206.3	MtrE
PF04207.3	MtrD
PF04211.3	MtrC
PF04215.3	SgaT_UlaA
PF04224.3	DUF417
PF04235.3	DUF418
PF04238.3	DUF420
PF04240.3	DUF422
PF04247.3	SirB
PF04258.4	Peptidase_A22B
PF04276.3	DUF443
PF04279.6	IspA
PF04284.3	DUF441
PF04290.3	DctQ
PF04298.3	Zn_peptidase_2
PF04306.3	DUF456
PF04323.3	DUF474
PF04332.5	DUF475
PF04333.3	VacJ
PF04342.3	DUF486
PF04346.3	EutH
PF04387.5	PTPLA
PF04401.4	DUF540
PF04403.3	PqiA
PF04474.3	DUF554
PF04479.3	RTA1
PF04506.4	Rft-1
PF04511.5	DER1
PF04515.3	DUF580
PF04535.3	DUF588
PF04544.3	Herpes_UL20
PF04547.3	DUF590
PF04549.5	CD47
PF04550.3	Phage_holin_2
PF04600.3	DUF571
PF04602.3	Arabinose_trans
PF04608.3	PgpA
PF04610.4	TrbL
PF04632.3	FUSC

PF04633.3	Herpes_BMRF2
PF04647.5	AgrB
PF04654.3	DUF599
PF04657.4	DUF606
PF04750.5	Far-17a_AIG1
PF04756.4	OST3_OST6
PF04789.6	DUF621
PF04791.6	LMBR1
PF04819.3	DUF716
PF04884.5	DUF647
PF04893.7	Yip1
PF04896.3	AmoC
PF04906.4	Tweety
PF04923.3	Ninjurin
PF04955.3	HupE_UreJ
PF04973.3	NMN_transporter
PF04976.3	DmsC
PF04982.4	HPP
PF04987.4	PigN
PF05007.3	Mannosyl_trans
PF05024.5	Gpi1
PF05038.3	Cytochrom_B558a
PF05072.4	Herpes_UL43
PF05078.3	DUF679
PF05090.4	VKG_Carbox
PF05095.3	DUF687
PF05106.3	Phage_holin_3
PF05128.3	DUF697
PF05140.4	ResB
PF05145.3	AmoA
PF05197.4	DUF714
PF05208.4	ALG3
PF05216.4	UNC-50
PF05230.2	MASE2
PF05241.2	EBP
PF05255.2	UPF0220
PF05313.3	Pox_P21
PF05346.2	DUF747
PF05425.3	CopD
PF05461.2	ApoL
PF05462.2	Dicty_CAR
PF05468.2	Bac_ATP_syn_I
PF05478.2	Prominin
PF05512.2	AWPM-19
PF05513.2	TraA
PF05514.2	HR_lesion
PF05525.3	Branch_AA_trans
PF05529.2	Bap31
PF05562.2	WCOR413
PF05571.3	DUF766

PF05579.4	Peptidase_S32
PF05602.3	CLPTM1
PF05620.2	DUF788
PF05628.3	Borrelia_P13
PF05631.4	DUF791
PF05640.4	DUF798
PF05653.4	DUF803
PF05656.4	DUF805
PF05661.3	DUF808
PF05675.3	DUF817
PF05684.3	DUF819
PF05745.2	CRPA
PF05758.2	Ycf1
PF05776.3	Papilloma_E5A
PF05805.2	L6_membrane
PF05817.4	Ribophorin_II
PF05832.3	DUF846
PF05857.2	TraX
PF05858.2	BIV_Env
PF05875.3	aPHC
PF05884.3	DUF856
PF05934.2	MCLC
PF05940.2	NnrS
PF05967.2	DUF887
PF05975.3	EcsB
PF05976.2	DUF893
PF05977.4	DUF894
PF05982.3	DUF897
PF05987.3	DUF898
PF05992.3	SbmA_BacA
PF06011.3	TRP
PF06027.3	DUF914
PF06077.2	LR8
PF06081.2	DUF939
PF06105.3	Aph-1
PF06123.3	CreD
PF06126.2	Herpes_LAMP2
PF06127.2	DUF962
PF06139.2	BphX
PF06146.3	PsiE
PF06149.2	DUF969
PF06151.4	Trehalose_recp
PF06157.2	DUF973
PF06161.2	DUF975
PF06166.2	DUF979
PF06168.2	DUF981
PF06173.3	DUF986
PF06177.2	DUF988
PF06181.2	DUF989
PF06182.2	DUF990

PF06195.3	DUF996
PF06197.3	DUF998
PF06237.3	DUF1011
PF06244.2	DUF1014
PF06271.2	RDD
PF06379.3	RhaT
PF06396.2	AGTRAP
PF06403.2	Lamprin
PF06423.2	GWT1
PF06432.2	GPI2
PF06450.2	NhaB
PF06454.2	DUF1084
PF06455.2	NADH5_C
PF06472.5	ABC_membrane_2
PF06496.2	DUF1097
PF06532.2	DUF1109
PF06539.2	DUF1112
PF06541.2	DUF1113
PF06550.2	DUF1119
PF06561.2	DUF1123
PF06570.2	DUF1129
PF06609.4	TRI12
PF06610.3	DUF1144
PF06638.2	Strabismus
PF06653.2	DUF1164
PF06656.2	Tenui_PVC2
PF06664.2	DUF1171
PF06681.4	DUF1182
PF06687.2	SUR7
PF06691.2	DUF1189
PF06699.2	PIG-F
PF06728.3	PIG-U
PF06736.2	DUF1211
PF06738.2	DUF1212
PF06749.3	DUF1218
PF06762.4	DUF1222
PF06770.2	Arif-1
PF06772.2	LtrA
PF06775.4	DUF1226
PF06790.2	UPF0259
PF06797.2	DUF1229
PF06800.3	Sugar_transport
PF06813.4	Nodulin-like
PF06814.3	Lung_7-TM_R
PF06826.3	Asp-Al_Ex
PF06836.2	DUF1240
PF06899.2	WzyE
PF06912.2	DUF1275
PF06921.2	VIRB2
PF06930.2	DUF1282

PF06933.2	SSP160
PF06939.2	DUF1286
PF06942.2	GlpM
PF06946.2	Phage_holin_5
PF06947.3	DUF1290
PF06965.2	Na_H_antiport_1
PF06993.2	DUF1304
PF07019.2	Rab5ip
PF07062.3	Clc-like
PF07074.3	TRAP-gamma
PF07077.2	DUF1345
PF07086.2	DUF1352
PF07095.2	IgaA
PF07098.2	DUF1360
PF07099.2	DUF1361
PF07114.2	DUF1370
PF07123.2	PsbW_2
PF07136.2	DUF1385
PF07155.2	DUF1393
PF07158.2	MatC_N
PF07168.2	FAE_3-kCoA_syn1
PF07185.2	DUF1404
PF07186.2	TraB_2
PF07187.2	DUF1405
PF07220.2	DUF1420
PF07226.2	DUF1422
PF07242.2	DUF1430
PF07256.2	DUF1435
PF07260.2	ANKH
PF07264.2	EI24
PF07274.2	DUF1440
PF07281.2	INSIG
PF07284.2	BCHF
PF07291.2	MauE
PF07298.2	NnrU
PF07301.2	DUF1453
PF07314.2	DUF1461
PF07331.2	DUF1468
PF07343.2	DUF1475
PF07344.2	Amastin
PF07399.2	DUF1504
PF07402.2	Herpes_U26
PF07415.2	Herpes_LMP2
PF07444.2	Ycf66_N
PF07456.2	Hpre_diP_synt_I
PF07457.2	DUF1516
PF07556.2	DUF1538
PF07613.2	DUF1576
PF07662.3	Nucleos_tra2_C
PF07666.2	MpPF26

PF07667.2	DUF1600
PF07672.4	MFS_Mycoplasma
PF07681.2	DoxX
PF07690.6	MFS_1
PF07694.3	5TM-5TMR_LYT
PF07695.2	7TMR-DISM_7TM
PF07698.2	7TM-7TMR_HD
PF07758.2	DUF1614
PF07760.2	DUF1616
PF07770.3	SFT2
PF07779.2	Cas1p
PF07786.3	DUF1624
PF07843.2	DUF1634
PF07851.3	TMPIT
PF07854.2	DUF1646
PF07856.3	DUF1650
PF07857.3	DUF1632
PF07884.4	VKOR
PF07895.2	DUF1673
PF07907.2	YibE_F
PF07916.2	TraG_N
PF07937.2	DUF1686
PF07947.4	YhhN
PF07948.2	Nairovirus_M
PF08006.2	DUF1700
PF08016.2	PKD_channel
PF08019.2	DUF1705
PF08229.2	SHR3_chaperone
PF08395.2	7tm_7
PF08426.1	ICE2
PF08449.2	UAA
PF08507.1	COPI_assoc
PF08551.1	DUF1751
PF08552.1	DUF1753
PF08560.1	DUF1757
PF08570.1	DUF1761
PF08690.1	GET2
PF08733.1	PalH
PF08739.1	PalI
PF09323.1	DUF1980
PF09335.1	SNARE_assoc
PF00695.9	vMSA
PF00860.11	Xan_ur_permease
PF00916.10	Sulfate_transp
PF00939.9	Na_sulph_symp
PF02133.6	Transp_cyt_pur
PF02386.7	TrkH
PF04922.3	DIE2_ALG10
PF05231.5	MASE1
PF05296.3	TAS2R

PF05297.2	Herpes_LMP1
PF05978.7	DUF895
PF06808.2	DctM
PF07290.2	DUF1449

Supplementary Table 2

Comparative modeling leverage of seven membrane proteins solved by the Center for the Structures of Membrane Proteins.

PDB Code	**Annotation**	**Number of UniProt sequences modeled**	**Percentage of residues modeled with higher sequence identity using this template**
2ns1	Chain A, crystal structure of the *E. coli* ammonia channel Amtb complexed with the signal transduction protein Glnk	2086	16
3bhs	Ammonium transporter family (Rh-like protein)	2009	17
2v8n	Chain A, wild-type structure of lactose permease	5957	24
2f2b	Chain A, crystal structure of integral membrane protein aquaporin Aqpm	2156	8
2cfq	Chain A, sugar free lactose permease at Neutral pH	5814	23
2o9d	Chain B, crystal structure of Aqpz mutant T183C	2037	8
3c02	Chain A, Aquaglyceroporin from *Plasmodium falciparum*	1991	8

Supplementary Table 3

Pfam families implicated in multidrug resistance phenotypes

Pfam_ID	Name	Description
PF00664	ABC_membrane	ABC transporter transmembrane region
PF00873	ACR_tran	AcrB/AcrD/AcrF family
PF00893	Multi_Drug_Res	Small Multidrug Resistance protein
PF01061	ABC2_membrane	ABC-2 type transporter
PF01554	MatE	MatE
PF03821	Mtp	Golgi 4-transmembrane spanning transporter
PF06472	ABC_membrane_2	ABC transporter transmembrane region 2