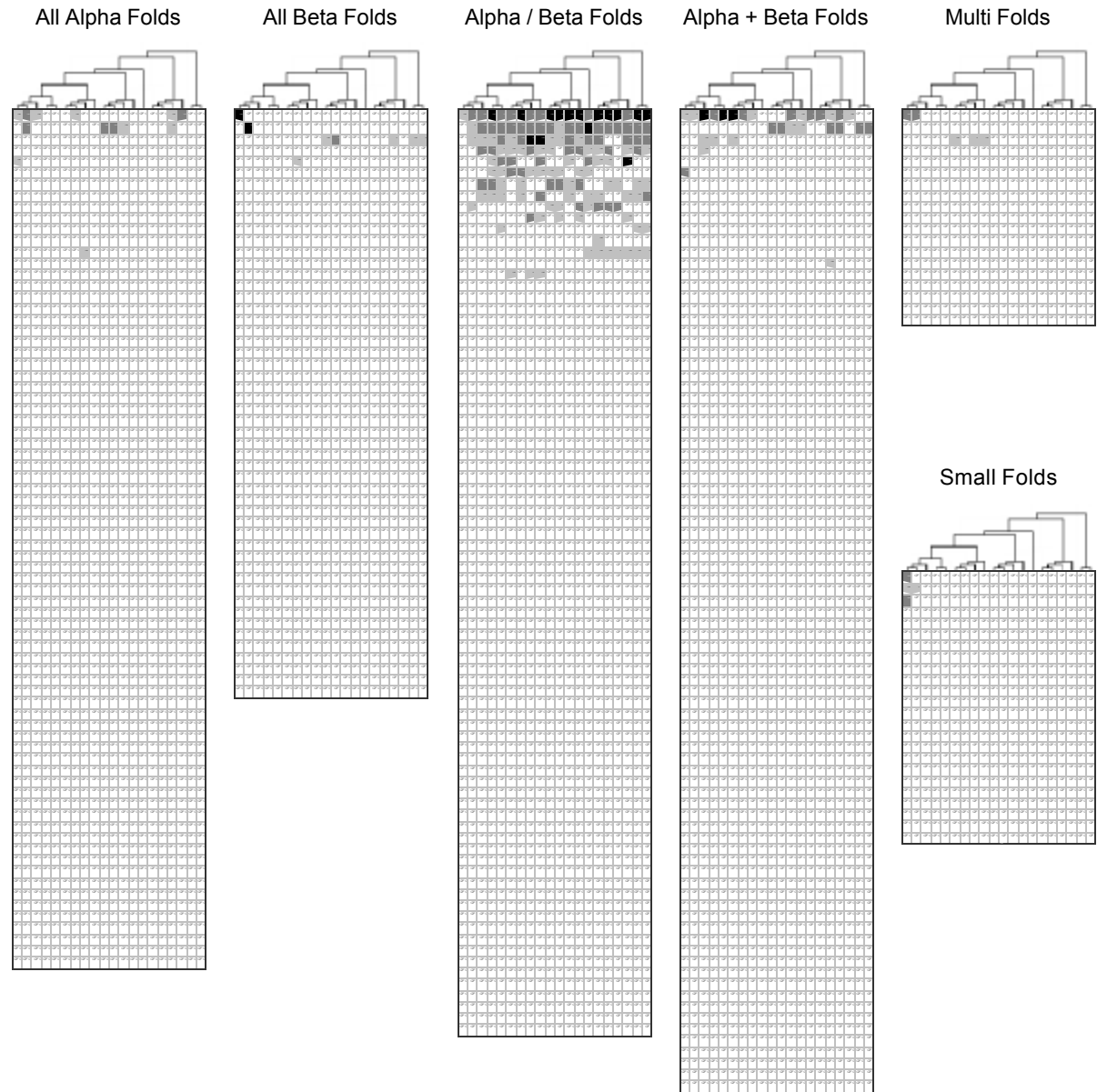
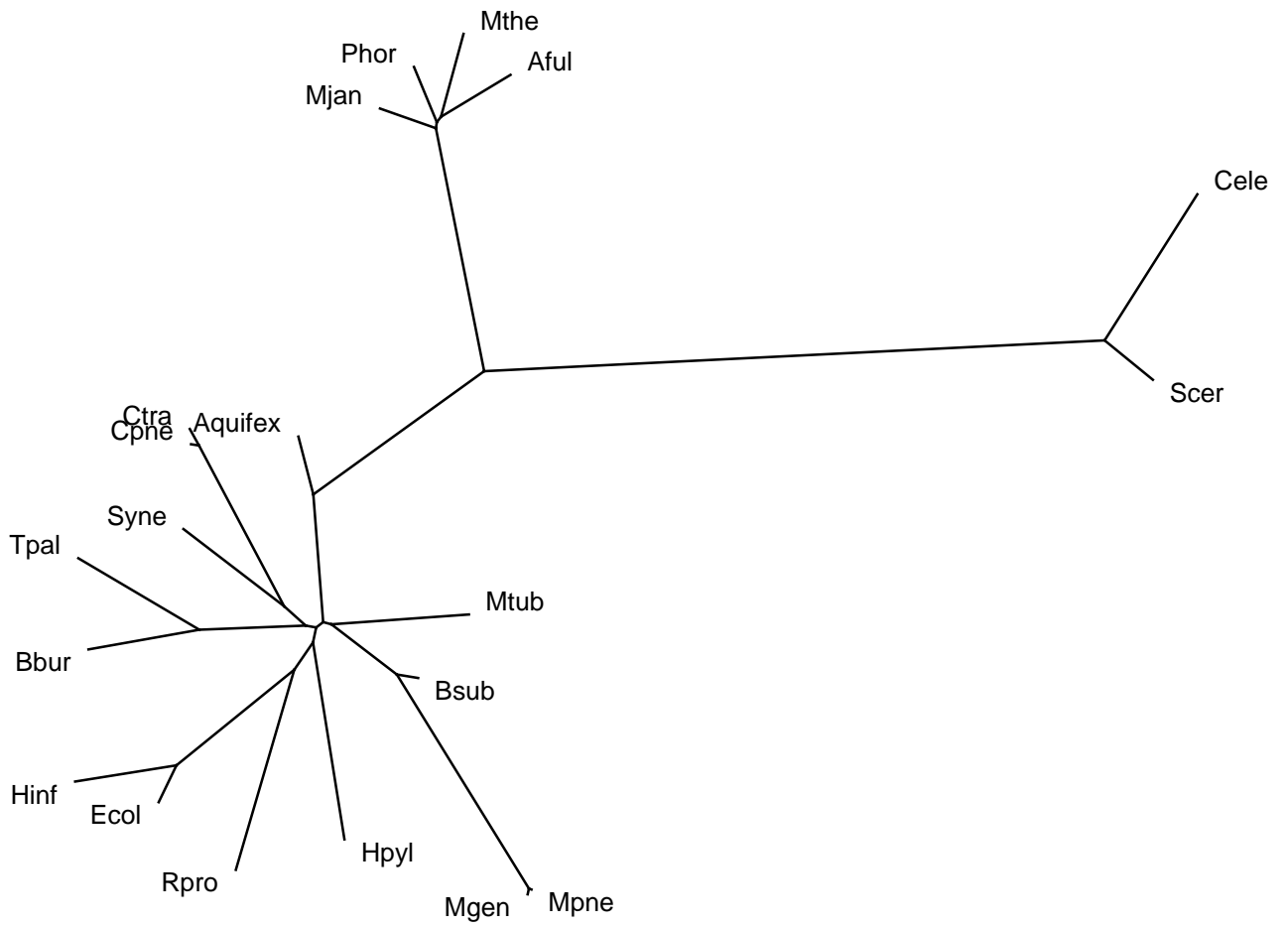


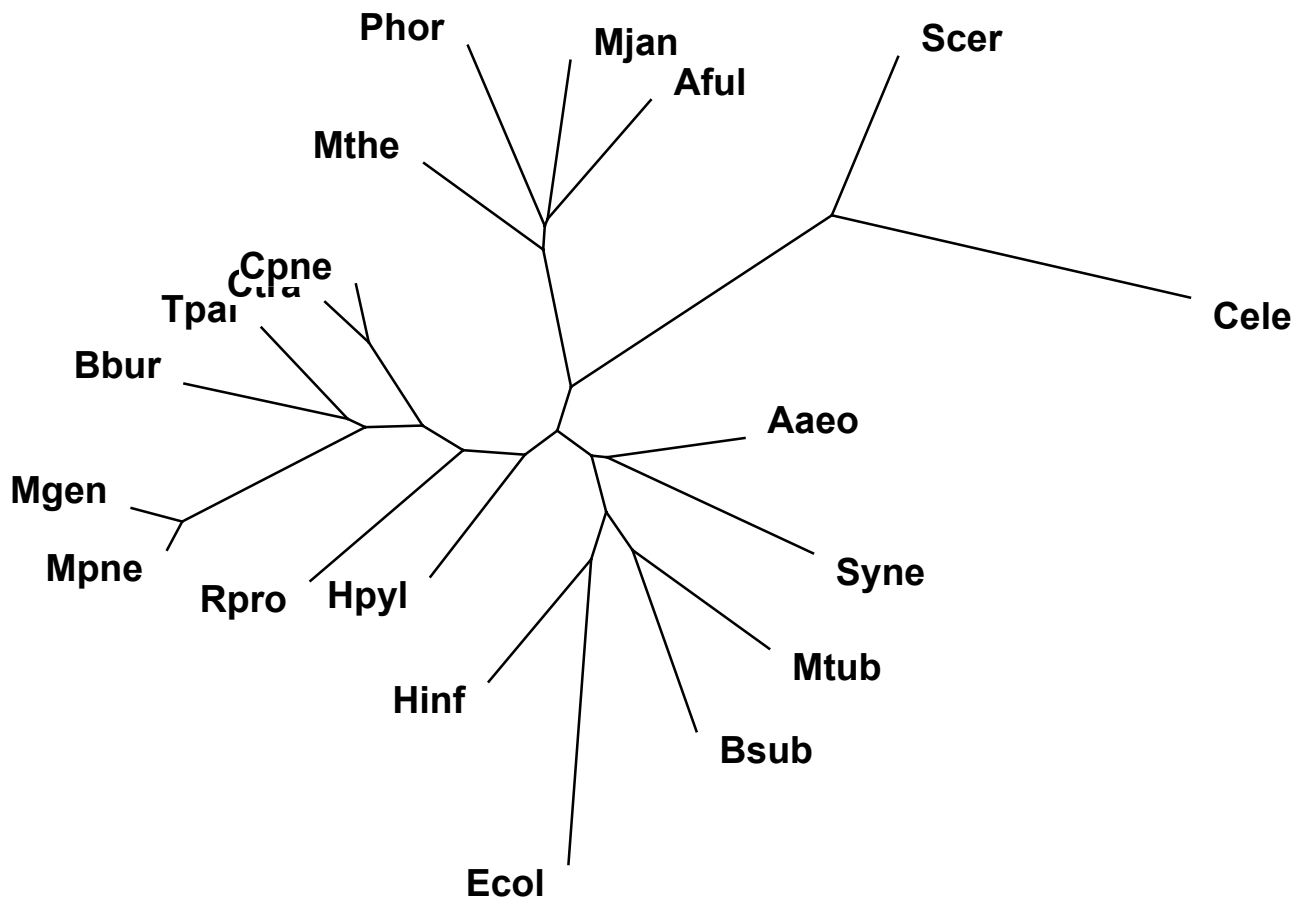
Figure IA



	FOLDS																				
	cele	scer	mjan	phor	mithe	aful	aaso	mtub	bsub	mpne	mgen	hpyl	rpro	ecol	hinf	bbur	tpal	syne	ctra	cpne	
P-loop cont. NTP hydrolase	5	8	21	18	17	14	20	7	9	24	33	16	23	9	17	26	21	10	20	18	3.29
Ferredoxin-like	4	7	29	6	35	24	11	6	.	.	.	7	5	10	10	2	5	6	.	.	4.34
beta/alpha (TIM)-barrel	.	5	13	9	14	12	10	10	12	6	7	7	5	12	11	8	7	8	10	9	3.1
Rossmann-fold	.	4	5	6	8	8	12	16	14	4	4	7	5	10	8	.	.	9	5	5	3.22
SAM-dep. met. transferases	.	.	12	10	5	7	8	7	4	4	.	10	5	4	6	4	.	5	5	4	3.53
Flavodoxin-like	.	.	.	4	8	8	.	.	8	.	.	5	.	8	6	3	.	11	.	.	3.14
alpha-alpha superhelix	5	7	7	.	.	.	9	5	6	.	.	1.91
FAD/NAD(P)-bndng domain	.	.	.	5	4	9	9	5	4	6	7	.	5	5	3.4
Adenine alpha hydrolase	.	.	7	6	5	.	5	.	6	9	5	6	.	.	4	3	.	4	4	.	3.17
PLP-dependent transferases	.	.	5	5	5	.	7	.	6	.	.	3	.	4	5	.	.	3	4	5	3.54
Protein kinases (PK)	10	8	5.1
Immunoglobulin-like	17	2.1
Ribonuclease H-like motif	.	5	4	5	.	7	6	8	5	6	3.47
Cl. II aaRS and biotin syn.	8	10	4	5	.	4	6	.	6	6	.	.	4.61
Acyl-CoA binding protein	.	7	11	13	3	3	1.105
alpha/beta-Hydrolases	9	5	.	4	.	5	4	.	.	.	3.56
Zincin-like	12	7.3
7-bladed beta-propeller	.	8	2.51
OB-fold	6	8	3	.	4	4	.	.	2.29
beta-Grasp	7	4.105
Glucocorticoid rcptr DNA-bnd	6	7.33

	SUPERFAMILIES																				
	cele	scer	mjan	phor	mithe	aful	aaso	mtub	bsub	mpne	mgen	hpyl	rpro	ecol	hinf	bbur	tpal	syne	ctra	cpne	
3.29.1	5	8	21	18	17	14	20	7	9	24	33	16	23	9	17	26	21	10	20	18	d1gky_ P-loop containing NTP hydrolases
4.34.1	.	.	25	4	31	20	6	3	.	8	6	d1fxd_ 4Fe-4S ferredoxins
3.22.1	.	4	5	6	8	8	12	16	14	4	4	7	5	10	8	2	2	9	5	5	d1xel_ NAD(P)-binding Rossmann-fold
3.53.1	.	.	12	10	5	7	8	7	4	4	.	10	5	4	6	4	2	5	5	4	d1vid_ SAM-dependent methyltransferases
3.4.1	.	.	5	4	9	9	5	4	6	7	.	5	5	.	.	3	3	3	.	.	d1grh_ FAD/NAD(P)-binding domain
3.54.1	.	5	5	5	5	5	7	.	6	.	.	3	.	4	5	.	.	3	4	5	d1map_ PLP-dependent transferases
5.1.1	10	8	d1hcl_ Protein kinases (PK), catalytic core
4.61.1	5	.	.	8	10	4	5	.	5	4	6	.	6	6	ds051_ Class II aaRS and biotin synthetases
3.1.5	.	.	5	3	7	.	.	.	4	2	2	.	.	.	d1ads_ NAD(P)-linked oxidoreductase
3.56.1	4	9	5	3	4	.	5	4	3	.	d1ax9_ alpha/beta-Hydrolases
1.105.4	.	6	11	12	3	3	d2maa_ Tropomyosin
3.47.1	.	4	3	4	.	4	3	5	4	5	.	.	.	d1ap8_ Translation initiation factor eIF4e
3.17.2	.	.	6	4	4	4	ds035_ adenine nucleotide alpha hydrolases
2.51.3	.	8	ds029_ Trp-Asp repeat (WD-repeat)
4.89.1	.	4	d1gsa_2 Glutathione synthetase ATP-binding
7.3.9	.	11	d1apo_ EGF/Laminin
3.82.1	.	.	4	4	5	4	3	4	4	4	d1rxm_ Periplasmic binding protein-like II
2.1.1	9	d1cd8_ Immunoglobulin
1.91.8	.	7	.	.	.	9	5	.	.	d1a17_ Tetratricopeptide repeat
2.29.4	5	6	2	4	.	.	.	3	.	4	3	ds025_ Nucleic acid-binding proteins
5.19.1	7	6	5	d1lci_ I-iretly luciferase-like
3.14.2	.	.	.	4	.	.	5	3	.	5	.	3	.	9	.	.	d2che_ CheY-like
3.83.1	8	.	6	4	d1afw1_ Thiolase
4.34.7	.	5	d1fht_ RNA-binding domain
1.91.3	4	5	.	.	d1awcb_ Ankyrin repeat
4.105.1	7	d1lit_ C-type lectin-like
7.33.1	6	d1gdc_ Glucocorticoid receptor DNA-binding





0.1

Figure 3 -- Conservation of the superfamilies in the 6 structural classes

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	cele	seer	mjan	phor	mthe	aful	aaeo	mtub	bsub	mpne	mgen	hpyl	rpro	ecol	hinf	bbur	tpal	syne	ctra	cpne
A	19	16	8	6	3	6	3	3	1	3	3	3	3	4	0	5	2	3	1	5
B	22	13	8	2	4	3	1	7	1	5	2	1	2	1	3	1	3	0	0	4
A/B	4	6	7	4	2	4	3	2	1	5	5	15	5	8	7	6	4	5	8	16
A+B	14	14	10	6	11	3	0	1	7	6	2	7	2	7	4	6	5	2	3	10
M	1	2	0	1	1	0	1	0	0	2	0	0	1	1	0	1	4	1	0	3
S	15	10	1	3	2	1	0	1	0	0	2	0	0	0	0	0	0	0	0	0

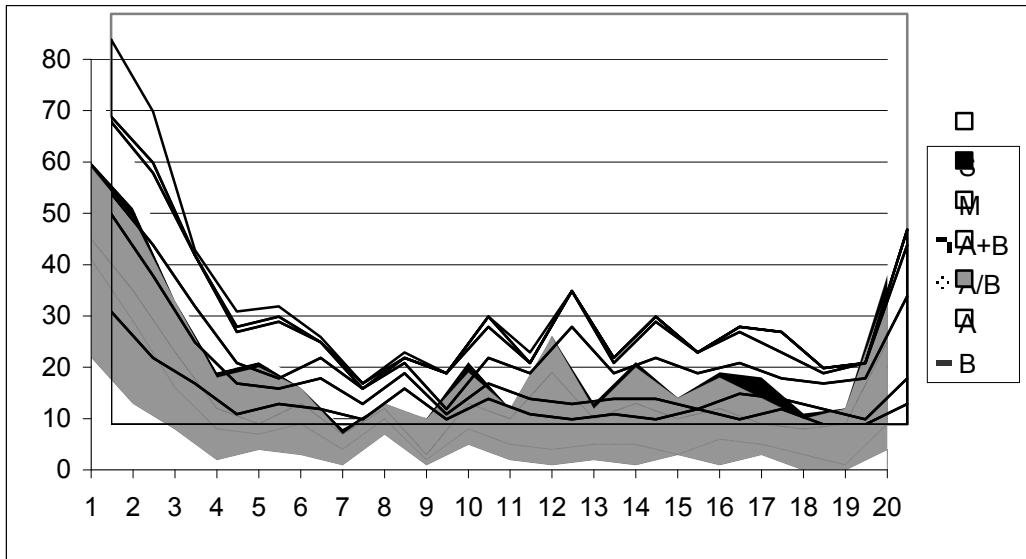
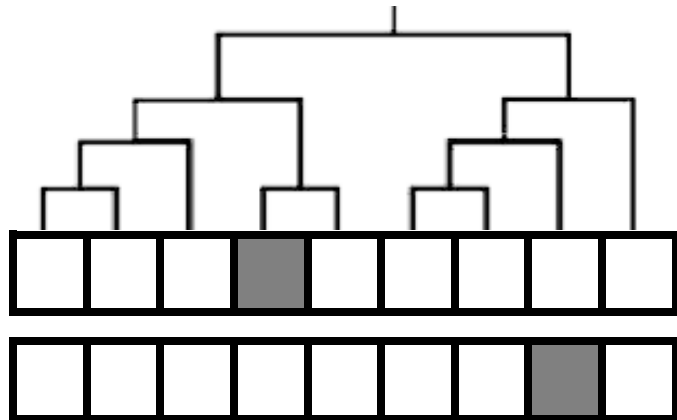
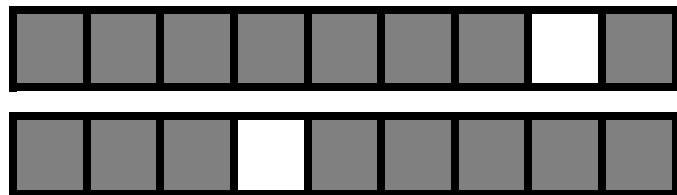


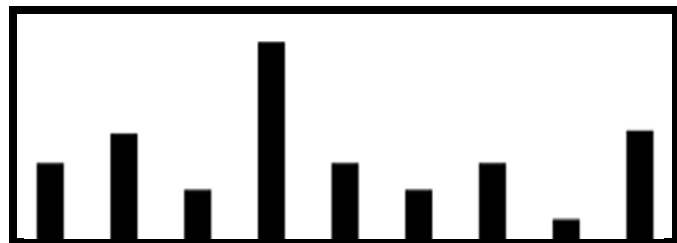
Figure 4 - Schematic of the different fold patterns



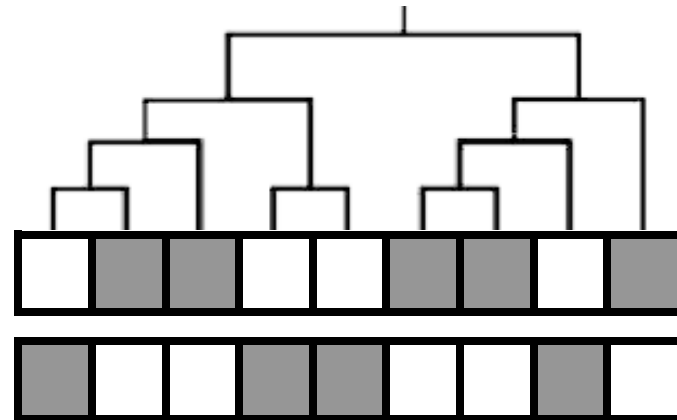
PRESENT IN A SINGLE ORGANISM



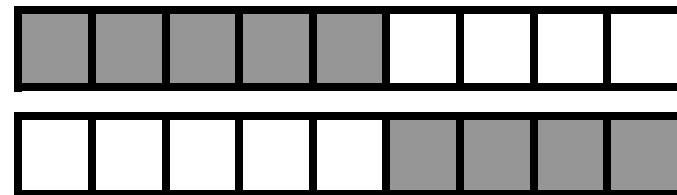
ABSENT FROM A SINGLE ORGANISM



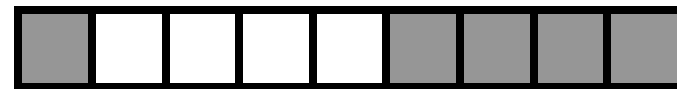
ABUNDANCE OF FOLDS OR ORGANISMS



COMPLEMENTARY PATTERN



SINGLE CLADE (COMPLEMENTS)



LATERAL GENE TRANSFER



GENE LOSS

Table I / A

abbrev.	Species Name	ORF Coverage			Amino Acid Coverage			Domain Matches			Domain Length	Duplication	
		Total	Matching	m/t	Total	Matching	m/t	Folds	Sfam	Dom		Fold	Sfam
Aao	<i>Aquifex aeolicus</i>	1522	527	34.6%	482512	116664	24.2%	162	205	690	169.1	4.26	3.37
Aful	<i>Archaeoglobus fulgidus</i>	2409	650	27.0%	663320	146655	22.1%	147	186	849	172.7	5.78	4.56
Bbur	<i>Borrelia burgdorferi</i>	1638	289	17.6%	432219	65816	15.2%	126	151	369	178.4	2.93	2.44
Bsub	<i>Bacillus subtilis</i>	4100	1121	27.3%	1217000	276596	22.7%	208	276	1460	189.4	7.02	5.29
Cele	<i>Caenorhabditis elegans</i>	19099	4586	24.0%	8096713	1136801	14.0%	247	304	7803	145.7	31.59	25.67
Cpne	<i>Chlamydia pneumoniae</i>	1052	274	26.0%	361694	66160	18.3%	136	165	367	180.3	2.70	2.22
Ctra	<i>Chlamydia trachomatis</i>	894	259	29.0%	312553	60295	19.3%	134	163	348	173.3	2.60	2.13
Ecol	<i>Echerischia coli</i>	4290	1191	27.8%	1363501	296762	21.8%	229	303	1611	184.2	7.03	5.32
Hinf	<i>Haemophilus influenzae Rd</i>	1707	528	30.9%	520930	125776	24.1%	190	243	710	177.1	3.74	2.92
Hpyl	<i>Helicobacter pylori</i>	1577	381	24.2%	500616	89025	17.8%	152	193	495	179.8	3.26	2.56
Mthe	<i>Methanobacterium thermoautotrophicum</i>	479	164	34.2%	174566	39680	22.7%	95	111	228	174.0	2.40	2.05
Mjan	<i>Methanococcus jannaschii</i>	1771	470	26.5%	501793	93299	18.6%	128	164	613	152.2	4.79	3.74
Mtub	<i>Mycobacterium tuberculosis</i>	677	178	26.3%	237651	43222	18.2%	101	118	251	172.2	2.49	2.13
Mgen	<i>Mycoplasma genitalium</i>	1871	522	27.9%	526205	105553	20.1%	135	179	675	156.4	5.00	3.77
Mpne	<i>Mycoplasma pneumoniae</i>	3924	1198	30.5%	1335687	291496	21.8%	199	253	1587	183.7	7.97	6.27
Phor	<i>Pyrococcus horikoshii</i>	2064	461	22.3%	568544	97276	17.1%	121	155	555	175.3	4.59	3.58
Rpro	<i>Rickettsia prowazekii</i>	837	264	31.5%	280233	60285	21.5%	135	160	350	172.2	2.59	2.19
Scer	<i>Saccharomyces cerevisiae</i>	6218	1699	27.3%	2906890	434481	14.9%	215	273	2346	185.2	10.91	8.59
Syne	<i>Synechocystis sp.</i>	3168	882	27.8%	1119717	196041	17.5%	199	255	1131	173.3	5.68	4.44
Tpal	<i>Treponema pallidum</i>	1031	252	24.4%	350676	58542	16.7%	123	150	346	169.2	2.81	2.31

Table I / B: Represented Superfamilies and Their Average Duplication

Levels in the Soluble Fold Classes in *A.fulgidus*, *E.coli*, Yeast, Worm and the Total of the 20 Genomes

	All-alpha			All-beta			Alpha/Beta			Alpha+Beta			Multidomain			Small			Total
	Sfams	Dup	Copy#	Sfams	Dup	Copy#	Sfams	Dup	Copy#	Sfams	Dup	Copy#	Sfams	Dup	Copy#	Sfams	Dup	Copy#	
aful	29	2.5	73	18	2.1	37	74	6.1	453	49	4.2	207	12	5.7	68	4	2.8	11	186
ecol	55	2.9	159	44	4.1	181	105	8.3	872	78	3.9	313	16	5.0	81	5	1.0	5	303
scer	56	7.9	448	35	9.5	333	88	9.3	823	72	5.2	351	14	14.6	204	13	14.3	187	273
cele	62	20.3	1319	52	27.8	1633	81	18.1	1482	72	15.7	1140	14	42.6	598	23	63.7	1631	304
20	97		3197	83		3069	117		8976	120		4046	19		1598	35		1898	471

Table II: Patterns of Complementary Superfamilies and Horizontal Transfer

- A/** Complementary clades between bacterial/archaeal and eukaryotic genes
- B/** Complementary clades between bacterial and eukaryotic/archaeal genomes
- C/** Other complementary patterns
- D/** Horizontal Transfer between Archaea and Bacteria
- E/** Horizontal Transfer between Eukaryotes and Bacteria

	afu	mjan	mthe	pfor	scer	cele	aaeo	syne	ecol	bsub	mtub	hirf	hpyl	ingen	mpne	bbur	tpal	ctra	cpne	lpro	Sfam	domain	SCOP Function	Swissprot	Swissprot Function		
A	3.25.1	d1fsz_1	Tubulin, GTPase domain	FTSZ_ECOLI	CELL DIVISION PROTEIN FTSZ	
	4.57.1	d1puc	Cell cycle regulatory proteins	CKS1_YEAST	CELL DIVISION CONTROL PROTEIN CKS1	
B	1.22.1	d1tafb	Histone-fold	T2D5_YEAST	TRANSCRIPTION INITIATION FACTOR TFIID	
	1.63.1	d1kxu_2	Cyclin-like	TF2B_RAT	TRANSCRIPTION INITIATION FACTOR IIB	
	7.35.3	d1qyp	Rubredoxin-like transcriptional factor domain	TFS2_YEAST	TRANSCRIPTION ELONGATION FACTOR S-II	
	1.34.1	d1coo	C' domain of RNA polymerase alpha subunit	RPOA_TREPA	DNA-DIRECTED RNA POLYMERASE	
	1.75.1	d1gln_1	Glu-tRNA synthetase AC-binding domain	SYE_BACSU	GLUTAMYL-TRNA SYNTHETASE (EC 6.1.1.17)	
	1.88.1	d1sig	RNA polymerase, sigma70 subunit	RPOS_ECOLI	RNA POLYMERASE SIGMA FACTOR
	4.104.1	d2def	Peptide deformylase catalytic core	DEF_HAEIN	POLYPEPTIDE DEFORMYLASE (EC 3.5.1.31)
	4.11.7	d1tif	Translation initiation factor, N' domain	IF3_BORBU	TRANSLATION INITIATION FACTOR IF-3
	4.28.1	d2reb_2	RecA protein, C-terminal domain	RECA_HAEIN	RECA PROTEIN (RECOMBINASE)
	4.36.1	d1ife	Translation initiation factor IF3	IF3_BORBU	TRANSLATION INITIATION FACTOR IF-3
.	4.88.1	d1div	Ribosomal protein L9	RL9_BACSU	50S RIBOSOMAL PROTEIN L9	
C	4.40.1	d2chsa	Chorismate mutase	CHMU_BACSU	CHORISMATE MUTASE (EC 5.4.99.5)	
	1.101.1	d5csma	Chorismate mutase II	CHMU_YEAST	CHORISMATE MUTASE (EC 5.4.99.5)	
	1.81.1	d1cem	Glycosyltransferases of the superhelical fold	GUN6_DICDI	ENDOGLUCANASE (EC 3.2.1.4)	
	2.21.1	d1yna	ConA-like lectins/glucanases	GUN1_TRIRE	ENDOGLUCANASE EG-1 (EC 3.2.1.4)	
	3.1.1	d1edt	Glycosyltransferases	GUNB_NEOPA	ENDOGLUCANASE B (EC 3.2.1.4)	
.	4.2.1	d153l	Lysozyme-like	CHIT_SOLTU	ENDOCHITINASE PRECURSOR (EC 3.2.1.14)	
.	2.65.2	d1hcz_2	Rudiment single hybrid motif	PYC_PICPA	PYRUVATE CARBOXYLASE (EC 6.4.1.1)	
.	2.65.3	d1f3z	Duplicated hybrid motif	PTGA_BACSU	PTS SYSTEM, GLUCOSE-SPECIFIC IIABC COMP	
D	1.86.1	d1aora1	Aldehyde FerOR C' domain	AOR_PYRFU	ALDEHYDE:FERREDOXIN OXIDOREDUCTASE	
	4.94.1	d1aora2	Aldehyde FerOR N' domain	AOR_PYRFU	ALDEHYDE:FERREDOXIN OXIDOREDUCTASE	
	3.1.10	d5ruba1	RuBisCo, C' domain	RBL_NITVU	RUBISCO LARGE SUBUNIT	
E	1.101.1	d5csma	Chorismate mutase II	CHMU_ARATH	CHORISMATE MUTASE (EC 5.4.99.5)	
	1.37.1	d1rec	EF-hand	TPC2_DROME	TROPONIN C	
	2.1.5	d1suh	Cadherin	CAD5_HUMAN	VASCULAR ENDOTHELIAL-CADHERIN	
	2.45.1	d1eal	Lipocalins	PGHD_HUMAN	PROSTAGLANDIN-H2
	3.7.1	d2bnh	Leucine-rich repeats	RINI_PIG	RIBONUCLEASE INHIBITOR
	4.70.1	d1axx	Cytochrome b5	NI1_MAIZE	NITRATE REDUCTASE (EC 1.6.6.1)
.	4.112.1	d1toh	Tyrosine hydroxylase	TY3H_HUMAN	TYROSINE 3-HYDROXYLASE (EC 1.14.16.2)	