

Table S9. Over-represented IPR domains in BSF genome

IPR ID	Max	Hill	Dmel	Zcuc	Gmor	Mdom	Lcup	Annotation
IPR015897	47 (Aaeg)	103	51	45	19	55	28	CHK kinase-like
IPR036508	134 (Aaeg)	182	98	87	46	112	87	Chitin binding domain superfamily
IPR004117	94 (Mdes)	126	59	66	42	79	54	Olfactory receptor, insect
IPR004875	17 (Aaeg)	46	0	17	0	2	0	DDE superfamily endonuclease domain
IPR001128	179 (Aaeg)	207	88	99	62	153	64	Cytochrome P450
IPR029033	35 (Mdes)	85	25	22	20	25	20	Histidine phosphatase superfamily
IPR005828	65 (Aaeg)	88	59	56	33	51	34	Major facilitator, sugar transporter-like
IPR010629	19 (Aaeg)	42	7	6	3	6	7	Insect allergen-related
IPR036875	31 (Aaeg)	54	10	12	9	18	13	Zinc finger, CCHC-type superfamily
IPR036058	23 (Mdes)	49	25	25	15	30	26	Kazal domain superfamily
IPR020471	12 (Agam)	29	9	10	10	6	8	Aldo/keto reductase
IPR036812	15 (Agam)	32	13	11	13	8	10	NADP-dependent oxidoreductase domain superfamily
IPR006600	19 (Mdes)	35	5	9	4	13	4	HTH CenpB-type DNA-binding domain
IPR020400	0	25	4	0	1	9	4	Cecropin, diptera
IPR008160	12 (Mdes)	34	9	15	11	19	5	Collagen triple helix repeat
IPR009882	3 (Mdes)	17	0	2	0	1	3	Gypsy
IPR013201	15(Bant)	28	10	8	4	9	10	Cathepsin propeptide inhibitor domain (I29)
IPR021109	60 (Mdes)	73	14	17	12	21	13	Aspartic peptidase domain superfamily
IPR036282	45 (Bant)	62	49	44	26	46	27	Glutathione S-transferase
IPR036640	29 (Aaeg)	41	24	17	17	25	19	ABC transporter type 1

The domain whose copy number in BSF genome (Hill) is higher than any other dipteran species (nine were investigated) for 30% was considered as a domain over-represented in BSF. The IPR domains were sorted based on the difference between the copy of BSF and the maximum value among other nine dipteran species; of these, the top 20 were listed in this table. The following domains were not listed in this list due to either unknown functional information or duplicated by a superfamily in this list: IPR004119, IPR007999, IPR013320, IPR000560, IPR002403, IPR023210, IPR011527, and IPR004046.