TBD

The Pickles

Febcember 32nd 3023

Results

. . . a quantitative comparison of the datasets in Table.

Table 1: Quantitative metrics of the generated and existing functional annotation sets. C, F, P, and A refer to the aspects of the GO: Cellular Component, Biological Function, Molecular Process, and Any/All.

			Annotations[a]				Annotated Genes[b]				Median Ann. per G.[c]			
Genome	Genes	Dataset	C	F	P	**A**	C	F	P	**A**	C	F	P	**A**
Arachis_hypogaea	100	GOMAP	153433	132944	493799	**780176**	57667	56855	67123	**67124**	2	2	6	**10**
Glycine_max		GOMAP	129215	114010	417575	**660800**	46020	47087	52871	**52872**	2	2	6	**11**
Hordeum_vulgarum		GOMAP	88130	80371	272835	**441336**	35237	36487	39733	**39734**	2	2	5	**10**
Medicago_truncatula.A17		GOMAP	107362	99719	364065	**571146**	42325	43736	50443	**50444**	2	2	6	**10**
Medicago_truncatula.R108		GOMAP	112343	108031	382322	**602696**	40332	50220	55706	**55706**	1	2	5	**9**
Oryza_sativa		GOMAP	72782	64783	248713	**386278**	28619	29876	35824	**35825**	2	2	6	**9**
Phaseolus_vulgaris		GOMAP	72005	64583	229630	**366218**	25934	25539	27432	**27433**	2	2	6	**11**
Triticum_aestivum		GOMAP	267742	218839	786028	**1272609**	95604	98224	107890	**107891**	2	2	6	**10**
Vigna_unguiculata		GOMAP	75867	68313	243278	**387458**	27173	27124	29772	**29773**	2	2	6	**11**
Zea_mays.B73.v3		GOMAP	135251	87953	291855	**515059**	34867	38099	39469	**39469**	3	2	6	**11**
Zea_mays.B73.v4		GOMAP	88831	82849	278952	**450632**	36717	37431	39324	**39324**	2	2	6	**10**
Zea_mays.Mo17		GOMAP	87573	79755	278043	**445371**	33618	35177	38620	**38620**	2	2	6	**10**
Zea_mays.PH207		GOMAP	90625	86106	288937	**465668**	35170	36843	40557	**40557**	2	2	6	**10**
Zea_mays.W22		GOMAP	95397	85616	290032	**471045**	36987	37764	40690	**40690**	2	2	6	**10**

[a] How many annotations in the C, F, and P aspect does this dataset contain? A = How many in total? $A = C + F + P$

[b] How many genes in the genome have at least one GO term from the C, F, P aspect annotated to them? A = How many at least one from any aspect? $(A = C \cup F \cup P)$

[c] Take a typical gene that is present in the annotation set. How many annotations does it have in each aspect? A = How many in total? Ask your favorite statistician why $A \neq C + F + P$