

Supplementary Information 10

CDS annotation result by human

The 33,409 representative clones in FANTOM2 were classified based on the gene name and CDS status. By excluding clones in unknown EST, unclassifiable and non-coding RNA, 17,200 clones were considered as protein coding genes. Among the remaining 16,209 TUs, 394 TUs were turned out to be protein coding genes supported by public proteins. In total, 17,594 FANTOM2 TUs were considered to have protein coding potential.

	complete	5'truncated	3'truncated	5',3'truncated	immature	5'UTR	3'UTR	artifact	non-coding	unknown	Total
DNA	4037	448	197	28	431	8	67	9	7	540	5772
Protein	838	13	1	0	2	0	0	0	0	1	855
homolog	1871	242	146	15	312	0	0	0	0	18	2604
similar	740	79	52	5	218	0	0	0	0	20	1114
weakly	560	55	48	5	207	0	0	0	0	30	905
cluster	64	20	1	1	20	1	23	0	0	172	302
domain	1753	129	136	5	136	1	0	0	0	24	2184
hyp_prot	2591	375	117	10	339	0	0	2	0	37	3471
Total	12454	1361	698	69	1665	10	90	11	7	842	17207
Translated	12267	1328	684	66	0	0	0	0	0	0	14345