

***De novo* prediction of localization of expression for the 30% fraction of genes characterized with high prediction scores**

The *de novo* predictions of localization of expression for the 30% fraction of genes characterized with high prediction scores obtained for each functional unit are verified with Gene Ontology terms from the Biological Process category. As a rule a gene is annotated with many GO:Biological Process terms. The *de novo* prediction of localization of expression for a gene is confirmed with GO term (the respective gene is marked with "yes" sign in the tables below), if and only if, the gene is annotated with at least one GO:Biological process term relevant to the developmental process the anatomical structures comprising the functional unit under study are known to be involved in. If the GO:Biological process annotation does not contain evidence in support of prediction of localization of expression a gene is assigned with "Not Annotated" ("NA") symbol in the tables. The GO:Biological Process annotation for *Drosophila* genes can be found in the data support files for this manuscript or in FlyBase database [1] at www.flybase.net.

Table 10: **Results of prediction of localization of gene expression for functional unit pep2ebrain:** embryonic central brain && protocerebrum primordium && procephalic ectoderm primordium

Prediction Rank	Supported with GO:Biological Process annotation	Gene ID	Gene name	Confidence Score (distance from the hyperplane)
1	yes	FBgn0001139	<i>gro</i>	13.8571
2	NA	FBgn0020496	<i>CtBP</i>	12.2987
3	yes	FBgn0014163	<i>fax</i>	12.1705
4	NA	FBgn0003498	<i>sqd</i>	11.8179
5	NA	FBgn0015622	<i>Cnx99A</i>	11.4479
6	yes	FBgn0010333	<i>Rac1</i>	11.4310
7	NA	FBgn0011760	<i>ctp</i>	11.2622
8	NA	FBgn0020910	<i>RpL3</i>	10.5409
9	NA	FBgn0010410	<i>RpL27A</i>	10.1206
10	NA	FBgn0003885	α <i>Tub84D</i>	9.7893
11	NA	FBgn0014857	<i>His3.3A</i>	9.5983
12	yes	FBgn0003053	<i>peb</i>	9.5226
13	yes	FBgn0004108	<i>Nrt</i>	9.5063
14	NA	FBgn0001105	<i>Gβ13F</i>	9.4496
15	yes	FBgn0014184	<i>Oda</i>	9.3732
16	yes	FBgn0004907	<i>14-3-3ζ</i>	9.3489
17	NA	FBgn0004867	<i>sop</i>	9.2384
18	NA	FBgn0024939	<i>RpL8</i>	9.0460
19	NA	FBgn0000097	<i>aop</i>	8.6087
20	NA	FBgn0020615	<i>SelD</i>	8.5423
21	yes	FBgn0003995	<i>vvl</i>	8.4689
22	yes	FBgn0024250	<i>brk</i>	8.4506
23	NA	FBgn0015805	<i>Rpd3</i>	8.3159
24	NA	FBgn0028519	<i>CG4500</i>	8.1256
continued on next page...				

Table 10 continued from previous page.

Prediction Rank	Supported with GO:Biological Process annotation	Gene ID	Gene name	Confidence Score (distance from the hyperplane)
25	yes	FBgn0000308	<i>chic</i>	7.9656
26	yes	FBgn0001235	<i>hth</i>	7.8935
27	NA	FBgn0004364	<i>18w</i>	7.6850
28	NA	FBgn0004177	<i>mts</i>	7.5565
29	NA	FBgn0011817	<i>nmo</i>	7.5545
30	yes	FBgn0000413	<i>da</i>	7.5429
31	NA	FBgn0003471	<i>β-Spec</i>	7.5041
32	yes	FBgn0004606	<i>zfh1</i>	7.3404
33	yes	FBgn0002777	<i>mnb</i>	7.3242
34	NA	FBgn0004237	<i>Hrb87F</i>	7.2800
35	NA	FBgn0003941	<i>RpL40</i>	7.2751
36	NA	FBgn0003377	<i>Sgs7</i>	7.2304
37	NA	FBgn0003942	<i>RpS27A</i>	7.2268
38	yes	FBgn0003326	<i>sca</i>	7.1966
39	NA	FBgn0000405	<i>CycB</i>	7.0991
40	NA	FBgn0003501	<i>Src64B</i>	7.0519
41	yes	FBgn0003310	<i>S</i>	6.9174
42	NA	FBgn0003002	<i>opa</i>	6.9040
43	NA	FBgn0020369	<i>Pros45</i>	6.8997
44	NA	FBgn0003866	<i>tsh</i>	6.8960
45	NA	FBgn0003177	<i>pyd</i>	6.6653
46	NA	FBgn0004133	<i>blow</i>	6.6602
47	yes	FBgn0002945	<i>nkd</i>	6.6601
48	NA	FBgn0003887	<i>βTub56D</i>	6.5789
49	NA	FBgn0004362	<i>HmgD</i>	6.5614
50	NA	FBgn0004427	<i>LysD</i>	6.5324
51	NA	FBgn0015763	<i>lic</i>	6.5020
52	NA	FBgn0015239	<i>Hr78</i>	6.4408
53	NA	FBgn0013343	<i>Syx1A</i>	6.4082
54	NA	FBgn0023441	<i>fus</i>	6.3593
55	NA	FBgn0003963	<i>ush</i>	6.3319
56	NA	FBgn0000042	<i>Act5C</i>	6.2664
57	NA	FBgn0004003	<i>wbl</i>	6.1980
58	yes	FBgn0003892	<i>ptc</i>	6.1305
59	yes	FBgn0001077	<i>ftz</i>	6.0059
60	NA	FBgn0014011	<i>Rac2</i>	5.9803
61	NA	FBgn0000150	<i>awd</i>	5.9343
62	yes	FBgn0004053	<i>zen</i>	5.9191
63	NA	FBgn0014870	<i>Psi</i>	5.8882
64	NA	FBgn0003118	<i>pnt</i>	5.8771
65	yes	FBgn0026181	<i>rok</i>	5.7088
66	yes	FBgn0003651	<i>svp</i>	5.6886
67	NA	FBgn0005278	<i>M(2)21AB</i>	5.6299
68	NA	FBgn0027343	<i>fz3</i>	5.5626
69	NA	FBgn0003210	<i>rb</i>	5.5013
70	yes	FBgn0052372	<i>CG32372</i>	5.4465

continued on next page...

Table 10 continued from previous page.				
Prediction Rank	Supported with GO:Biological Process annotation	Gene ID	Gene name	Confidence Score (distance from the hyperplane)
71	yes	FBgn0004179	<i>Csp</i>	5.4383
72	yes	FBgn0000273	<i>Pka-C1</i>	5.4012
73	NA	FBgn0000046	<i>Act87E</i>	5.3949
74	NA	FBgn0041186	<i>Slbp</i>	5.3550
75	NA	FBgn0015625	<i>CycB3</i>	5.3462
76	NA	FBgn0004509	<i>Fur1</i>	5.3093
77	yes	FBgn0001104	<i>G-α65A</i>	5.3028
78	yes	FBgn0010105	<i>comm</i>	5.2662
79	yes	FBgn0000340	<i>cno</i>	5.2191
80	yes	FBgn0005592	<i>btl</i>	5.1844
81	yes	FBgn0002973	<i>numb</i>	5.1702
82	yes	FBgn0024836	<i>stan</i>	5.1629
83	yes	FBgn0014343	<i>mirr</i>	5.1539
84	NA	FBgn0004425	<i>LysB</i>	5.1538
85	yes	FBgn0015380	<i>drl</i>	5.1158
86	NA	FBgn0000043	<i>Act42A</i>	5.1015
87	NA	FBgn0024234	<i>gbb</i>	5.0911
88	yes	FBgn0000499	<i>dsh</i>	5.0850
89	NA	FBgn0004828	<i>His3.3B</i>	5.0326
90	NA	FBgn0002989	<i>okr</i>	5.0274
91	NA	FBgn0016696	<i>Pitslre</i>	5.0183
92	NA	FBgn0012037	<i>Ance</i>	4.9827
93	yes	FBgn0004647	<i>N</i>	4.8935
94	NA	FBgn0001565	<i>Hlc</i>	4.8580
95	NA	FBgn0004399	<i>psq</i>	4.8255
96	NA	FBgn0003512	<i>Sry-δ</i>	4.8240
97	yes	FBgn0003430	<i>slp1</i>	4.8109
98	yes	FBgn0003986	<i>vnd</i>	4.7892
99	NA	FBgn0014028	<i>SdhB</i>	4.7829
100	NA	FBgn0038224	<i>CG3321</i>	4.7388
101	NA	FBgn0011672	<i>Mvl</i>	4.7353
102	yes	FBgn0005672	<i>spi</i>	4.7237
103	NA	FBgn0004456	<i>mew</i>	4.7229
104	yes	FBgn0000459	<i>disco</i>	4.7048
105	NA	FBgn0001404	<i>egh</i>	4.7037
106	yes	FBgn0000570	<i>elav</i>	4.6986
107	NA	FBgn0024923	<i>TER94</i>	4.6936
108	yes	FBgn0003749	<i>trh</i>	4.6703
109	NA	FBgn0015218	<i>eIF-4E</i>	4.6636
110	NA	FBgn0014388	<i>sty</i>	4.5951
111	NA	FBgn0025140	<i>pit</i>	4.5889
112	yes	FBgn0015014	<i>tgo</i>	4.5820
113	yes	FBgn0003391	<i>shg</i>	4.5685
114	NA	FBgn0001257	<i>ImpL2</i>	4.5494
115	NA	FBgn0020378	<i>Sp1</i>	4.5112
116	NA	FBgn0004551	<i>Ca-P60A</i>	4.5034
continued on next page...				

Table 10 continued from previous page.

Prediction Rank	Supported with GO:Biological Process annotation	Gene ID	Gene name	Confidence Score (distance from the hyperplane)
117	NA	FBgn0000615	<i>exu</i>	4.4831
118	NA	FBgn0032408	<i>CG6712</i>	4.4170
119	NA	FBgn0004449	<i>Ten-m</i>	4.4159
120	yes	FBgn0004009	<i>wg</i>	4.3906
121	NA	FBgn0010278	<i>Ssrp</i>	4.3492
122	NA	FBgn0002413	<i>dco</i>	4.3335
123	NA	FBgn0000568	<i>Eip75B</i>	4.2971
124	yes	FBgn0003165	<i>pum</i>	4.2872
125	NA	FBgn0013811	<i>Dhc62B</i>	4.2754
126	yes	FBgn0013765	<i>cnn</i>	4.2496
127	NA	FBgn0013974	<i>Gyc76C</i>	4.2472
128	yes	FBgn0020391	<i>Nrk</i>	4.2173
129	NA	FBgn0005624	<i>Psc</i>	4.2062
130	NA	FBgn0011715	<i>Snr1</i>	4.2045
131	yes	FBgn0003719	<i>tld</i>	4.1924
132	yes	FBgn0010575	<i>sbb</i>	4.1730
133	yes	FBgn0013272	<i>Gp150</i>	4.1729
134	NA	FBgn0003068	<i>per</i>	4.1613
135	NA	FBgn0004426	<i>LysC</i>	4.1175
136	NA	FBgn0024811	<i>Crk</i>	4.0998
137	NA	FBgn0001090	<i>bnb</i>	4.0960
138	yes	FBgn0003660	<i>Syb</i>	4.0479
139	yes	FBgn0010341	<i>Cdc42</i>	4.0164
140	NA	FBgn0015617	<i>Cdk7</i>	4.0100
141	NA	FBgn0004897	<i>fd96Ca</i>	3.9875
142	NA	FBgn0020278	<i>loco</i>	3.9615
143	yes	FBgn0001319	<i>kn</i>	3.9563
144	yes	FBgn0003285	<i>rst</i>	3.9456
145	NA	FBgn0000566	<i>Eip55E</i>	3.9351
146	NA	FBgn0017418	<i>ari-1</i>	3.9311

References

- [1] Drysdale RA, Crosby MA (2005) FlyBase: genes and gene models. *Nucleic Acids Res* 33:D390–5.