

***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 8: **Results of prediction of localization of gene expression for functional unit mat2pep: procephalic ectoderm primordium && cellular blastoderm && maternal**

Prediction Rank	Supported with GO:Biological Process annotation	Gene ID	Gene name	Confidence Score (distance from the hyperplane)
1	yes	FBgn0011761	<i>dhd</i>	53.9952
2	NA	FBgn0015565	<i>yin</i>	52.5244
3	yes	FBgn0003015	<i>osk</i>	51.6832
4	yes	FBgn0000615	<i>exu</i>	44.1211
5	yes	FBgn0004649	<i>yl</i>	32.2157
6	yes	FBgn0013272	<i>Gp150</i>	30.9283
7	NA	FBgn0010288	<i>Uch</i>	30.2230
8	yes	FBgn0017561	<i>Ork1</i>	28.2992
9	yes	FBgn0002962	<i>nos</i>	27.6248
10	yes	FBgn0026320	<i>Tom</i>	26.5294
11	NA	FBgn0004873	<i>hts</i>	23.6667
12	NA	FBgn0015776	<i>nrv1</i>	22.4076
13	yes	FBgn0000166	<i>bcd</i>	22.1905
14	yes	FBgn0001090	<i>bnb</i>	21.7687
15	yes	FBgn0000216	<i>Brd</i>	21.7446
16	yes	FBgn0004364	<i>18w</i>	20.1106
17	NA	FBgn0000182	<i>BicC</i>	19.1090
18	NA	FBgn0005671	<i>Vha55</i>	18.9957
19	yes	FBgn0003970	<i>vas</i>	18.4916
20	yes	FBgn0021776	<i>mira</i>	18.1972
21	yes	FBgn0023094	<i>cyc</i>	16.2562
22	yes	FBgn0000180	<i>bib</i>	16.0548
23	NA	FBgn0011705	<i>rost</i>	15.8194
24	NA	FBgn0028519	<i>CG4500</i>	15.4161
continued on next page...				

Table 8 continued from previous page.				
Prediction Rank	Supported with GO:Biological Process annotation	Gene ID	Gene name	Confidence Score (distance from the hyperplane)
25	NA	FBgn0005771	<i>noc</i>	15.3789
26	yes	FBgn0004236	α <i>Tub67C</i>	15.3351
27	yes	FBgn0002633	<i>HLHm7</i>	15.2871
28	NA	FBgn0000392	<i>cup</i>	15.2272
29	NA	FBgn0029174	<i>FKBP59</i>	14.8504
30	NA	FBgn0003187	<i>qua</i>	14.6951
31	yes	FBgn0010452	<i>trn</i>	14.2813
32	NA	FBgn0012037	<i>Ance</i>	14.2331
33	yes	FBgn0005695	<i>gcl</i>	14.0786
34	NA	FBgn0000183	<i>BicD</i>	13.7730
35	yes	FBgn0001235	<i>hth</i>	13.6069
36	NA	FBgn0000562	<i>egl</i>	13.5788
37	NA	FBgn0000097	<i>aop</i>	13.2934
38	NA	FBgn0004914	<i>Hnf4</i>	13.2717
39	yes	FBgn0000251	<i>cad</i>	13.1426
40	NA	FBgn0003204	<i>ras</i>	13.1191
41	yes	FBgn0000575	<i>emc</i>	12.9465
42	NA	FBgn0002673	<i>twe</i>	12.8778
43	NA	FBgn0019686	<i>lok</i>	12.7563
44	NA	FBgn0024846	<i>p38b</i>	12.5961
45	yes	FBgn0002629	<i>m4</i>	12.4736
46	yes	FBgn0000577	<i>en</i>	11.5534
47	NA	FBgn0005672	<i>spi</i>	11.5020
48	NA	FBgn0003023	<i>otu</i>	11.4709
49	yes	FBgn0014163	<i>fax</i>	11.3598
50	yes	FBgn0002631	<i>HLHm5</i>	11.2983
51	NA	FBgn0015239	<i>Hr78</i>	10.9673
52	NA	FBgn0002528	<i>LanB2</i>	10.6957
53	yes	FBgn0001254	<i>ImpE2</i>	10.5246
54	yes	FBgn0002561	<i>l(1)sc</i>	10.5130
55	yes	FBgn0003733	<i>tor</i>	10.2383
56	NA	FBgn0021765	<i>scu</i>	10.1786
57	NA	FBgn0020497	<i>emb</i>	9.7753
58	NA	FBgn0004872	<i>piwi</i>	9.5132
59	yes	FBgn0003117	<i>pnr</i>	9.4732
60	NA	FBgn0010316	<i>dap</i>	9.4639
61	yes	FBgn0014133	<i>bif</i>	9.4065
62	NA	FBgn0024234	<i>gbb</i>	9.3837
63	yes	FBgn0001987	<i>Gli</i>	9.3539
64	NA	FBgn0002069	<i>Aats-asp</i>	9.3367
65	yes	FBgn0011592	<i>fra</i>	9.2284
66	yes	FBgn0000463	<i>Dl</i>	8.8257
67	NA	FBgn0032693	<i>Cyp310a1</i>	8.7707
68	yes	FBgn0000011	<i>ab</i>	8.7675
69	NA	FBgn0004101	<i>bs</i>	8.6142
70	yes	FBgn0010105	<i>comm</i>	8.5711
continued on next page...				

Table 8 continued from previous page.				
Prediction Rank	Supported with GO:Biological Process annotation	Gene ID	Gene name	Confidence Score (distance from the hyperplane)
71	yes	FBgn0000411	<i>D</i>	8.5129
72	yes	FBgn0000492	<i>Dr</i>	8.4884
73	yes	FBgn0004009	<i>wg</i>	8.4225
74	NA	FBgn0004362	<i>HmgD</i>	8.3760
75	NA	FBgn0002948	<i>nod</i>	8.2316
76	NA	FBgn0003507	<i>srp</i>	8.1902
77	NA	FBgn0011823	<i>Pen</i>	7.9986
78	NA	FBgn0000405	<i>CycB</i>	7.9496
79	yes	FBgn0000591	<i>E(spl)</i>	7.8492
80	NA	FBgn0003483	<i>spn-E</i>	7.6450
81	NA	FBgn0003884	α <i>Tub84B</i>	7.5595
82	NA	FBgn0001308	<i>Khc</i>	7.3953
83	NA	FBgn0001215	<i>Hrb98DE</i>	7.3923
84	yes	FBgn0002526	<i>LanA</i>	7.2772
85	yes	FBgn0002732	<i>mα</i>	7.2556
86	yes	FBgn0001325	<i>Kr</i>	7.2540
87	NA	FBgn0020633	<i>Mcm7</i>	7.1390
88	NA	FBgn0011604	<i>Iswi</i>	7.0895
89	yes	FBgn0003495	<i>spz</i>	7.0775
90	yes	FBgn0001075	<i>ft</i>	6.9668
91	NA	FBgn0015218	<i>eIF-4E</i>	6.9414
92	NA	FBgn0001257	<i>ImpL2</i>	6.8665
93	NA	FBgn0000150	<i>awd</i>	6.8446
94	yes	FBgn0004170	<i>sc</i>	6.7448
95	yes	FBgn0003430	<i>slp1</i>	6.4661
96	NA	FBgn0024923	<i>TER94</i>	6.4253
97	yes	FBgn0015380	<i>drl</i>	6.4185
98	NA	FBgn0020545	<i>kraken</i>	6.4003
99	yes	FBgn0005634	<i>zip</i>	6.3568
100	NA	FBgn0015765	<i>Mpk2</i>	6.3489
101	NA	FBgn0020369	<i>Pros45</i>	6.3325
102	NA	FBgn0003942	<i>RpS27A</i>	6.3138
103	NA	FBgn0015625	<i>CycB3</i>	6.3066
104	yes	FBgn0014179	<i>gmc</i>	6.2684
105	yes	FBgn0015805	<i>Rpd3</i>	6.1754
106	NA	FBgn0002524	<i>lace</i>	6.0995
107	NA	FBgn0003887	β <i>Tub56D</i>	6.0608
108	NA	FBgn0004237	<i>Hrb87F</i>	6.0451
109	yes	FBgn0003525	<i>stg</i>	5.9833
110	NA	FBgn0001941	<i>ifc</i>	5.9770
111	NA	FBgn0003471	β - <i>Spec</i>	5.9538
112	NA	FBgn0000927	<i>fs(1)Ya</i>	5.9130
113	NA	FBgn0000042	<i>Act5C</i>	5.8574
114	NA	FBgn0002283	<i>l(3)73Ah84D</i>	5.7829
115	NA	FBgn0003885	α <i>Tub84D</i>	5.7761
116	yes	FBgn0001148	<i>gsb</i>	5.7600
continued on next page...				

Table 8 continued from previous page.

Prediction Rank	Supported with GO:Biological Process annotation	Gene ID	Gene name	Confidence Score (distance from the hyperplane)
117	yes	FBgn0010349	<i>Dhc64C</i>	5.7546
118	yes	FBgn0011598	<i>grp</i>	5.7446
119	NA	FBgn0004381	<i>Klp68D</i>	5.7117
120	yes	FBgn0001123	<i>G-α60A</i>	5.6853
121	yes	FBgn0011817	<i>nmo</i>	5.6682
122	NA	FBgn0020910	<i>RpL3</i>	5.6398
123	yes	FBgn0003498	<i>sqd</i>	5.5945
124	yes	FBgn0011260	<i>Sema-2a</i>	5.5443
125	NA	FBgn0015295	<i>shark</i>	5.5302
126	NA	FBgn0020622	<i>Pi3K21B</i>	5.5179
127	NA	FBgn0010438	<i>mtSSB</i>	5.4838
128	NA	FBgn0010410	<i>RpL27A</i>	5.4172
129	NA	FBgn0014857	<i>His3.3A</i>	5.3835
130	yes	FBgn0000308	<i>chic</i>	5.3499
131	NA	FBgn0004867	<i>sop</i>	5.3374
132	NA	FBgn0000158	<i>bam</i>	5.3253
133	yes	FBgn0004893	<i>bowl</i>	5.2975
134	NA	FBgn0003124	<i>polo</i>	5.2972
135	NA	FBgn0024939	<i>RpL8</i>	5.2501
136	NA	FBgn0011361	<i>mtacp1</i>	5.2215
137	yes	FBgn0010382	<i>CycE</i>	5.1200

References

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