

categories:

- 1 – no align
- 2 – align
- 2a – no public hits, no genscan prediction
- 2a1 - < 500bp to public, genscan
- 2a2 - <1000bp to public, genscan
- 2a3 - < 5000bp to public, genscan
- 2b1 – genscan with introns or exon-intron-mix
- 2b2 – genscan exons (at least 2 exons)
- 2c1 – public evidence introns or exon-intron-mix
- 2c2 – public evidence exons (at least 2 exons)
- 2c3 – public evidence, single exon ESTs

public data - amount of mRNAs in each category:

- 1 15
- 2 35
- 2a 0
- 2b1 5
- 2b2 14
- 2c1 0
- 2c2 14
- 2c3 2

EST singletons:

- 1 11
- 2 39
- 2a 7
- 2a1 2
- 2a2 2
- 2a3 3
- 2b1 24
- 2b2 1
- 2c1 2
- 2c2 2
- 2c3 3

EST clusters:

- 1 9
- 2 41
- 2a 6
- 2a1 3
- 2a2 2
- 2a3 1
- 2b1 11
- 2b2 7
- 2c1 2
- 2c2 11
- 2c3 4

alignment of 50 publicly available mRNAs

1 - 1

>gi|78771598|dbj|AB240046.1| *Oryzias latipes* OIPGRMC1 mRNA for progesterin receptor membrane component 1, complete cds no A

2 - 2c2

>gi|38045382|gb|AY353247.1| *Oryzias latipes* RhBG-like protein mRNA, complete cds scaff2539, overlap with ESTs

3 - 1

>gi|62176937|emb|AJ635228.1| *Oryzias latipes* mRNA for cytoglobin-2 (cygb-2 gene) no A

4 - 2b2

>gi|77627615|dbj|AB211172.1| *Oryzias latipes* gadd45alpha mRNA for growth arrest and DNA damage-inducible gene 45 alpha, complete cds scaff5538, overlap with genscan prediction

5 - 2c2

>gi|15004312|gb|AF332510.1| *Oryzias latipes* membrane-associated transporter protein B mRNA, complete cds, scaff6012, scaff too short, overlap with EST

6 - 1

>gi|76150371|emb|AM087010.1| *Oryzias latipes* mRNA for autophagy protein 5 (atg5 gene) no A

7 - 2b1

>gi|74267588|dbj|AB208019.1| *Oryzias latipes* gene for hoxD4b, complete cds, scaff10596, overlap with exons and intron of genscan prediction

8 - 2b1

>gi|74267574|dbj|AB208012.1| *Oryzias latipes* gene for hoxC13a, complete cds, scaff2457, overlap with exons and introns of genscan prediction

9 - 2b1

>gi|74267554|dbj|AB208002.1| *Oryzias latipes* gene for hoxB6b, complete cds, scaff1130, overlap with exons and intron of genscan prediction

10 - 2b1

>gi|74267528|dbj|AB207989.1| *Oryzias latipes* gene for hoxA13b, complete cds, scaff13884, overlap with exons and intron of genscan prediction

11 - 2b2

>gi|32478197|gb|AY319970.1| *Oryzias latipes* brain aromatase (CYP19B) mRNA, complete cds, scaff400, overlap with genscan prediction

12 - 2c3

>gi|66796119|dbj|AB180742.1| *Oryzias latipes* RH1 gene for rod opsin, complete cds, scaff5553, overlap with ESTs, no intron-exon-structure

13 - 2b2

>gi|71361196|dbj|AB185848.1| *Oryzias latipes* coll mRNA for collagenase, complete cds, scaff21875, overlap with genscan prediction

14- 2c2

>gi|71360941|emb|AM051359.1| *Oryzias latipes* mRNA for beclin 1 (atg6 gene), scaff2280, overlap with EST

15 - 2b2

>gi|57162066|emb|AJ871609.1| *Oryzias latipes* mRNA for alpha 2,8 sialyltransferase (siat8e gene), scaff2321, overlap with genscan prediction

16 - 2b2

>gi|57162060|emb|AJ871606.1| *Oryzias latipes* mRNA for ganglioside GD3 synthase (siat8a gene), scaff6612, overlap with genscan prediction

17 – 2c2

>gi|32328900|dbj|AB075198.2| *Oryzias latipes* wap65 mRNA for warm-temperature-acclimation-related-65 kDa-protein, complete cds, scaff15668, overlap with ESTs

18 – 2b2

>gi|67077828|dbj|AB193469.1| *Oryzias latipes* timp-3 mRNA for tissue inhibitor of metalloproteinase-3, complete cds, scaff1295, overlap with genscan prediction

19 - 1

>gi|48857058|gb|AY627639.1| *Oryzias latipes* spindlin protein C (SpinC) mRNA, complete cds, no A

20 – 2c2

>gi|63002599|dbj|AB178523.1| *Oryzias latipes* plg mRNA for plasminogen, complete cds, scaff16119, overlap with ESTs, contig too short

21 – 2c2

>gi|51465797|dbj|AB188297.1| *Oryzias latipes* OIPC4 mRNA for general transcriptional coactivator, complete cds, scaff3859, overlap with ESTs

22 - 1

>gi|44886087|dbj|AB164051.1| *Oryzias latipes* cGK I beta mRNA for cGMP-dependent protein kinase I beta, complete cds, no full A

23 – 2c2

>gi|3288571|emb|Y11255.1|OLANNEXM4 *O. latipes* mRNA for annexin max4, scaff5593, overlap with ESTs

24 - 1

>gi|2344868|emb|Z97020.1|OLZ97020 *Oryzias latipes* mRNA for PAX2 protein no full A

25 – 2b2

>gi|2344866|emb|Z97019.1|OLZ97019 *Oryzias latipes* mRNA for sonic hedgehog, scaff35, overlap with genscan prediction

26 - 1

>gi|60649903|emb|AJ888977.1| *Oryzias latipes* mRNA for beta3-glucuronyltransferase (b3gat1 gene), no good A

27 -2b2

>gi|58864765|emb|AJ879587.1| *Oryzias latipes* mRNA for alpha3-fucosyltransferase (fut11 gene), scaff1287, overlap with genscan prediction

28 - 2b2

>gi|15667413|emb|AJ298300.1|OLA298300 *Oryzias latipes* mRNA for Rx3 protein, scaff7330, overlap with genscan prediction

29 – 2c2

>gi|12657449|emb|AJ278700.1|OLA278700 *Oryzias latipes* mRNA for nuclear receptor related 1 (nurr1 gene), scaff1910, overlap with genscan and fugu predictions

30 – 2c2

>gi|7635916|emb|AJ250405.1|OLA250405 *Oryzias latipes* mRNA for Rx2 transcription factor, scaff7422, overlap with ESTs

31 – 2b2

>gi|7635914|emb|AJ250404.1|OLA250404 *Oryzias latipes* mRNA for Vsx2 transcription factor, scaff210, overlap with genscan prediction

32 – 2c3

>gi|6688923|emb|AJ245396.1|OLA245396 *Oryzias latipes* mRNA for SOX3 protein, scaff2570, overlap with single exon ESTs

33 – 2c2

>gi|4585815|emb|AJ238012.1|OLA238012 *Oryzias latipes* mRNA for ferritin H3,

scaff6973, overlap with ESTs

34 – 2c2

>gi|4585813|emb|AJ238011.1|OLA238011 *Oryzias latipes* mRNA for pyruvate dehydrogenase, scaff6998, overlap with ESTs

35 - 1

>gi|4585811|emb|AJ238010.1|OLA238010 *Oryzias latipes* mRNA for heat shock protein 10, no A

36 – 2b2

>gi|4468636|emb|AJ131390.1|OLA131390 *Oryzias latipes* mRNA for tailless, scaff11195, overlap with genscan prediction

37 – 2b2

>gi|3115325|emb|AJ000939.1|OLJ000939 *Oryzias latipes*, *Otx2* gene, scaff31116, overlap with genscan prediction

38 - 1

>gi|3115323|emb|AJ000938.1|OLJ000938 *Oryzias latipes*, *Pax6* gene, no full A

39 – 2b2

>gi|62318451|dbj|AB193548.1| *Oryzias latipes* LEP mRNA for leptin, complete cds, scaff4865, overlap with genscan prediction

40 - 1

>gi|46397551|gb|AY518736.1| *Oryzias latipes* clone OLa20.03b *Fas* mRNA, complete cds, no full A

41 – 2c2

>gi|46397545|gb|AY518733.1| *Oryzias latipes* clone MF01SSA088G09 caspase-8 mRNA, complete cds, scaff273, overlap with ESTs and genscan predictions

42 - 1

>gi|62005801|dbj|AB190862.1| *Oryzias latipes* *gata-3* mRNA for transcription factor GATA-3, complete cds, no full A

43 – 2c2

>gi|50820169|gb|AY682722.1| *Oryzias latipes* alcohol dehydrogenase Class VI (ADH8) mRNA, complete cds, scaff546, overlap with ESTs

44 - 1

>gi|45378921|gb|AY560904.1| *Oryzias latipes* secreted frizzled-related protein 1 mRNA, complete cds, no full A

45 – 2b2

>gi|51465795|dbj|AB162905.1| *Oryzias latipes* SCP3 mRNA for synaptonemal complex protein 3, complete cds, scaff1012, partial overlap with genscan prediction

46 - 1

>gi|48147236|dbj|AB180923.1| *Oryzias latipes* *oleed* mRNA for embryonic ectoderm development protein, complete cds, no full A

47 – 2b1

>gi|52214090|dbj|AB160983.1| *Oryzias latipes* mRNA for yippee-like b, complete cds, scaff1277, partial overlap with genscan predicted exon and intron

48 - 1

>gi|40891599|gb|AY363113.1| *Oryzias latipes* zinc finger protein 183 mRNA, complete cds, no full A

49 - 1

>gi|46561999|gb|AY575076.1| *Oryzias latipes* SPARC mRNA, complete cds, no full A

50 – 2c2

>gi|50251129|dbj|AB099891.1| *Oryzias latipes* *ctsS* mRNA for cathepsin S, complete cds, scaff8882, overlap with ESTs and genscan prediction

Alignment of 50 randomly selected EST clusters

1 – 2c3
CL3Contig4, scaff17724, overlap mit EST, no intron-exon-structure
2 – 2c2
CL5Contig13, many hits, best hit on scaff 9116, overlap mit ESTs
3 – 2c2
CL14Contig3, many hits, best hit on scaff 5411,
4 – 2c2
CL24Contig1, scaff 9442, overlap with ESTs
5 – 2c2
CL34Contig2, many hits, best hit on scaff3623, overlap with ESTs
6 – 2b1
CL35Contig1, many hits, best hit on scaff14927, overlap with intron of gescan prediction
7 – 2b1
CL40Contig1, many hits, best hit on scaff13201, overlap with gescan intron
8 – 2b2
CL67Contig1, scaff13534, partial overlap with gescan prediction exons
9 – 2c2
CL109Contig1, scaff4191, overlap with ESTs
10 – 2b2
CL146Contig1, scaff12733, overlap with gescan prediction
11 – 2c1
CL3552Contig1, scaff749, partial overlap of one exon with single exon ESTs
12 -1
CL3442Contig1, no good A
13 – 2a1
CL3159Contig1, scaff10172, no overlap, < 500 to gescan, < 2000bp to public evidence
14 – 2b1
CL3033Contig1, scaff236, partial overlap with gescan prediction
15 – 2b1
CL2988Contig1, scaff88, partial overlap with gescan prediction, introns and exons
16 – 2b2
CL2837Contig1, scaff14043, partial overlap with gescan prediction
17 – 2a2
CL2779Contig1, scaff4409, no overlap, < 1000bp to gescan
18 – 2c2
CL2607Contig1, scaff24199, overlap with ESTs
19 – 2c2
CL2529Contig1, scaff104, overlap with EST
20 - 1
CL2469Contig1, no good A
21 – 2b1
CL2386Contig1, scaff12003, overlap with gescan intron
22 – 2c3
CL2242Contig1, scaff4274, overlap with single exon EST
23 - 1

CL2138Contig1, no A
24 – 2c3
CL2049Contig1, scaff7672, overlap with single exon EST
25 – 2c3
CL1971Contig1, scaff7800, overlap with ESTs
26 - 1
CL1863Contig1, no A
27 – 2a1
CL1785Contig1, scaff319, no overlap, < 500bp to genscan
28 – 2c1
CL1666Contig1, scaff677, overlap with EST intron, genscan predicted exons
29 - 1
CL1539Contig1, no full A
30 - 1
CL1412Contig1, no full A
31 – 2b2
CL1351Contig1, scaff2564, partial overlap with genscan prediction
32 – 2b1
CL1239Contig1, scaff3512, overlap with genscan predicted exon and intron
33 - 1
CL1135Contig1, no full A
34 – 2a3
CL1087Contig1, scaff2299, no overlap, < 3000bp to genscan and public
35 – 2b1
CL933Contig1, scaff12027, partial overlap of genscan predicted exon
36 - 1
CL800Contig1, no A
37 – 2c2
CL778Contig1, scaff2154, overlap with ESTs
38 – 2a1
CL610Contig1, scaff153, no overlap, < 500bp to mapped gene
39 – 2c2
CL528Contig1, scaff5759, overlap with ESTs
40 – 2b1
CL465Contig1, scaff166, ??
41 – 2c2
CL362Contig1, scaff2379, overlap with ESTs
42 - 1
CL261Contig1, no A
43 – 2b2
CL67Contig3, scaff26197, overlap with genscan prediction
44 – 2b1
CL293Contig1, scaff4743, overlap with genscan prediction
45 – 2c2
CL497Contig1, scaff2198, overlap with ESTs
46 – 2b2
CL768Contig1, scaff2219, overlap with genscan prediction
47 – 2a2
CL1089Contig1, scaff11007, no overlap, < 1000bp to genscan, < 2000bp to public
48 – 2b1
>CL1371Contig1, scaff9814, partial overlap with genscan predicted exon and intron

49 – 2b2

CL1660Contig1, scaff9321, overlap with genscan prediction

50 – 2b1

CL1937Contig1, scaff230, partial overlap with genscan predicted exon

Alignment of 50 randomly selected EST singletons

1 – 2a2

McF0011M19-MGRbd1, scaff3402, no overlap, < 1000bp to genscan

2 - 1

McF0012N03-MGRbd1, no good A

3 – 2b2

McF0013O10-MGRbd1, scaff3232, overlap with genscan prediction

4 - 1

McF0014P24-MGRbd1, no good A

5 - 1

McF0015D01-MGRbd1, no good A

6 – 2b1

McF0016F08-MGRbd1, scaff9556, overlap with genscan intron and exon

7 – 2c3

McF0017G09-MGRbd1, scaff7985, overlap with ESTs and genscan intron

8 – 2b1

McF0018H09-MGRbd1, scaff5601, no overlap

9 – 2c2

McF0019I04-MGRbd1, many hits, best hit on scaff2768, overlap with ESTs

10 – 2b1

McF0020J05-MGRbd1, scaff4585, overlap with genscan intron

11 - 2b1

McF0001MGR-1A04bd1, scaff9946, little overlap with genscan prediction

12 – 2c3

McF0002C16-MGRbd1, scaff4854, overlap with many single exon ESTs,

13 – 2b1

McF0003D01-MGRbd1, scaff1822, overlap with intron of genscan prediction

14 – 2b1

McF0004E14-MGRbd1, scaff2532, partial overlap with genscan exon

15 - 1

McF0005G11-MGRbd1, noA

16 – 2c3

McF0006H11-MGRbd1, scaff167, overlap with many single exon ESTs

17 - 1

McF0007I13-MGRbd1, no good A

18 – 2b1

McF0008J19-MGRbd1, scaff750, overlap with genscan intron

19 – 2b1

McF0009K16-MGRbd1, scaff3933, partial overlap with genscan prediction

20 – 2b1

McF0010L07-MGRbd1, scaff9653, partial overlap with genscan exon

21 – 2b1

McF0011N19-MGRbd1, scaff15, overlap with genscan predicted intron

22 - 1

McF0012O04-MGRbd1, no A

23 – 2c1
McF0013P03-MGRbd1, scaff5762, overlap with EST intron and exon
24 – 2b1
McF0014A03-MGRbd1, scaff8937, overlap with EST intron and genscan predicted
exon
25 – 2b1
McF0015D22-MGRbd1, scaff5601, partial overlap with genscan intron
26 – 2a3
McF0016F16-MGRbd1, scaff5653, no overlap, < 2000bp to genscan, < 5000bp to
public evidence
27 – 2a1
McF0017G10-MGRbd1, scaff8581, no overlap, < 500bp to genscan, < 1000bp to
public evidence
28 – 2b1
McF0018H17-MGRbd1, scaff937, partial overlap with genscan exon and intron
29 – 2b1
McF0019I08-MGRbd1, scaff193, overlap with genscan intron
30 – 2c1
McF0020J12-MGRbd1, scaff3406, partial overlap with ESTs and genscan
31 – 2b1
McF0021K10-MGRbd1, scaff3522, overlap with genscan intron
32 – 2b1
McF0022L10-MGRbd1, scaff1531, overlap with genscan intron
33 – 2b1
McF0023M19-MGRbd2, scaff83863, partial overlap with genscan exon
34 - 1
McF0024O09-MGRbd2, no good A
35 – 2c2
>McF0025P07-MGRbd1, scaff6470, overlap with ESTs
36 – 2a1
McF0026A21-MGRbd1, scaff1183, no overlap, < 500bp to genscan
37 – 2b1
McF0027B16-MGRbd1, scaff160, partial overlap with genscan
38 – 2a3
McF0028C21-MGRbd1, scaff25963, no overlap, < 3000bp to genscan
39 – 2b1
McF0029D08-MGRbd1, scaff12430, small overlap with genscan exon
40 -1
McF0030F02-MGRbd1, no A
41 -1
McF0031G23-MGRbd1, no good A
42 – 2a3
McF0032I03-MGRbd1, scaff74, no overlap, < 3000bp to genscan
43 – 2b1
McF0033J21-MGRbd1, scaff1770, overlap ???
44 – 2b1
McF0034L22-MGRbd1, scaff7489, overlap with genscan intron
45 – 2b1
McF0035M05-MGRbd1, scaff14660, overlap with genscan intron
46 - 1
McF0036N23-MGRbd1, no full A

47 – 2b1

McF0037O17-MGRbd1, scaff4895, overlap with genscan intron

48 - 1

McF0039P17-MGRbd1, no A

49 – 2a2

McF0040B17-MGRbd1, scaff52, no overlap, < 1000 bp to genscan prediction

50 – 2b1

McF0041D07-MGRbd1, scaff7177, overlap with genscan intron and exon