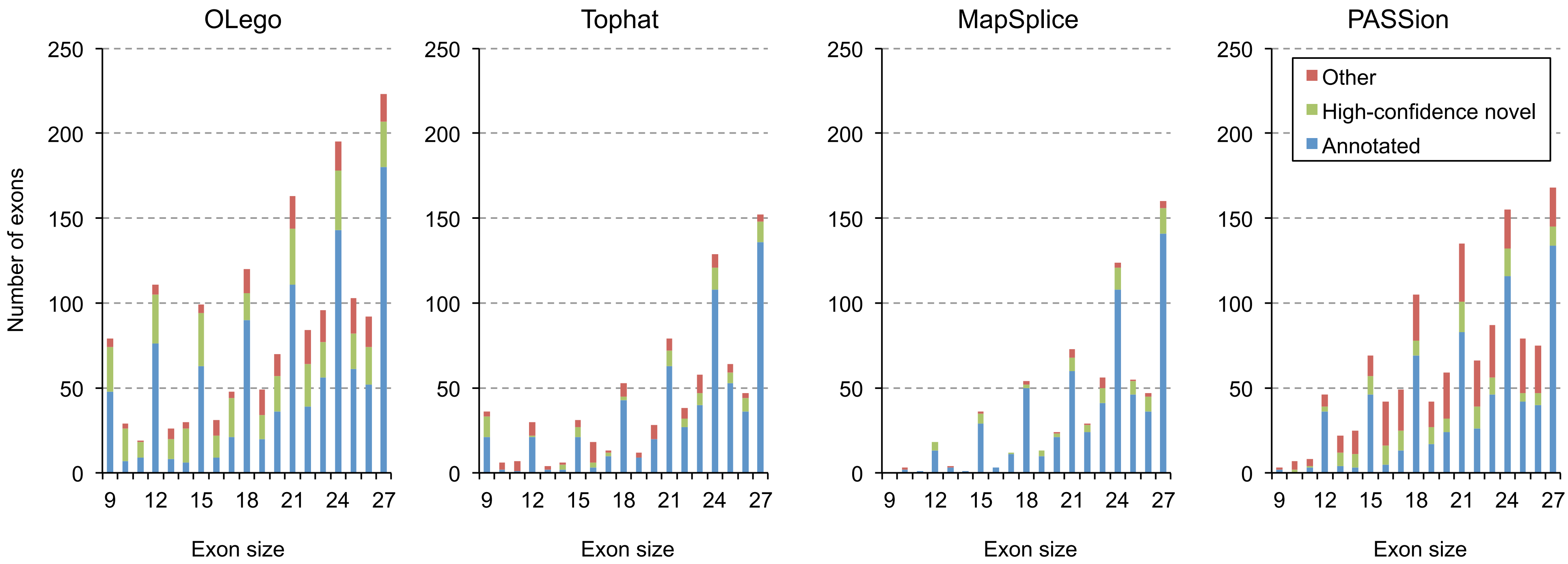
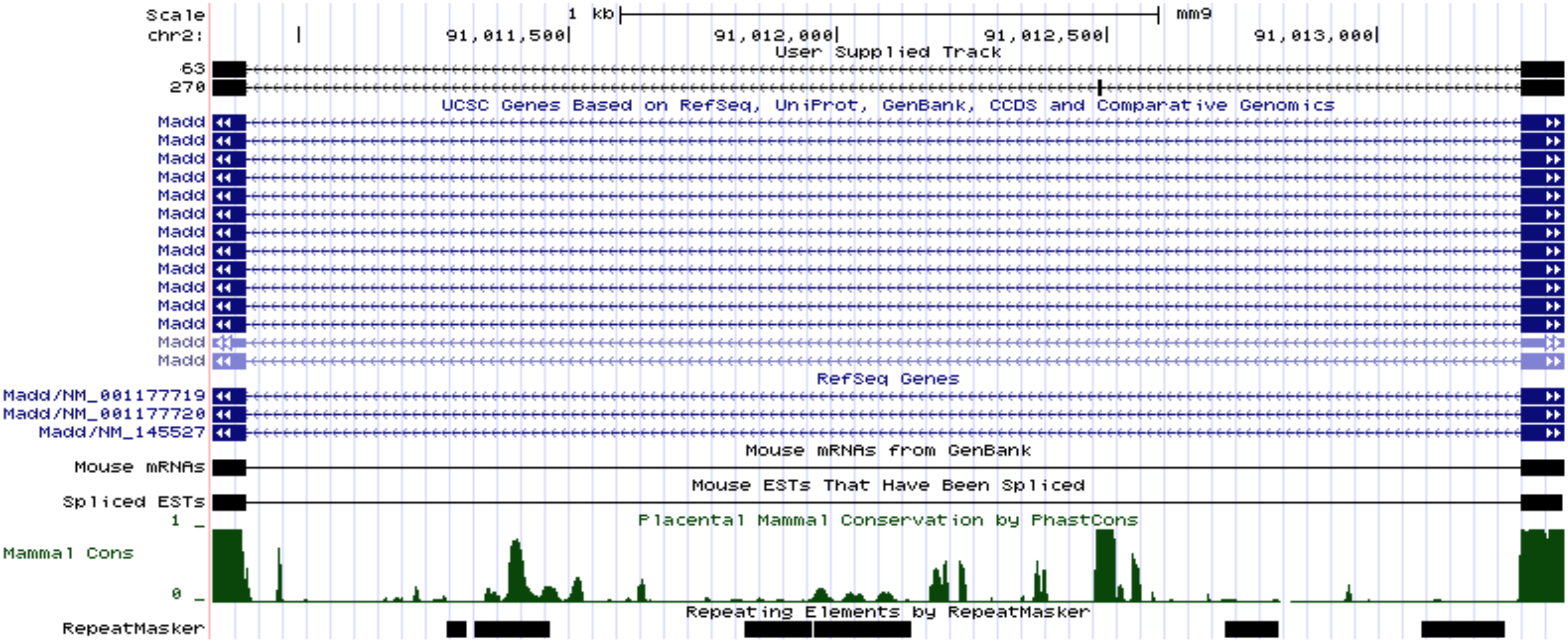


Supplementary Figure 1: Number of micro-exons identified by OLego and the other three programs.

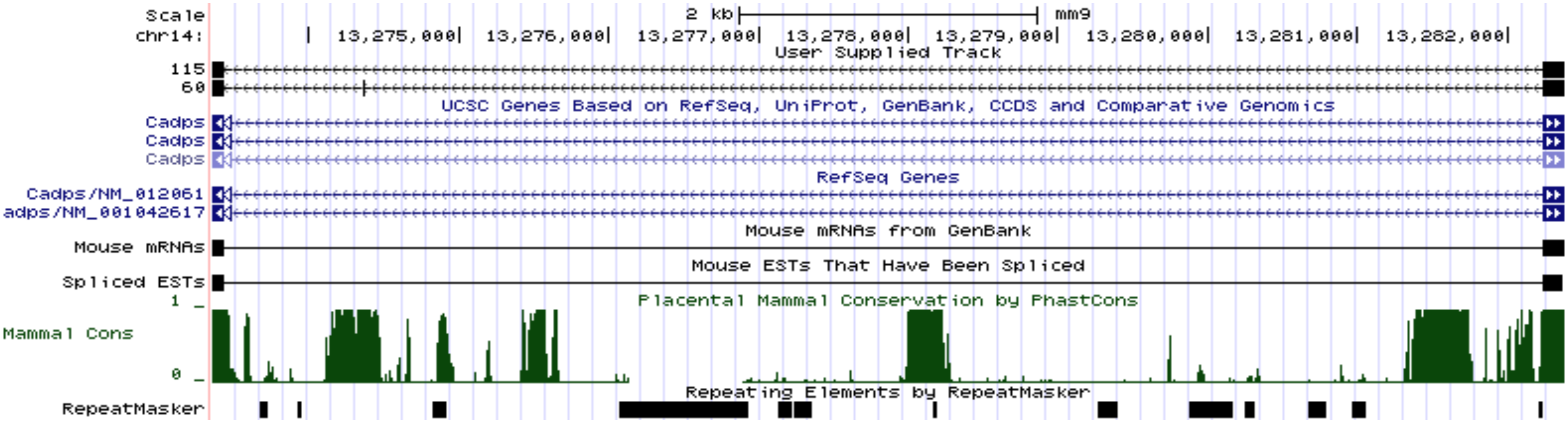
Exons are binned by their sizes (9~27nt), and in each bin, they are classified into three groups: annotated micro-exons in previous gene models (blue), high-confidence novel micro-exons (exons with both flanking constitutive splice sites annotated; green), and other exons (red).

Supplementary Figure 2: UCSC genome browser screenshots of the validated micro-exons.

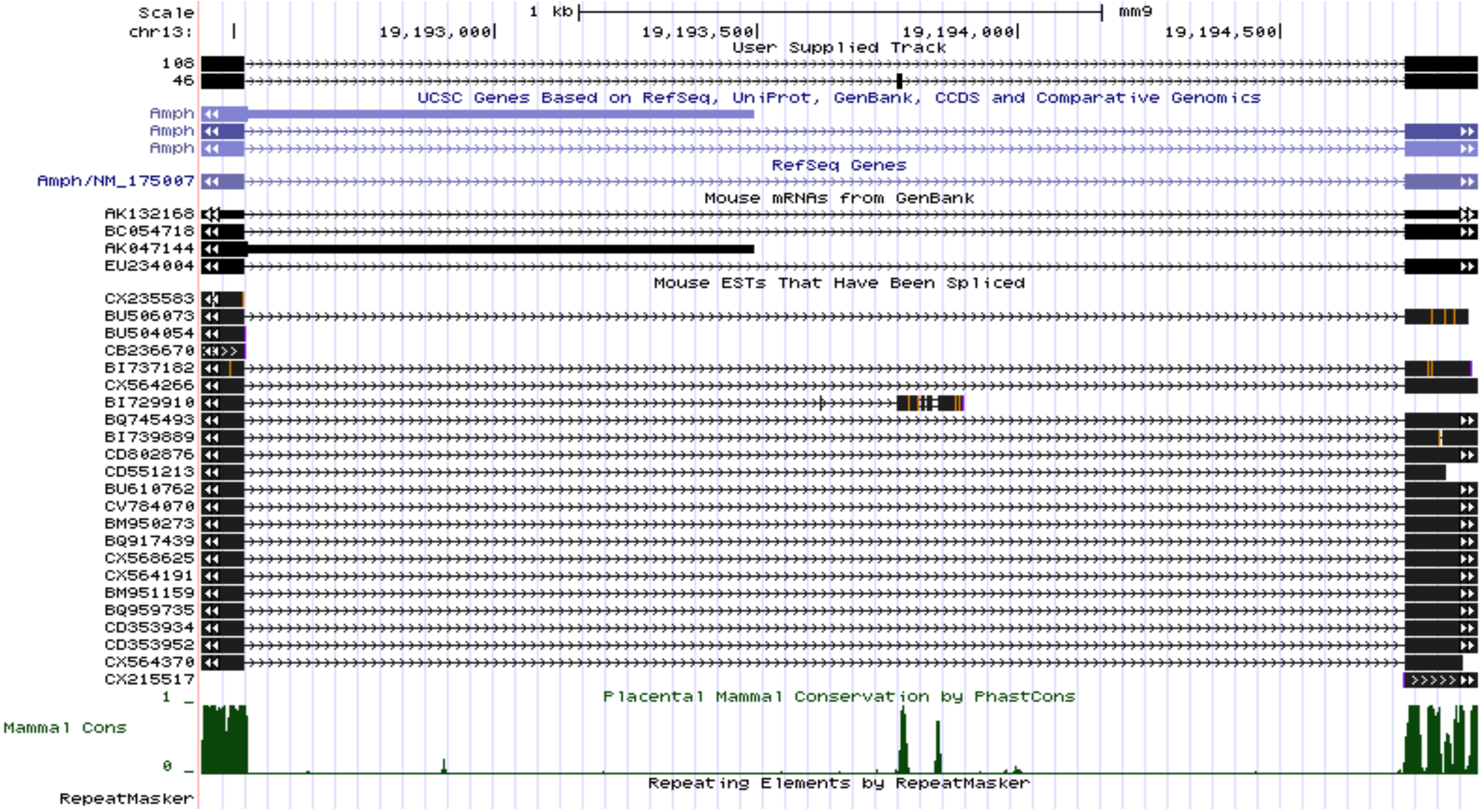
Structures of the inclusion and exclusion isoforms are shown in the user supplied track. The number of supporting reads is indicated to the left of each isoform. Known transcript structures (UCSC genes, RefSeq genes and ESTs) are also shown in the screenshots. The micro-exons are in the same order as they are listed in Supplementary Table 2.


Scale
1 kb
mm9
chr2:
91,011,500
91,012,000
91,012,500
91,013,000
User Supplied Track
63
270
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics
Madd
RefSeq Genes
Madd/NM_001177719
Madd/NM_001177720
Madd/NM_145527
Mouse mRNAs from GenBank
Mouse mRNAs
Mouse ESTs That Have Been Spliced
Spliced ESTs
Placental Mammal Conservation by PhastCons
1
Mammal Cons
0
Repeating Elements by RepeatMasker
RepeatMasker

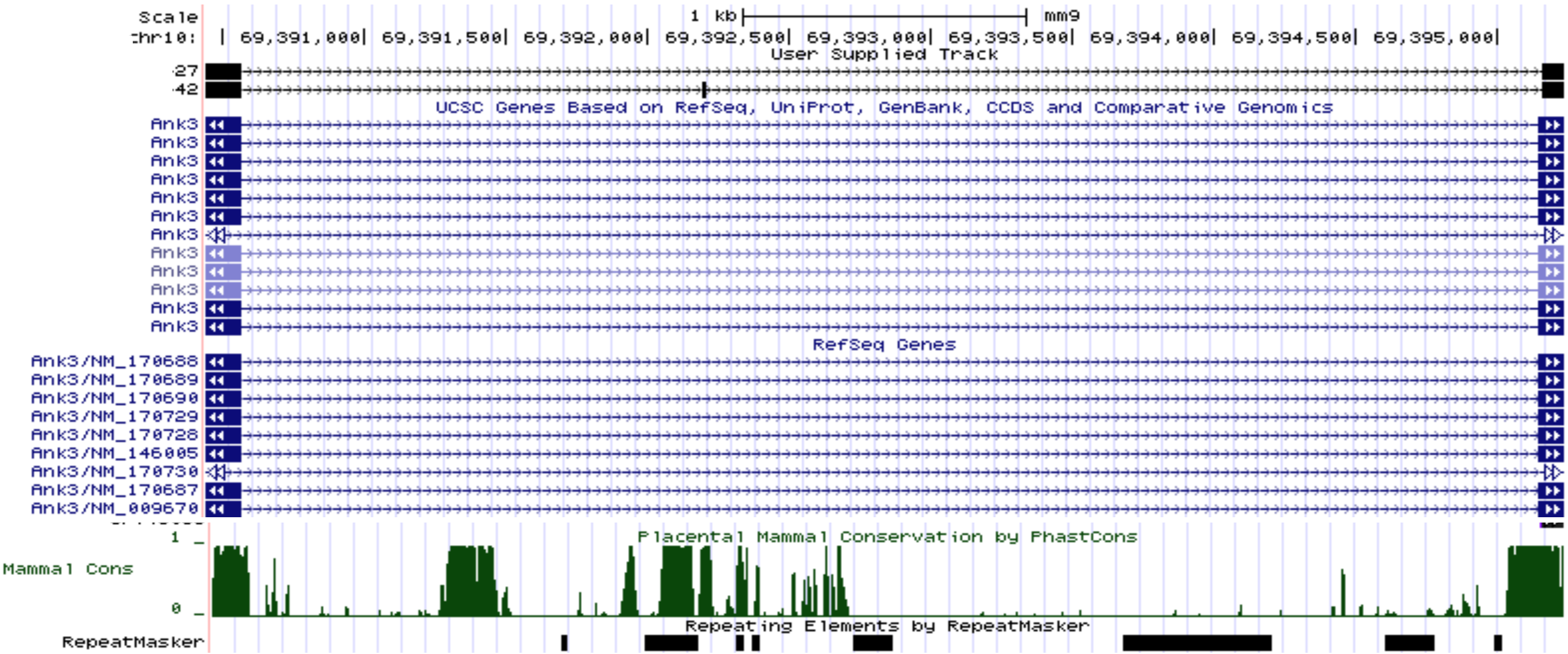
Madd


Scale
2 kb
mm9
chr14:
13,275,000
13,276,000
13,277,000
13,278,000
13,279,000
13,280,000
13,281,000
13,282,000
User Supplied Track
115
60
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics
Cadps
Cadps
Cadps
RefSeq Genes
Cadps/NM_012061
adps/NM_001042617
Mouse mRNAs from GenBank
Mouse mRNAs
Mouse ESTs That Have Been Spliced
Spliced ESTs
Placental Mammal Conservation by PhastCons
1
Mammal Cons
0
Repeating Elements by RepeatMasker
RepeatMasker

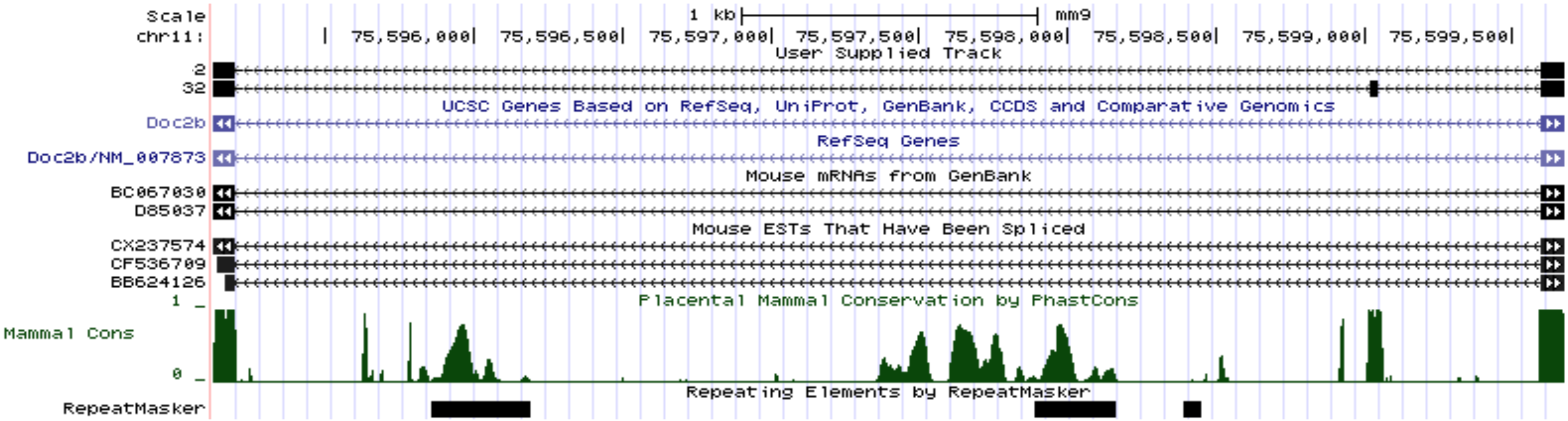
Cadps


Scale
chr13:
1 kb
mm9
19,193,000
19,193,500
19,194,000
19,194,500
User Supplied Track
108
46
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics
Amph
Amph
Amph
RefSeq Genes
Amph/NM_175007
Mouse mRNAs from GenBank
AK132168
BC054718
AK047144
EU234004
Mouse ESTs That Have Been Spliced
CX235583
BU506073
BU504054
CB236670
BI737182
CX564266
BI729910
BQ745493
BI739889
CD802876
CD551213
BU610762
CV784070
BM950273
BQ917439
CX568625
CX564191
BM951159
BQ959735
CD353934
CD353952
CX564370
CX215517
Placental Mammal Conservation by PhastCons
1
Mammal Cons
0
Repeating Elements by RepeatMasker
RepeatMasker

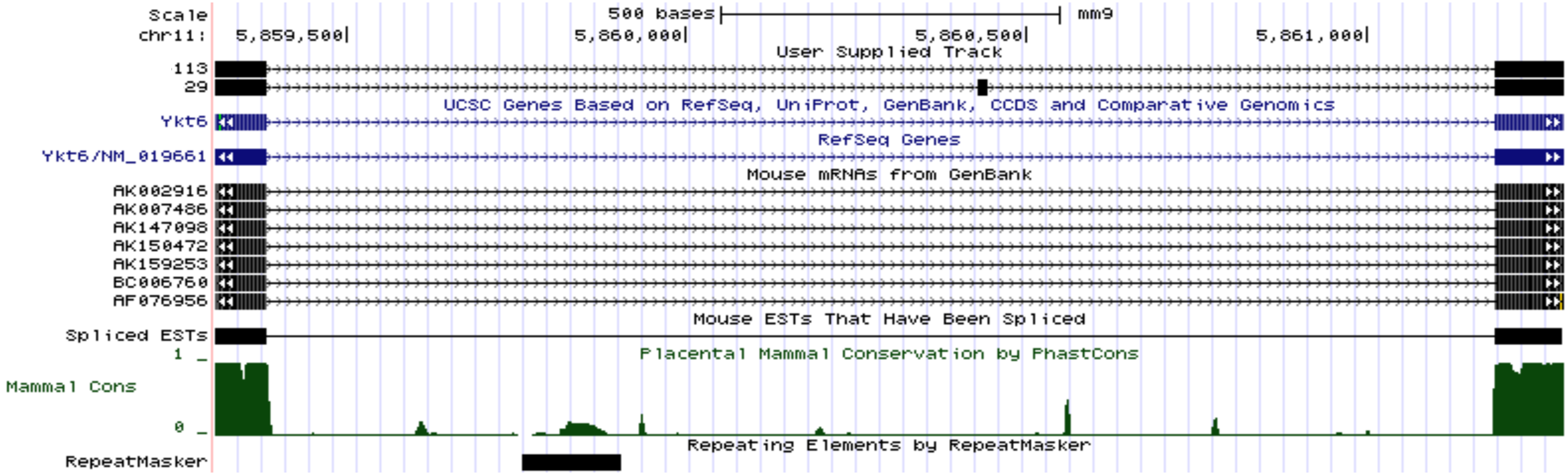
Amph


Scale
1 kb
mm9
hr10:
69,391,000
69,391,500
69,392,000
69,392,500
69,393,000
69,393,500
69,394,000
69,394,500
69,395,000
User Supplied Track
27
42
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics
Ank3
Ank3
Ank3
Ank3
Ank3
Ank3
Ank3
Ank3
Ank3
Ank3
Ank3
Ank3
RefSeq Genes
Ank3/NM_170688
Ank3/NM_170689
Ank3/NM_170690
Ank3/NM_170729
Ank3/NM_170728
Ank3/NM_146005
Ank3/NM_170730
Ank3/NM_170687
Ank3/NM_009670
Placental Mammal Conservation by PhastCons
1
Mammal Cons
0
Repeating Elements by RepeatMasker
RepeatMasker

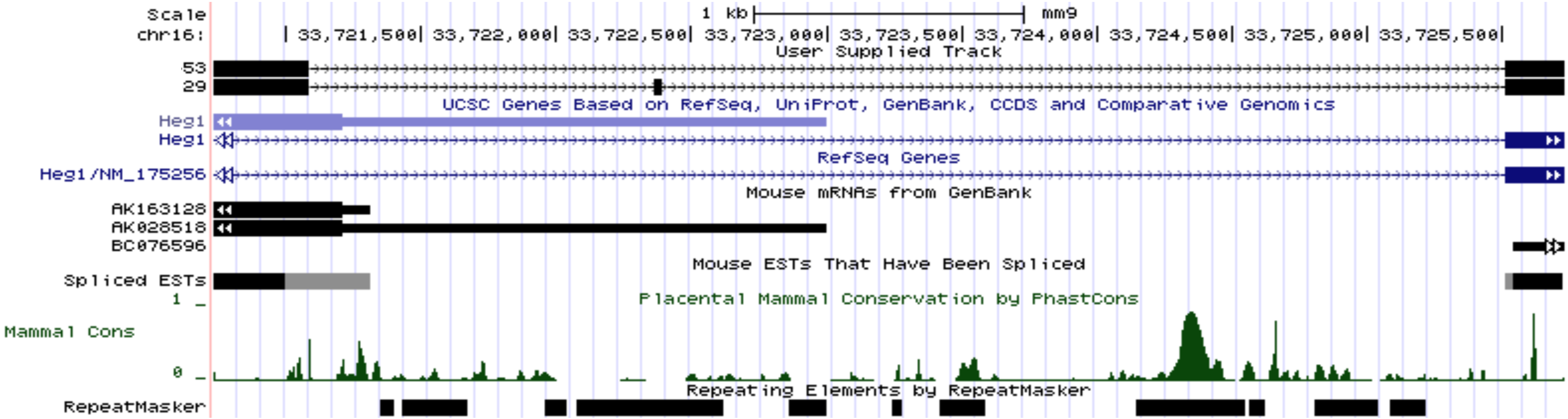
Ank3


Scale
chr11:
1 kb
mm9
75,596,000
75,596,500
75,597,000
75,597,500
75,598,000
75,598,500
75,599,000
75,599,500
User Supplied Track
2
32
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics
Doc2b
RefSeq Genes
Doc2b/NM_007873
Mouse mRNAs from GenBank
BC067030
D85037
Mouse ESTs That Have Been Spliced
CX237574
CF536709
BB624126
1
Placental Mammal Conservation by PhastCons
Mammal Cons
0
Repeating Elements by RepeatMasker
RepeatMasker

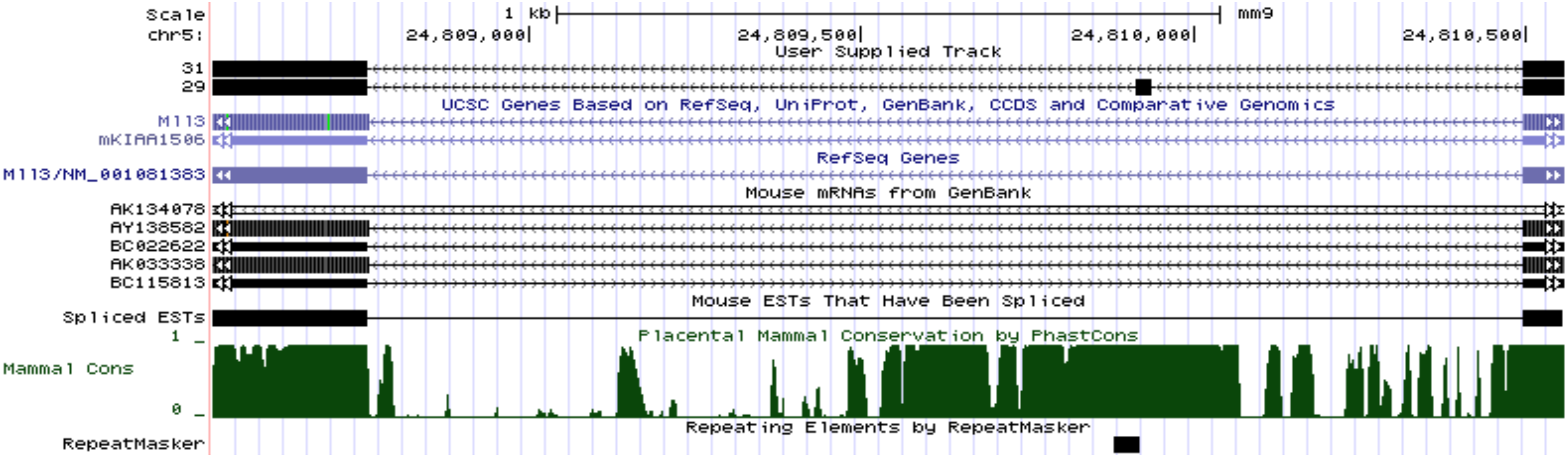
Doc2b


Scale
500 bases
mm9
chr11:
5,859,500
5,860,000
5,860,500
5,861,000
User Supplied Track
113
29
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics
Ykt6
RefSeq Genes
Ykt6/NM_019661
Mouse mRNAs from GenBank
AK002916
AK007486
AK147098
AK150472
AK159253
BC006760
AF076956
Mouse ESTs That Have Been Spliced
Spliced ESTs
Placental Mammal Conservation by PhastCons
1
Mammal Cons
0
Repeating Elements by RepeatMasker
RepeatMasker

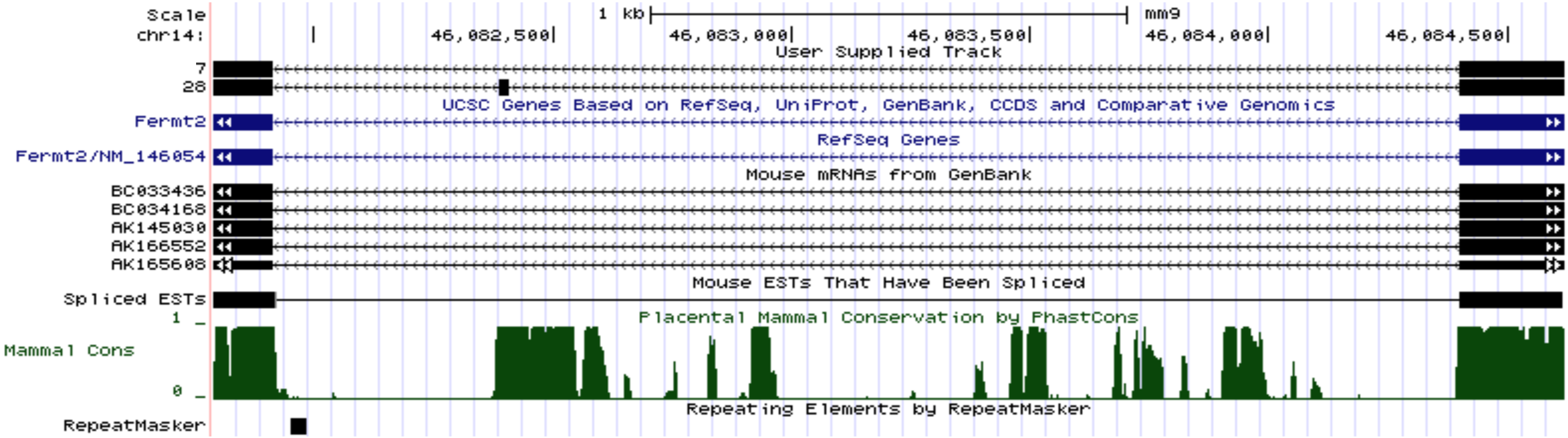
Ykt6


Scale
chr16:
1 kb
mm9
33,721,500
33,722,000
33,722,500
33,723,000
33,723,500
33,724,000
33,724,500
33,725,000
33,725,500
User Supplied Track
53
29
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics
Heg1
Heg1
RefSeq Genes
Heg1/NM_175256
Mouse mRNAs from GenBank
AK163128
AK028518
BC076596
Mouse ESTs That Have Been Spliced
Spliced ESTs
1
Placental Mammal Conservation by PhastCons
Mammal Cons
0
Repeating Elements by RepeatMasker
RepeatMasker

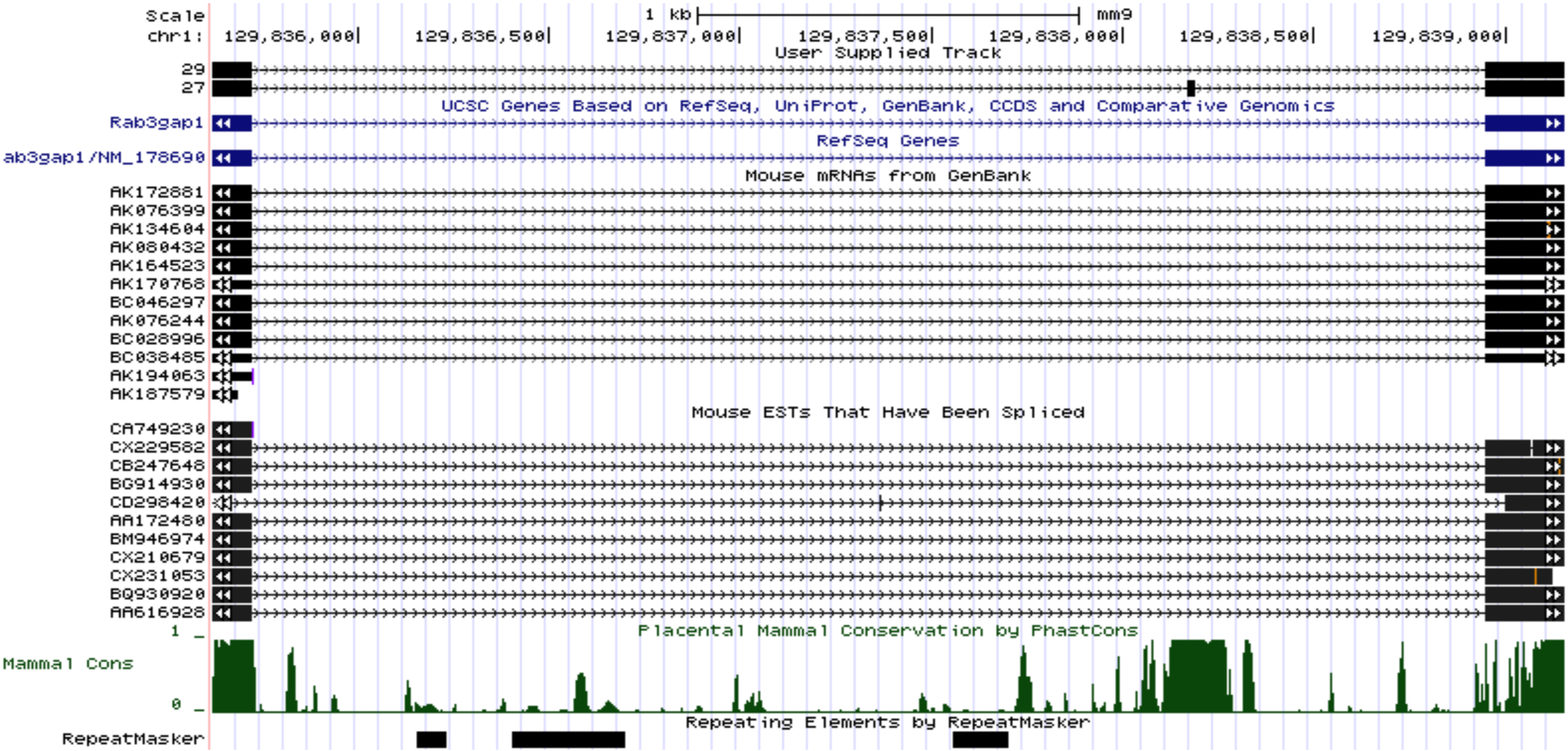
Heg1


Scale
1 kb
mm9
chr5:
24,809,000
24,809,500
24,810,000
24,810,500
User Supplied Track
31
29
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics
Mll3
mKIAA1506
RefSeq Genes
Mll3/NM_001081383
Mouse mRNAs from GenBank
AK134078
AY138582
BC022622
AK033338
BC115813
Mouse ESTs That Have Been Spliced
Spliced ESTs
Placental Mammal Conservation by PhastCons
1
Mammal Cons
0
Repeating Elements by RepeatMasker
RepeatMasker

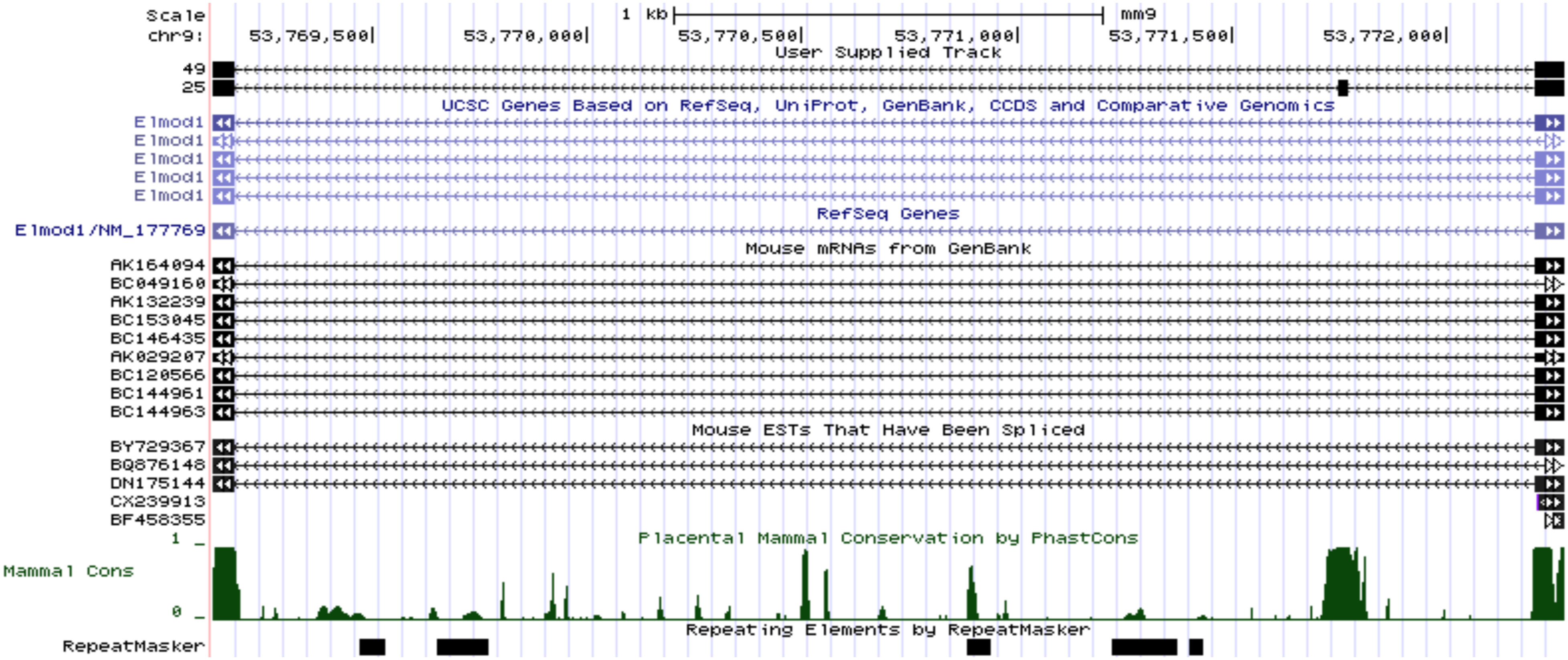
Mll3


Scale
1 kb
mm9
chr14:
46,082,500
46,083,000
46,083,500
46,084,000
46,084,500
User Supplied Track
7
28
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics
Fermt2
RefSeq Genes
Fermt2/NM_146054
Mouse mRNAs from GenBank
BC033436
BC034168
AK145030
AK166552
AK165608
Mouse ESTs That Have Been Spliced
Spliced ESTs
Placental Mammal Conservation by PhastCons
1
Mammal Cons
0
Repeating Elements by RepeatMasker
RepeatMasker

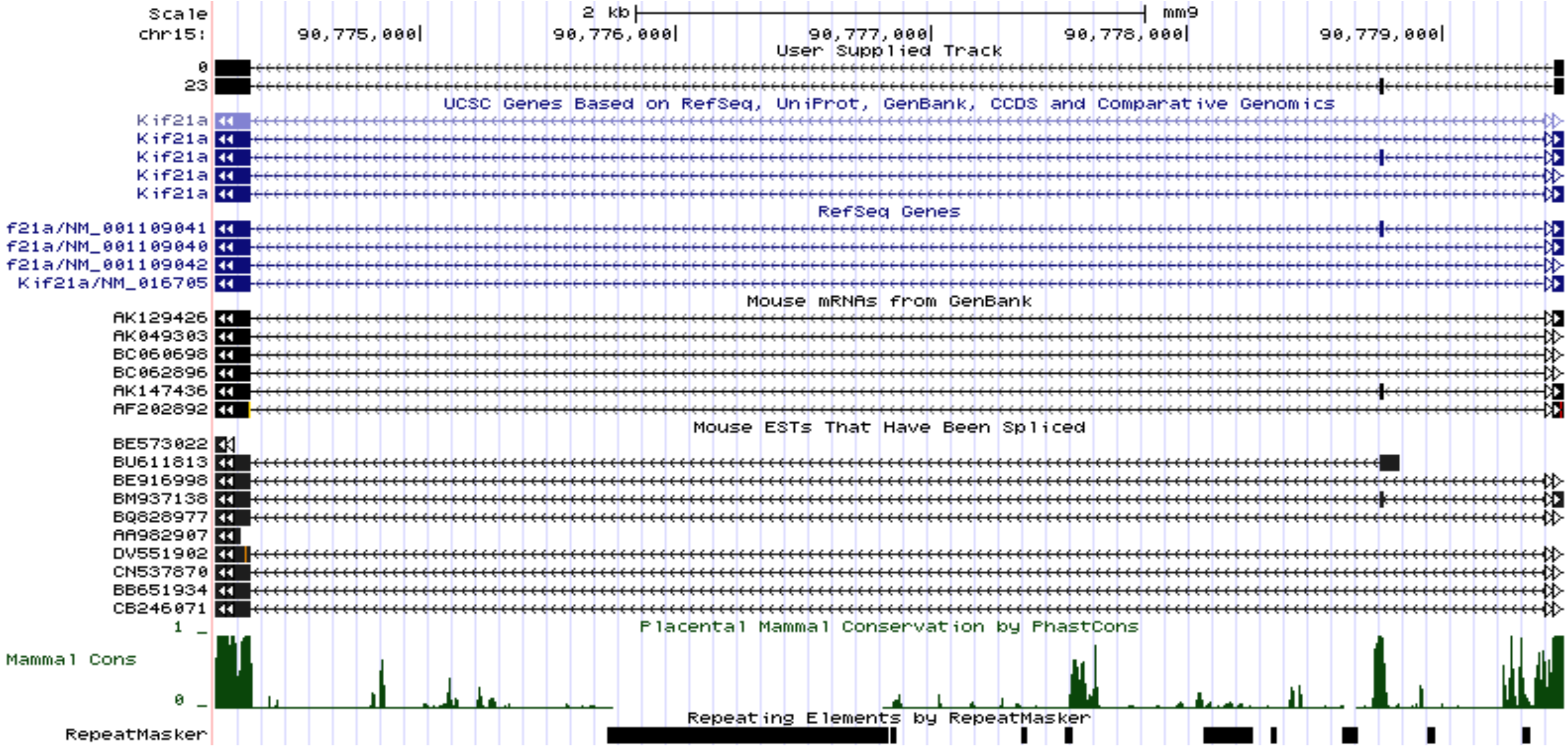
Fermt2


Scale
chr1:
1 kb
mm9
129,836,000
129,836,500
129,837,000
129,837,500
129,838,000
129,838,500
129,839,000
User Supplied Track
29
27
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics
Rab3gap1
RefSeq Genes
ab3gap1/NM_178690
Mouse mRNAs from GenBank
AK172881
AK076399
AK134604
AK080432
AK164523
AK170768
BC046297
AK076244
BC028996
BC038485
AK194063
AK187579
Mouse ESTs That Have Been Spliced
CA749230
CX229582
CB247648
BG914930
CD298420
AA172480
BM946974
CX210679
CX231053
BQ930920
AA616928
Placental Mammal Conservation by PhastCons
1
Mammal Cons
0
Repeating Elements by RepeatMasker
RepeatMasker

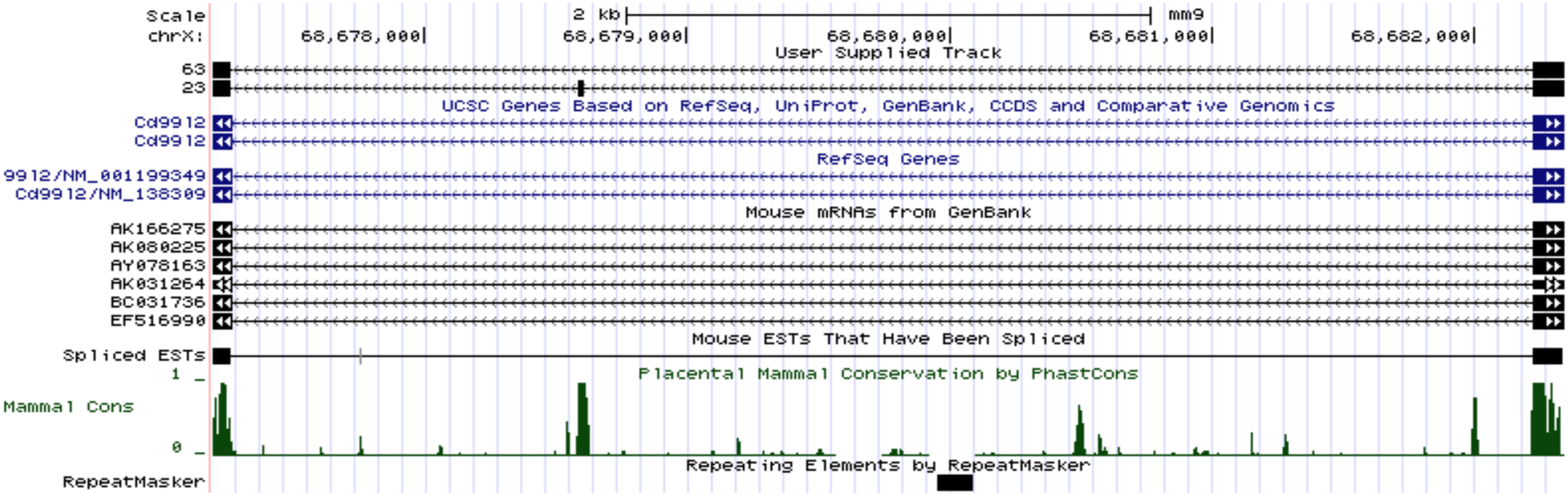
Rab3gap1


Scale
chr9:
1 kb
mm9
53,769,500
53,770,000
53,770,500
53,771,000
53,771,500
53,772,000
User Supplied Track
49
25
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics
Elmod1
Elmod1
Elmod1
Elmod1
Elmod1
RefSeq Genes
Elmod1/NM_177769
Mouse mRNAs from GenBank
AK164094
BC049160
AK132239
BC153045
BC146435
AK029207
BC120566
BC144961
BC144963
Mouse ESTs That Have Been Spliced
BY729367
BQ876148
DN175144
CX239913
BF458355
Placental Mammal Conservation by PhastCons
1
Mammal Cons
0
Repeating Elements by RepeatMasker
RepeatMasker

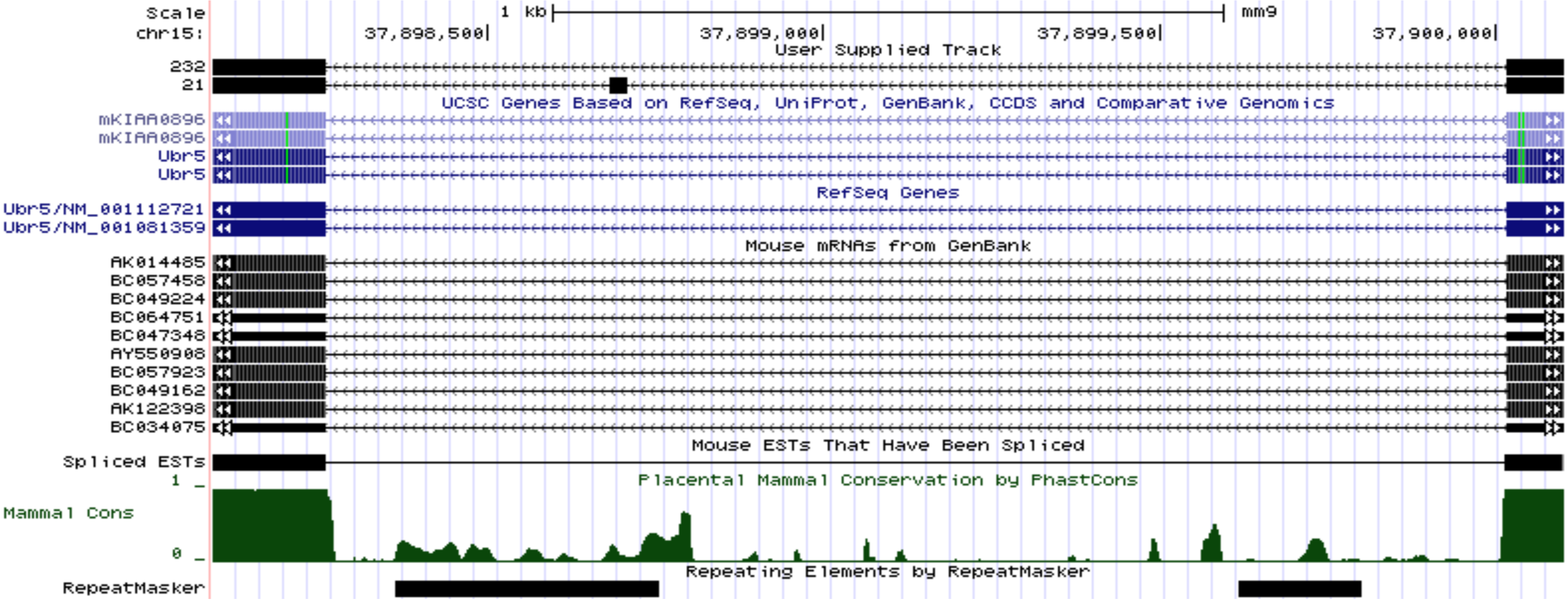
Elmod1


Scale
chr15:
2 kb
mm9
90,775,000
90,776,000
90,777,000
90,778,000
90,779,000
User Supplied Track
0
23
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics
Kif21a
Kif21a
Kif21a
Kif21a
Kif21a
RefSeq Genes
f21a/NM_001109041
f21a/NM_001109040
f21a/NM_001109042
Kif21a/NM_016705
Mouse mRNAs from GenBank
AK129426
AK049303
BC060698
BC062896
AK147436
AF202892
Mouse ESTs That Have Been Spliced
BE573022
BU611813
BE916998
BM937138
BQ828977
AA982907
DV551902
CN537870
BB651934
CB246071
1 _
Placental Mammal Conservation by PhastCons
Mammal Cons
0 _
Repeating Elements by RepeatMasker
RepeatMasker

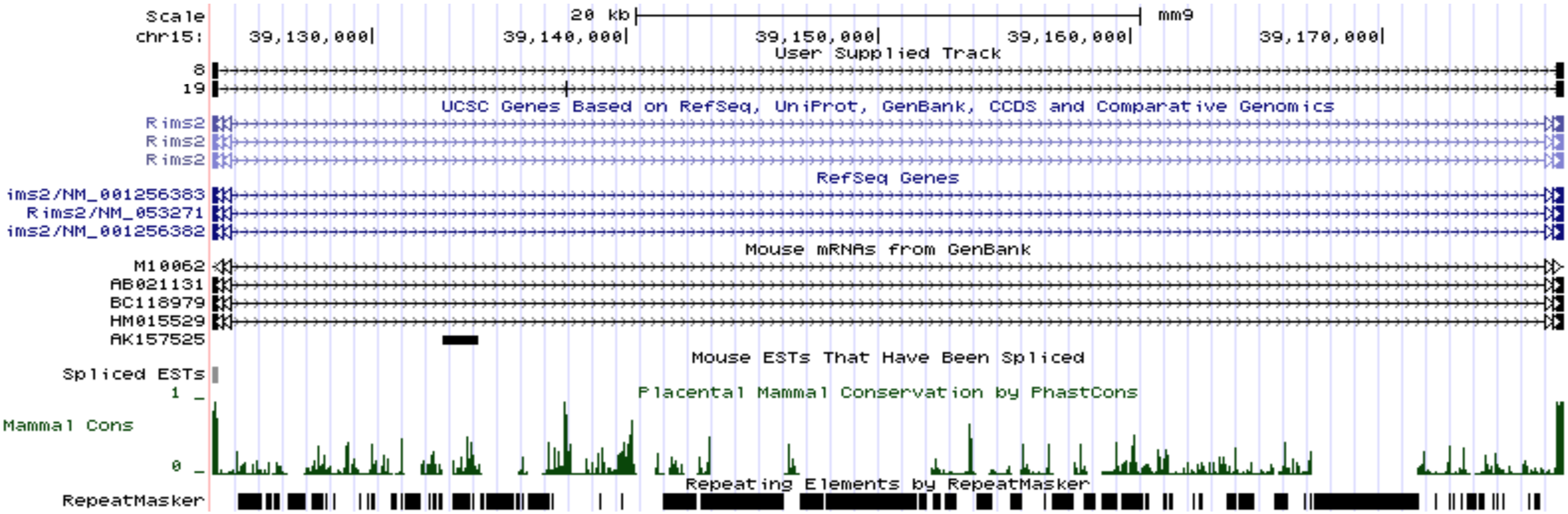
Kif21a


Scale
chrX:
2 kb
mm9
68,678,000
68,679,000
68,680,000
68,681,000
68,682,000
User Supplied Track
63
23
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics
Cd9912
Cd9912
RefSeq Genes
9912/NM_001199349
Cd9912/NM_138309
Mouse mRNAs from GenBank
AK166275
AK080225
AY078163
AK031264
BC031736
EF516990
Mouse ESTs That Have Been Spliced
Spliced ESTs
Placental Mammal Conservation by PhastCons
1
Mammal Cons
0
Repeating Elements by RepeatMasker
RepeatMasker

Cd9912


Scale
chr15:
1 kb
mm9
37,898,500
37,899,000
37,899,500
37,900,000
User Supplied Track
232
21
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics
mKIAA0896
mKIAA0896
Ubr5
Ubr5
RefSeq Genes
Ubr5/NM_001112721
Ubr5/NM_001081359
Mouse mRNAs from GenBank
AK014485
BC057458
BC049224
BC064751
BC047348
AY550908
BC057923
BC049162
AK122398
BC034075
Mouse ESTs That Have Been Spliced
Spliced ESTs
1 _
Placental Mammal Conservation by PhastCons
Mammal Cons
0 _
Repeating Elements by RepeatMasker
RepeatMasker

Ubr5


Scale
chr15:
20 kb
mm9
39,130,000
39,140,000
39,150,000
39,160,000
39,170,000
User Supplied Track
8
19
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics
Rims2
Rims2
Rims2
RefSeq Genes
ims2/NM_001256383
Rims2/NM_053271
ims2/NM_001256382
Mouse mRNAs from GenBank
M10062
AB021131
BC118979
HM015529
AK157525
Mouse ESTs That Have Been Spliced
Spliced ESTs
Placental Mammal Conservation by PhastCons
1 _
Mammal Cons
0 _
Repeating Elements by RepeatMasker
RepeatMasker

Rims2