

Supplementary Material Figure 1

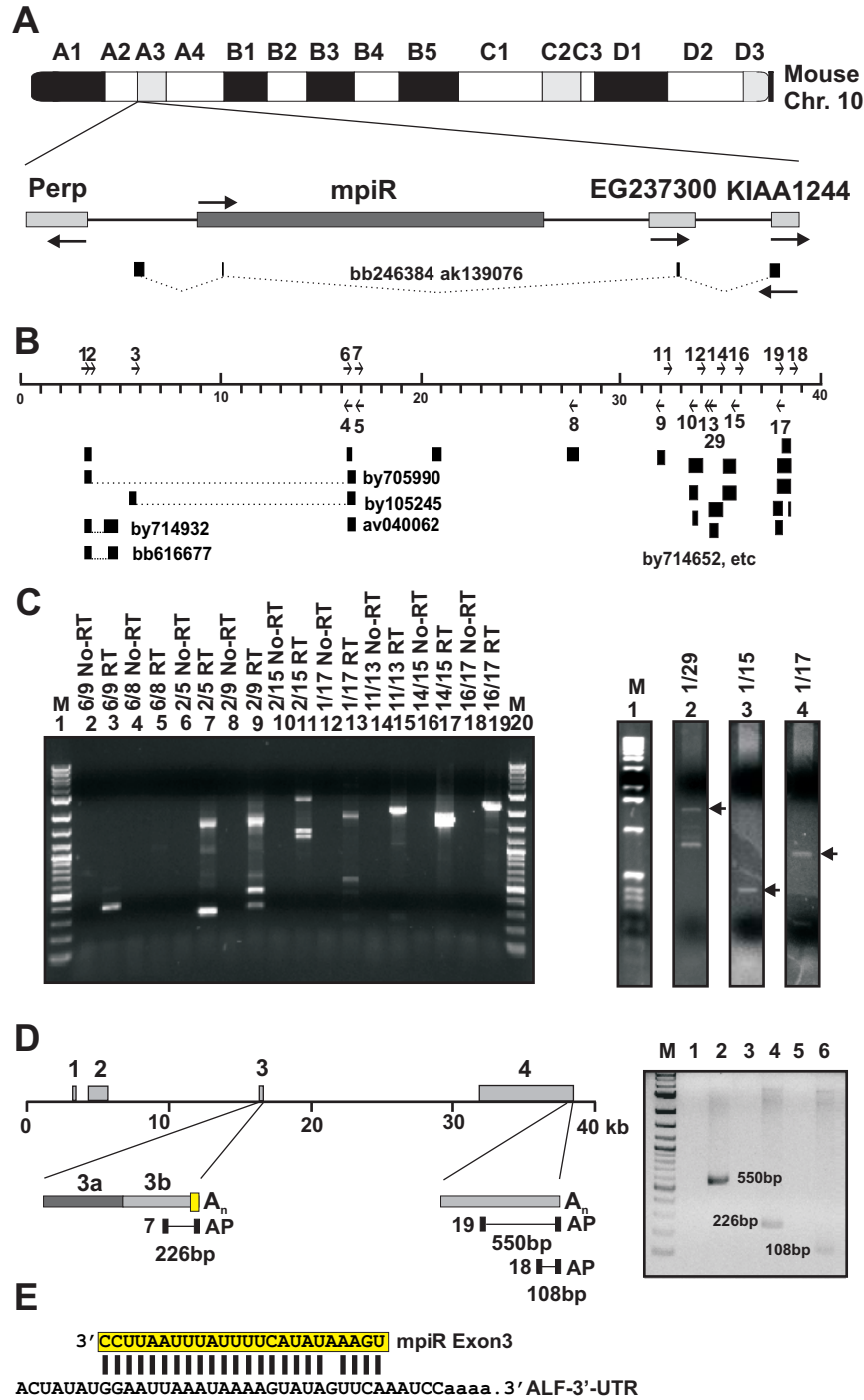


Figure S1. A. Chromosomal location 10A2/3 and the local environment surrounding the *mpiR* gene includes flanking *Perp* and *KIAA1244* genes, as well as a hypothetical protein *EG237300*. The *mpiR* gene is transcribed in the direction opposite to *Perp*. A partial opposite strand EST is shown which spans a large genomic region including the *KIAA1244*, *EG237300*, and *mpiR* genes. **B.** ESTs derived from *mpiR* are shown as black boxes and are in some cases labeled. Dashed lines between the boxes indicate sequence connections. The locations of primers used for RT-PCR analysis are shown by arrows at the top or bottom of the 40 kb genomic region. Fragmented ESTs are notably present beneath Exon 4. **C.** RT-PCR products obtained with various primer sets using testis RNA (left) and liver RNA (right) under RT- and no-RT reaction conditions. **D.** 3'-RACE analysis was used to identify 3'-ends of *mpiR* transcripts. One end corresponds to the longer version of Exon3 present in transcripts v.1 and v.2. The other end defines Exon 4. RT-PCR RACE results are shown in the gel to the right (lanes 2, 4, and 6). Lanes 1, 3, and 5 are no-RT control reactions. Primers used are designated 7, 18, 19, and AP. **E.** Yellow boxed sequences show those complementary to the 3'-UTR of ALF mRNA.

Supplementary Material Figure 2

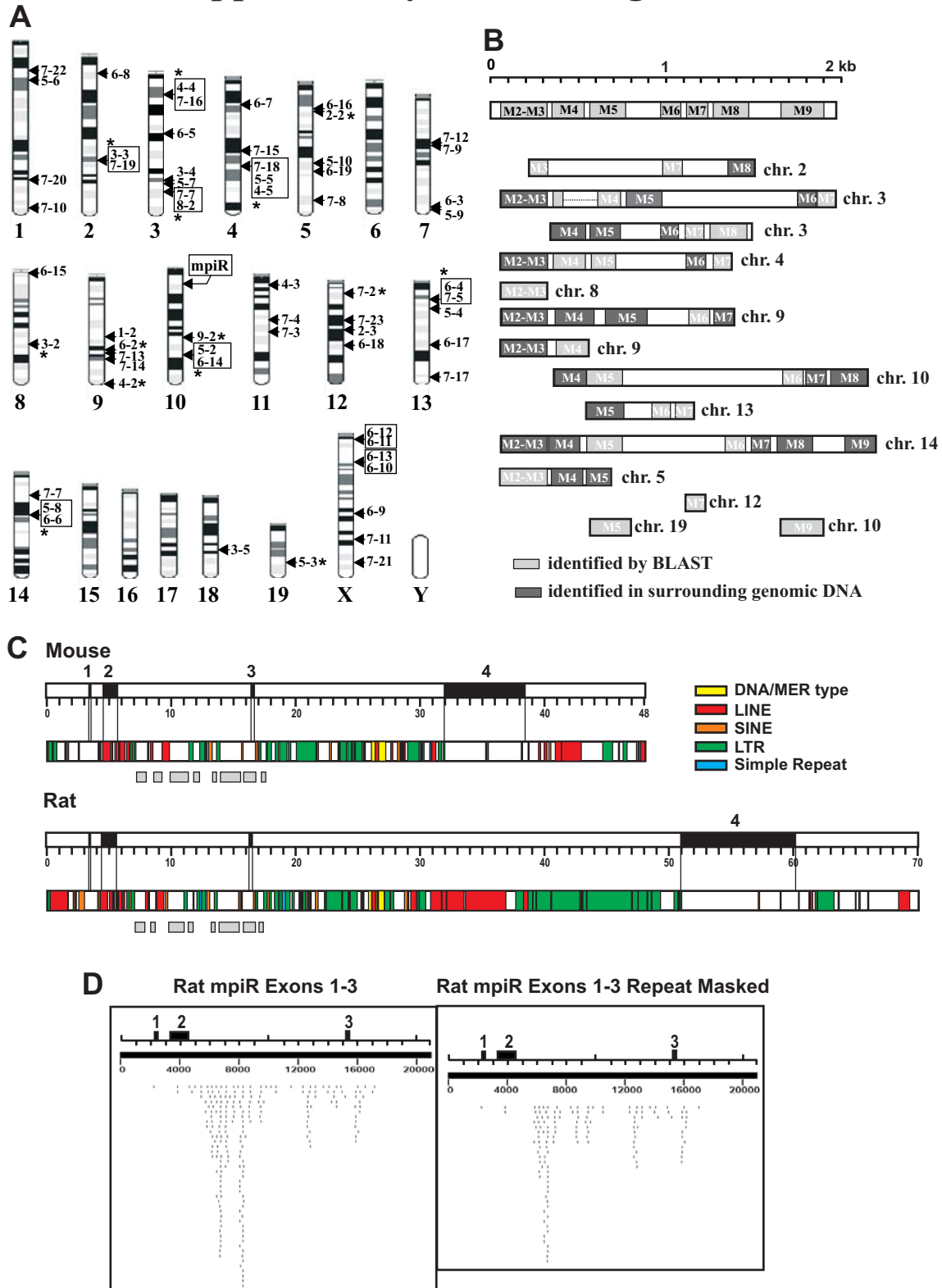


Figure S2. Analysis of the *mpiR* Exon4 multicopy repeats. **A.** The locations of identified homologous units that could be unambiguously identified by BLAST search were mapped onto mouse chromosomes. Boxed loci are those identified by BLAST to consist of multiple elements. **B.** Visual curation and pairwise computer analysis of selected repeats (shown by asterisks in A) show that many were in fact clustered modules similar in their organization to those identified in Exon4, whereas a few appeared to be solitary copies. Units identified by BLAST or by visual inspection are indicated by differences in shading. Spacing between elements varies from case to case. **C.** Comparison of repetitive element distribution in the mouse and rat *mpiR* genes. Gray boxes show region of conserved organization. **D.** piRNA matches to rat *mpiR* exons 1 to 3 shows that many are still observed even in a repeat masked sequence.

Supplementary Material Figure 4

```

M chr 17 -----GG.CTTCCATGTAGAAAAGT.GG..TC.G.AG.CCA..AAAG.GA.TATGAATG.A.A...AAGTCACCTCAGGAACCTT-.ATGC..G
M lin-28 -----GG.CTTCCATGTAGAAAAGT.GG..TC.G.A..CCA..AAAG.GA.TATGAATG.A.A...AGTCACCTCAGGAACCTT-.ATGC..G
H lin-28 -----GG.CTTCCATGTAGAAAAGT-GG..TC.G.AG.CCA..AAAG.GAATATGAATG.A.A...AAGTCACCTCAGGAACCTT..ATGC..G
M chr 7-1 -----G..TTCCATGTAGAAAAGT-GG..TC.G.AG.CCA..AAA..GAATATGAATG.A.A...AAG.CACTCAGGA.CTT-.ATGC..G
M chr 7-2 -----GG..TTCCATGTAGAAAAGT-GG..TC.G.AG.CCA..AAA..GAATATGAATG.A.A...AAG.CACTCAGGA.CTT-.ATGC..G
H chr 1 .....GG.CTTCCATGTAGAAAAGT-GG..TC.G.AG.CCA..AAAG.GAATATGAATG.A.A...AAGTCACCTCAGGAACCTT..ATGC..G
H chr 2 -----GG.CTTCCATGTAGAAA.T-G...TC.G.AG.CCA..AAAG.GAATATGAATG.A.A...AAGTCACCTCAGGAACCTT..AT.C...
H chr 5 -----GG.CTTC..T.TAGAAAAGT-GG..TC.G..G.CCA..AAAG.GAATATGAATG.A.A...AAGTCACCT.AG..ACTT..ATGC..G
M mpiR -----GG.CTTCCAT.TAGAA...-GG.....G.AG--...AAAG.GAA-----ATG-A.A...A.GTCACT.A.GAACCTT...T.C..G

M chr 17 .GC--A..AA..TTATGTCAAAG.GGCCA.AAGA..GTT--...-AGGAGAC..-A..AATGTA.CT.CATGTTTACTGCTAGAAA.CA
M lin-28 .GC--A..AA..TTATGTCAAAG.GGCCACAAGA.TGTT--A...-AGGAGAC..-AC.AATGTA.CTCCATGTTTACTGCTAGAAA.CA
H lin-28 .GC-.A..AA..TTATGTCAAAGTGGCCACAAGA.TGTT..A...-AGGAGAC..-AC.AATGTA.CTCCATGTTTACTGCTA.AAA.CA
M chr 7-1 .GC--A..AA..TTA.GT.AAAGTGGCCACAAGA.TGTT--A...-AGGAGAC..-AC.AATGTA...CCATGTTTACTGCTAGAA..CA
M chr 7-2 .GC--A..AA..T.ATGTCAAAGTGGCCACAAGA.TGTT--A...-A.GAGAC..-AC.AATGTA...CCATGTTTACTGCTAGAA..CA
H chr 1 .GC-.A..AA..TTATGTCAAAGTGGCCACAAGA.TGTT..A...-AGGAGAC..-AC.AATGTA.CTCCATGTTTACTGCTA.AAA.CA
H chr 2 .-----TTATGTCAAAGTGGCCACAAGA.TGTT..A...-AGGAGAC..-A..AAT.TA.CTC.AT.TTACTGCTAGAAA.CA
H chr 5 .GC--A..AA..TTATGTCAAAGTGGCCACAAGA.TGTT..A...-AGGAGAC..-AC.AATGTA.CTCCATGTTTACTGCTAGAAA.CA
M mpiR .....AA..TTATGT.AA-GTG.C.A.A..A.TGTT..A...A.GA.AC...AC.AATGT..CT.CATGTT..CTGC.AGAAA.CA

M chr 17 AAGCTTT.TG..AAATCTTGAAATTTATGGGG.GG---...-GAAAAG..TGT.A.T---GTCTGT..TTT.C.-GATC...TTCC..
M lin-28 AAGCTTT.TG..AAATCTTGAAATTTATGGGG.GG---...-T..GAAAAG..TGT.A.T---GTCTGT..TTT.C.-.ATC...TTCC..
H lin-28 AAGCTTT.TG..AAATCTTGAAATTTATGGGG.GG.....-T..GAAAAG..TGT.A.T---GTCTGT..TTT.C.TGATC...TTCC..
M chr 7-1 AAGCTTT.TG..AAATCTTGAAATTTATG.....-T..GAAAAG..TGT.A.T---GTCTGT..TTT.C.TGAT...TTCC..
M chr 7-2 AAGCTTT.TG..AAATCTTG.ATTTATG.....-T..GAAAAG..TGT.A.T---GTCTGT..TTT.C.TGAT...TTCC..
H chr 1 AAGCTTT.TG..AAATCTTGAAATTTATGGGG.GG.....-T..GAAAAG..TGT.A.T---GTCTGT..TTT.C.TGATC...TTCC..
H chr 2 AAGCTTT.TG..AAATCTTGAAATTTATGGGG.GG.....-T..GAAAAG..TGT.A.T---GTCTGT..TTT.C.TGATC...TTCC..
H chr 5 AAG.TTT.TG..AAATCTTGAAATTTATGGGG.GG.....-...G.-AG..TGT.A.....-G.CTGT..TTT...TGATC...CC..
M mpiR AA.CTTT.TG..A.ATCTTG.A.TT..ATGGGG.GG.....-T..G..AG...TA..T...GT.TGT..T.T.C.TGATC...TC...

M chr 17 .A.T.C.G.A...A.G...C.....GCTT.G.T.A.CC.CA.C...G...G.GTGTTT.ATTT..TGTTGA.TTT.....
M lin-28 .A.T.C.G.A...A.G...C.....GCTT.G.T.A.CC.CA.C...G...G.GTGTTT.ATTT..T.GTTGATTT.....
H lin-28 .A.T.C.G.A...A.G...C.....GCTT.G.T.A.CC.CA.C...G...G.GTGTTT.ATTT.ATGTTT.GATTTT.....
M chr 7-1 .A.T.C.G.A...A.G...C-----T-----G.GTGTTT.ATTT.ATGTTT.ATTT.....
M chr 7-2 .A.T.C.G.A...A.G...C-----T-----G..TGTTT.ATTT.ATG.TT.ATTT.....
H chr 1 .A.T.C.G.A...A.G...C.....GCTT.G.T.A.CC.CA.C...G...G.GTGTTT.ATTT.ATGTTT.GATTTT--...
H chr 2 .A.T.C.G.A...A.G...C.....GCTT.G.T.A.CC.CA.C...G...G.GTGTTT.ATTT.ATGTTT.GATTTT-----
H chr 5 .A.T.C.G.A...A.G.....GCTT.G.T.A.CC.CA.C...G...G.GTGTTT.ATTT.ATGTTT.GATTTT-----
M mpiR .....A...A-----GCTT.G.T.A.CC.CA.C...G-----TGTTT-AT.T.AT..TTGATT.T-----

M chr 17 ...C.GG...TT...TTCC.ATT.T.TAATCA-T...T.ACACA..CTGACT.T---CCTTC.CTT.TC--TTCCCT..G.A...A.A
M lin-28 ...C.GG...CTT...TTCC.ATT.T.TAATCA-T...T.ACACA.GCTGACT.T---CCTTCCCTT.TC--TTCCCT..GAAAATA.A
H lin-28 ...C.GG..ACTT...TTCC.ATT.T.TAATCA-T.T..TAACACA.GCTGACT.T--CCTTCCCTT.TC..TTCCCT..GAAAATA.A
M chr 7-1 ...C.....ACTT...TCC.ATT.T.TAATCA-T.T..TA...CA.G.TGACT.T---CCTTCC.TT.....GAAAATA.A
M chr 7-2 ...C.....ACTT...TCC.ATT.T.TAATCA-T.T..TA...CA.G.TG.CT.T---CCTTCCCTT.TC--TTCCCT..GAAAATA.A
H chr 1 ...C.GG..ACTT...TTCC.ATT.T.TAATCA-T.T..TAACACA.GCTGACT.T--CCTTCCCTT.TC..TTCCCT..GAAAATA.A
H chr 2 ...C.GG..A.TT...TTCC.ATT.T.TA.TCA.T.T..TAACACA.GCTGAC.T--C.TT.CCTT.TC..TTCCCT..GAAAATA
H chr 5 ...C.GG..ACT...TTCC.ATT.T.TAA.CA-T.T..TAACACA.GCTGACT.T...CCTTCCCTT.TC..TTCCCT..GAAAAT..A
M mpiR .....GG..ACTT...T-CC--TT.-.T..T.--T.T..TAACACA..CTGA..T-----TTCC.TT.TC..TT.CCT..GAA.ATA.A

```

Figure S4. Alignment of a region within Intron 2 of the mouse *mpiR* gene with similar sequences from the mouse and human *lin-28* genes (M *lin-28*, H *lin-28*), as well as additional selected examples from different mouse and human chromosomes (M chr 17, M chr 7-1, M chr 7-2, H chr 1, H chr 2, H chr 5).