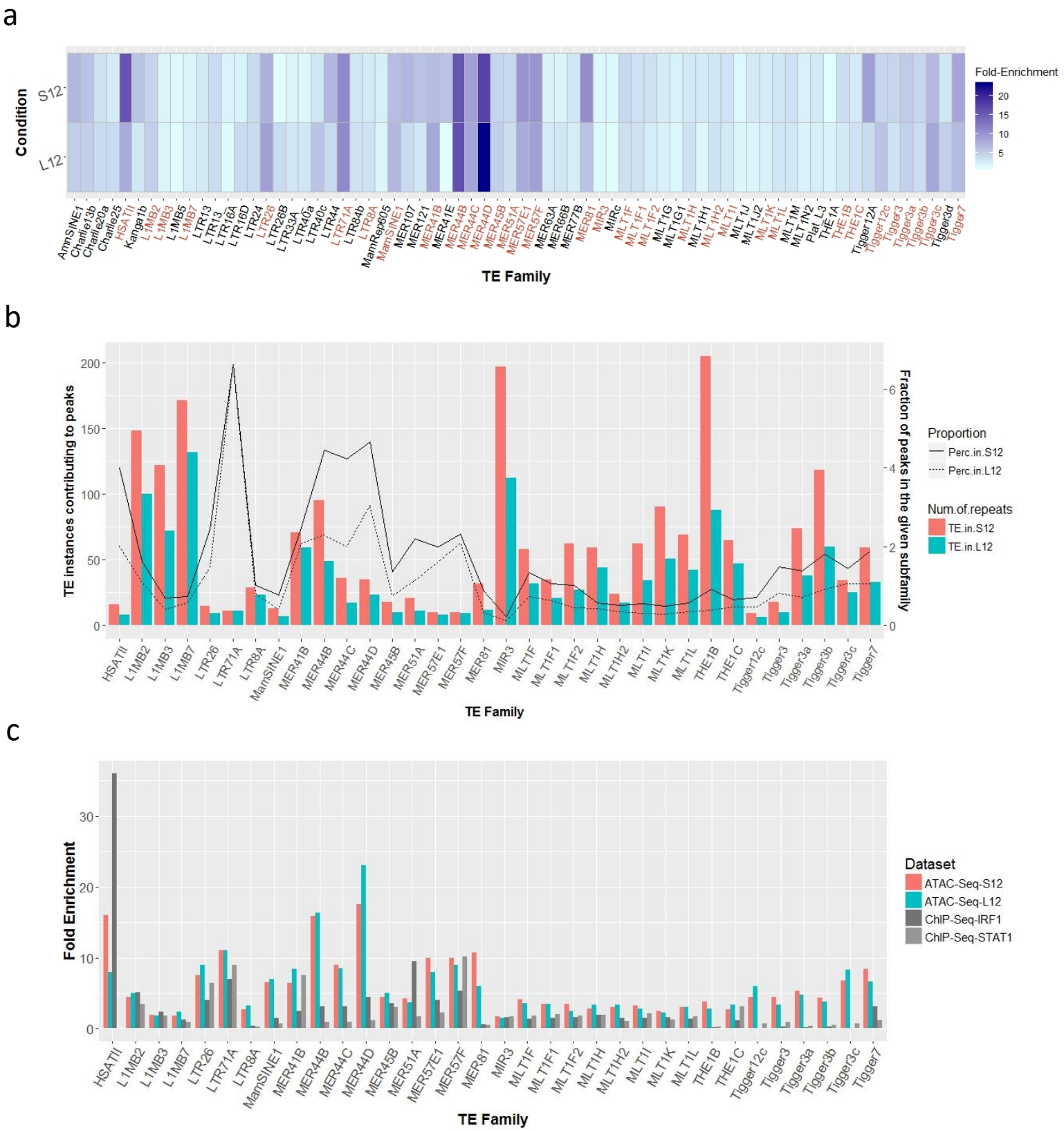


Figure S1. (a) TE enrichment in accessible chromatin, comparison between S12 and L12 conditions. Darker colour indicates greater fold-enrichment. All subfamilies with q-value < 0.05 shown. Red colour indicates subfamilies which passed the significance threshold in both condition (the 34 “immune" subfamilies). (b) Absolute number of TE instances (bars) and fraction of all instances in each subfamily (lines) contributing to accessible chromatin. All 34 immune TE subfamilies shown. (d) Comparison of TE enrichment results in our ATAC-Seq peaks and the ChIP-Seq datasets published by Chuong *et al.* All 34 immune TE subfamilies shown.

a

b

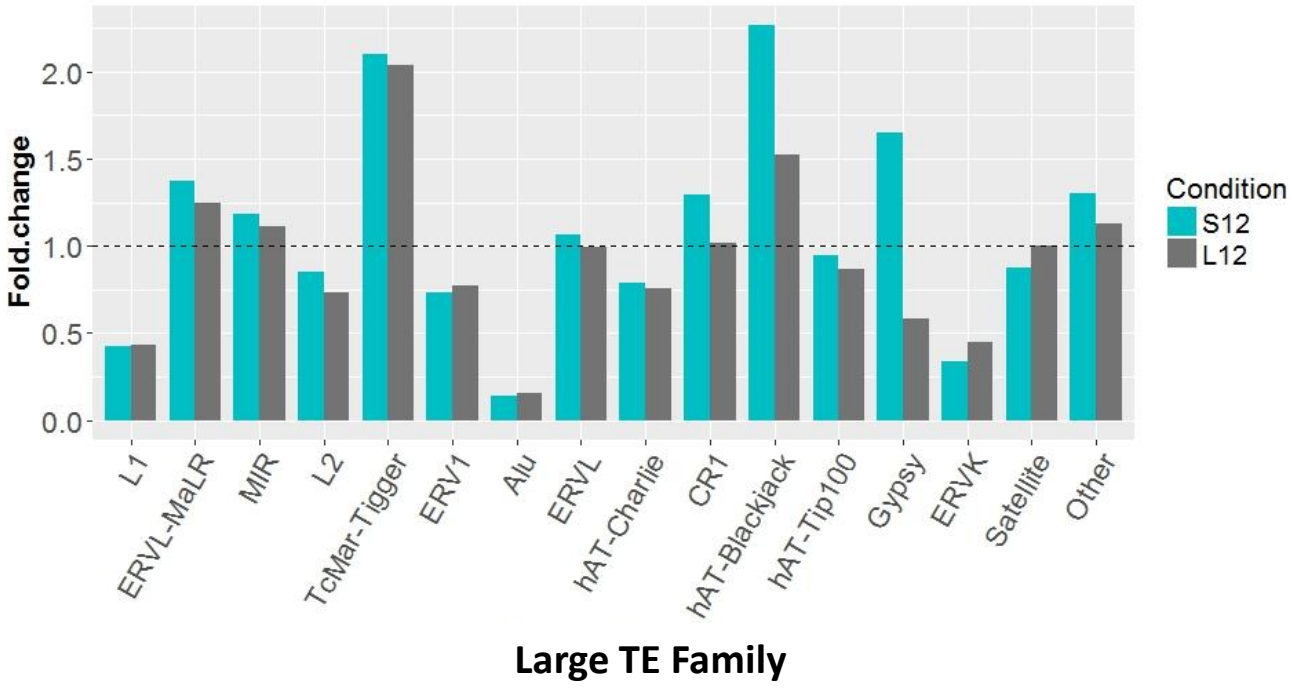
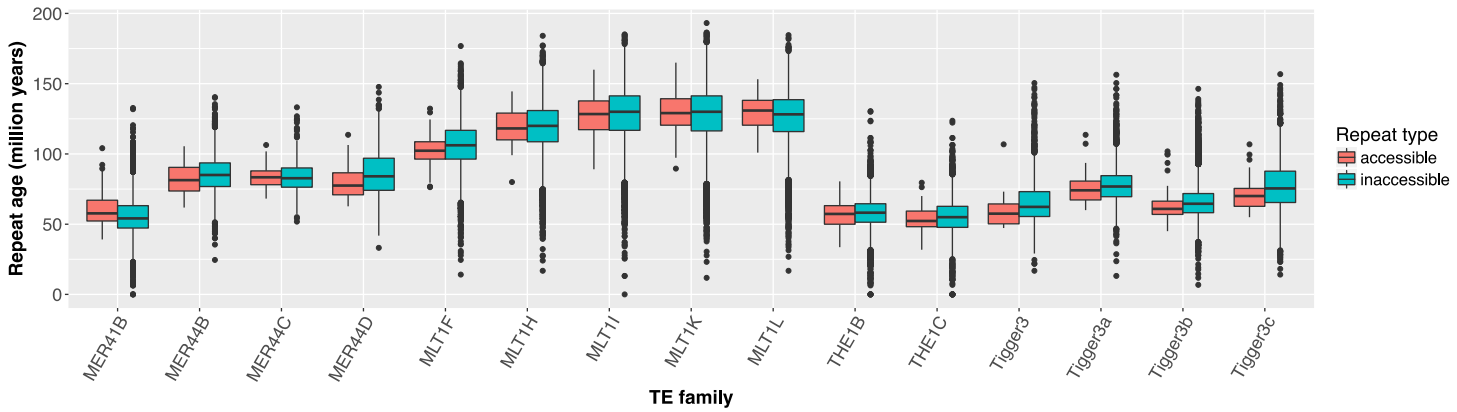


Figure S2. Enrichment of (a) the 4 major TE classes and (b) the large TE families/categories in accessible chromatin (expected/observed number of TEs in ATAC-Seq peaks). Note that the most enriched large TE family is TcMar-Tigger, a category of DNA transposons which includes immune TE subfamilies MER44B, MER44C, MER44D, Tigger3, Tigger3a, Tigger3b, Tigger3c, Tigger7 and Tigger12c.

a



b

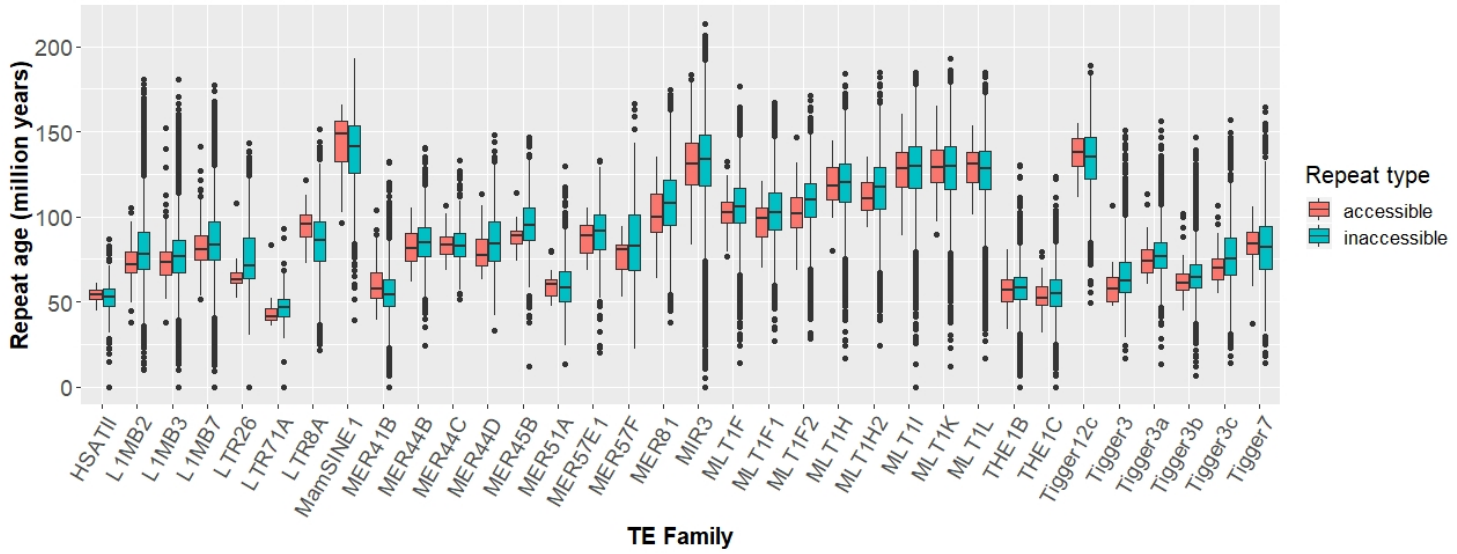


Figure S3. Age of TE instances in accessible and inaccessible chromatin for each repeat subfamily with (a) the set of 14 related subfamilies and (b) the full set of 34 immune subfamilies. We find several subfamilies which show age differences between the two groups. For example, the accessible instances are younger on average for L1MB2 (p = 3.7e-9), Tigger3b (p = 6.6e-5), MLT1F2 (p = 9.3e-5), Tigger3c (p = 6.4e-3), MLT1F (p = 0.017), MLT1F1 (p = 0.027), MER44B (p =0.028), Tigger3 (p = 0.028), MER44D (p = 0.032) and THE1B (p = 0.035). Conversely, they are older for LTR8A (p = 1.2e-4) and MER41B (p = 3.6e-3).


AP-1(bZIP)


BATF(bZIP)


Atf4(bZIP)


Atf3(bZIP)


Fos12(bZIP)


Fra1(bZIP)


Jun-AP1(bZIP)


MafA(bZIP)


MafK(bZIP)


Bach2(bZIP)


EWS:ERG-fusion(ETS)


NFAT(RHD)


IRF1(IRF)


IRF2(IRF)


IRF4(IRF)


ISRE(IRF)


PRDM1(Zf)


CHOP(bZIP)


CEBP(bZIP)


GATA3(Zf)


ELF5(ETS)


GSC(Homeobox)


SCL(bHLH)


Ap4(bHLH)

NFkB-p65(RHD)

NFkB-p65-Rel(RHD)

NFY(CCAAT)

Bcl6(Zf)

STAT1(Stat)

STAT6(Stat)

Foxh1(Forkhead)

Figure S4. Top 30 enriched motifs (found in over 1/3 of accessible instances in at least one repeat subfamily and showed at least a 20% increase compared with inaccessible repeats). The STAT1 motif is also included.

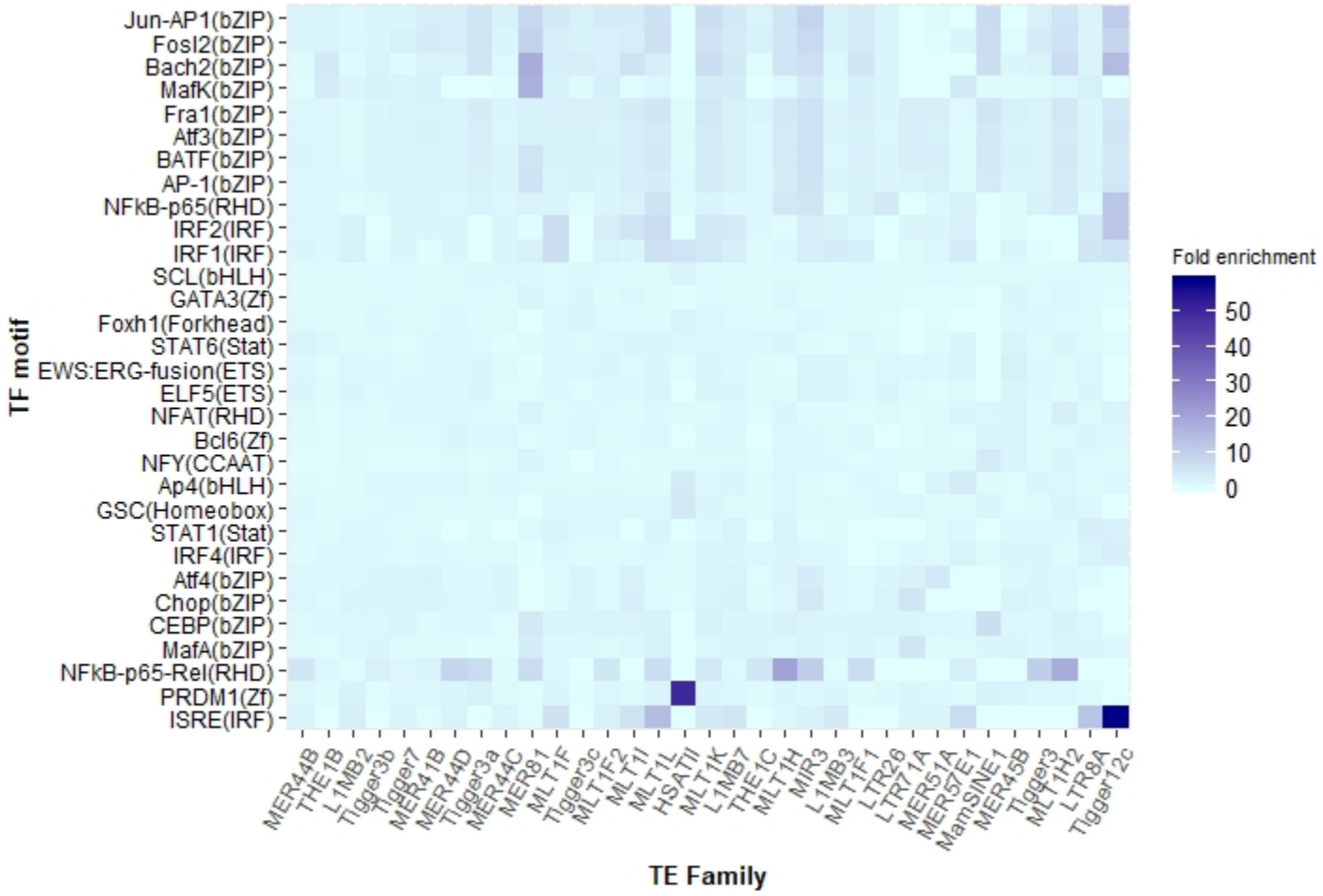


Figure S5. Fold-enrichment for top 30 enriched motifs (found in over 1/3 of accessible instances in at least one repeat subfamily and showed at least a 20% increase compared with inaccessible repeats) + the STAT1 motif.

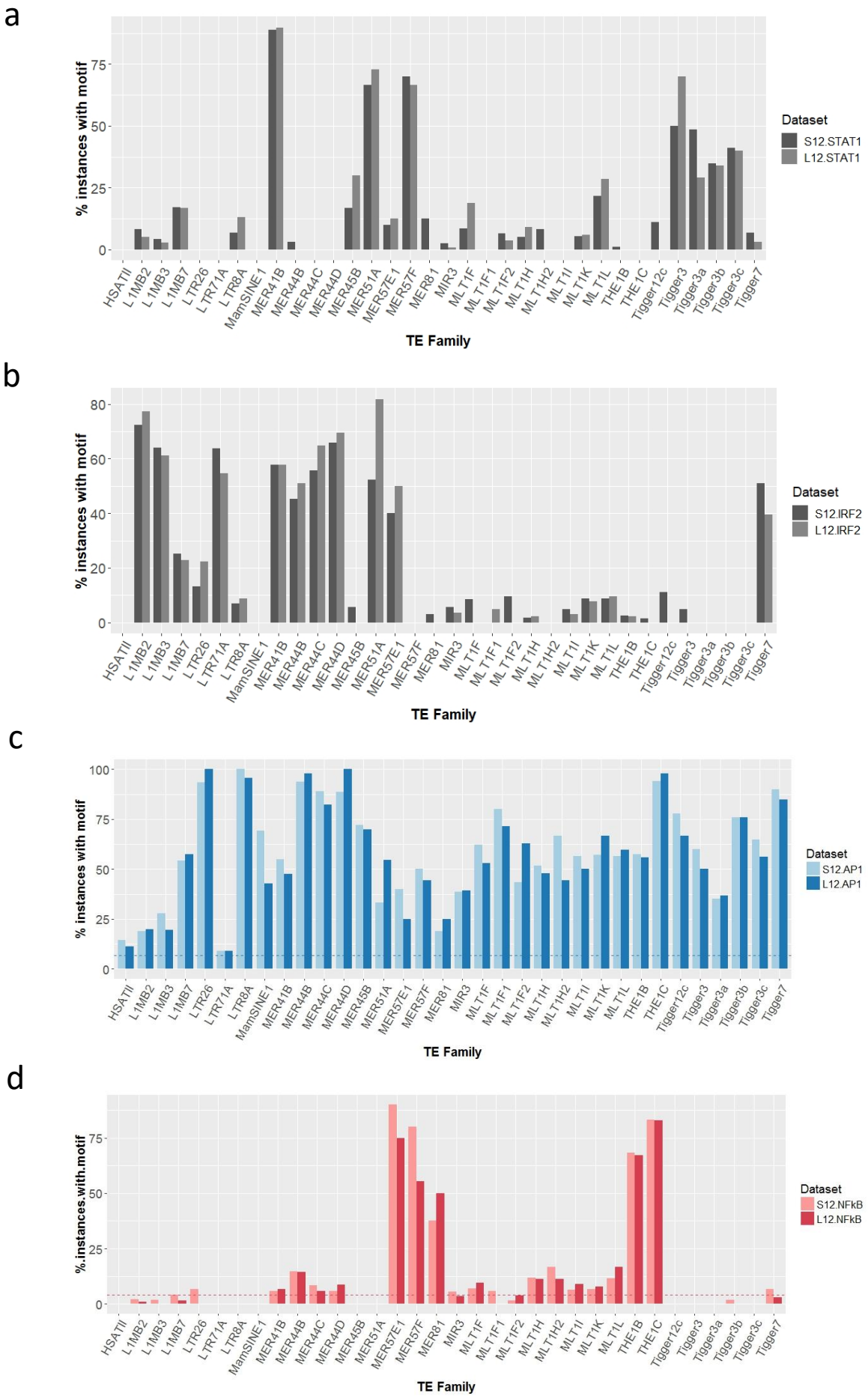


Figure S6. Proportion of accessible repeats with the STAT-1 (a), IRF2 (b), AP-1 (c) and NF-kB (d) motifs in all 34 immune subfamilies. We include here IRF2 rather than IRF1 as a representation of the IRF motifs in PARs, as the two motifs are very similar and IRF2 is the most enriched in PARs.

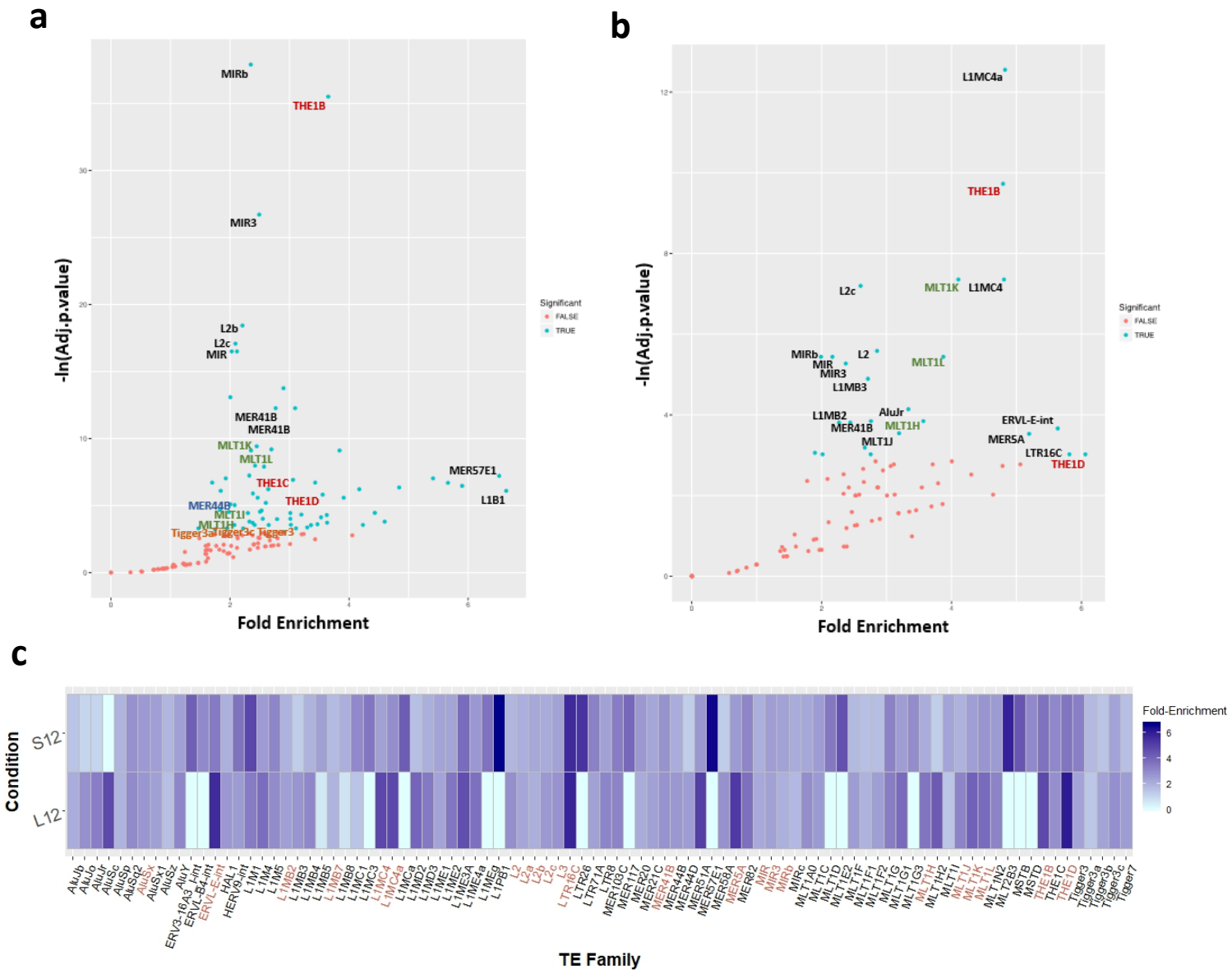
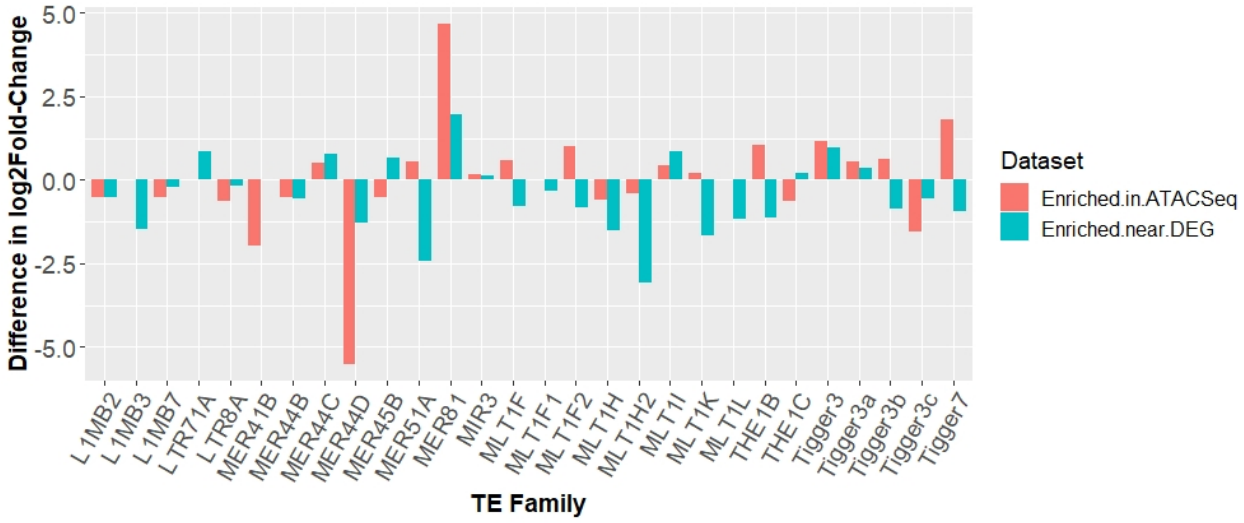
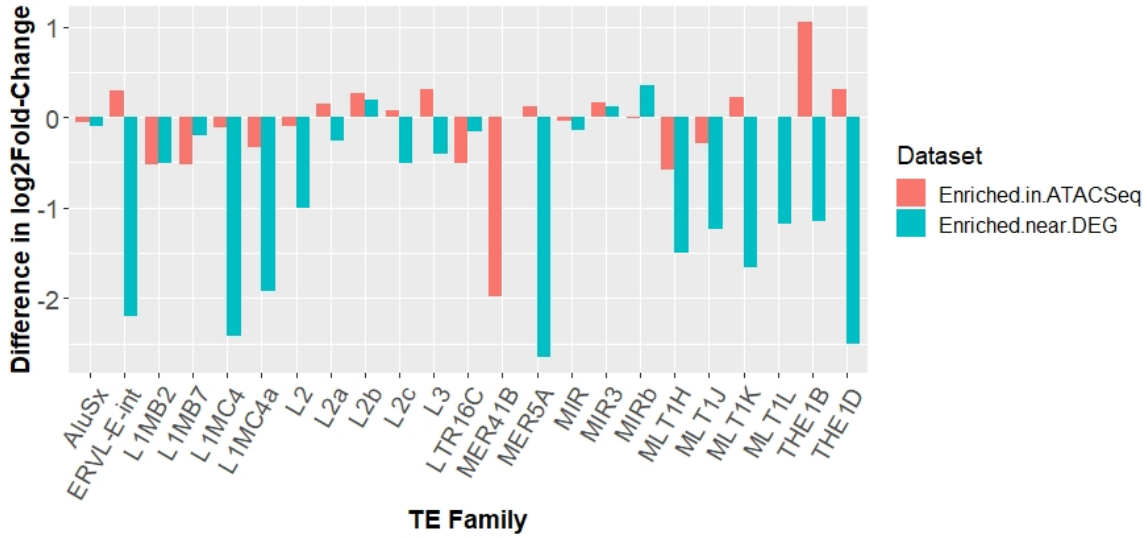


Figure S7. (a-b) Enrichment of PARs near differentially expressed genes in S12 and L12 samples. (c) Comparison of fold-enrichment between S12 and L12. Darker colour indicates greater fold-enrichment. All subfamilies with q-value < 0.05 are shown. Red colour indicates subfamilies significant in both conditions.

a



b



c

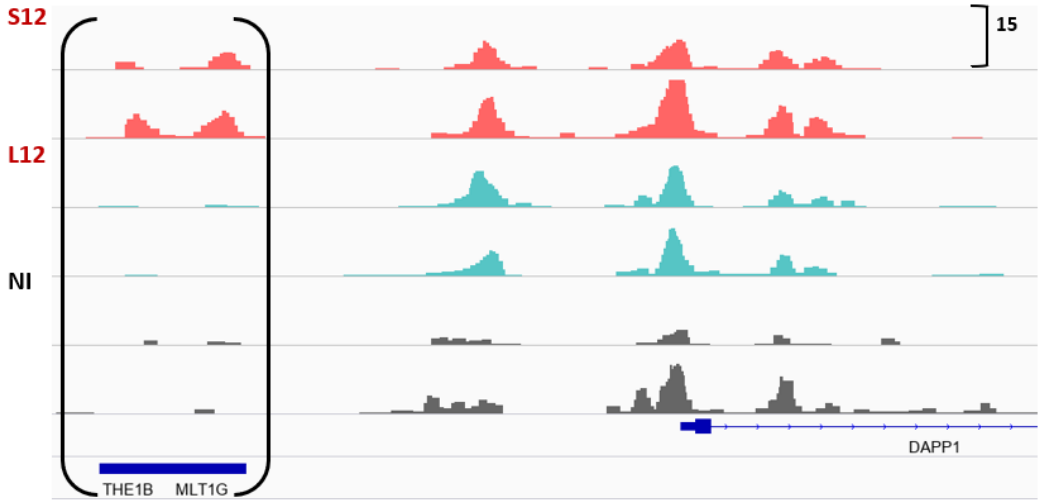


Figure S8. (a) Difference in fold-enrichment between S12 and L12 for the enrichment in accessible chromatin (orange) and enrichment near DEG (turquoise). Positive differences indicate increased enrichment for S12 and negative indicate increased for L12. Many families show similar differences in both analyses but the difference is small: L1MB2, MER44B, MER44C, MER44C, MER81, MLT1H, MLT1H2, MLT1I, Tigger3, Tigger3a, Tigger3c. (b) Similar as in (a), but differences in fold-enrichment are shown for the 23 subfamilies enriched near DEGs in both S12 and L12. Note the greater enrichment in L12. (c) Example of THE1B and MLT1G instances which are only accessible in the S12 samples and found 2.7 kb upstream of the DAPP1 gene, which is differentially expressed only in S12 (log2-fold-change 3.26 and 0.63 for S12 and L12, respectively).

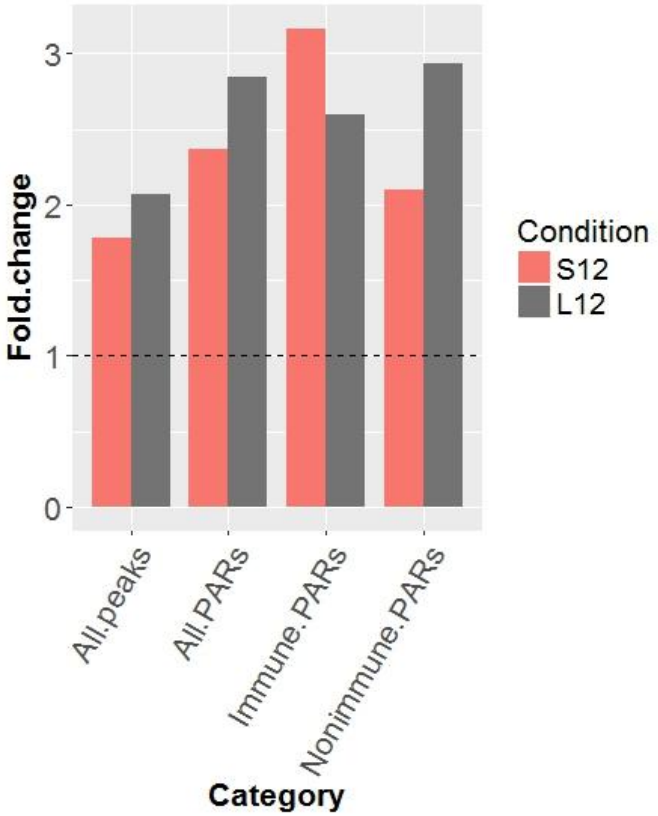


Figure S9. Enrichment of all ATAC-Seq peaks, PARs, immune PARs and non-immune PARs in QTL-regions. “Immune PARs” are defined as belonging to the 34 “immune” subfamilies enriched in accessible chromatin. “Non-immune PARs” are accessible repeat instances belonging to all the remaining subfamilies which have not passed the significance threshold. Number of peaks contributing to QTL-regions in each category, for S12 and L12 conditions respectively: 645 and 224 ATAC-Seq peaks; 170 and 57 PARs; 57 and 13 immune PARs; 111 and 44 non-immune PARs.

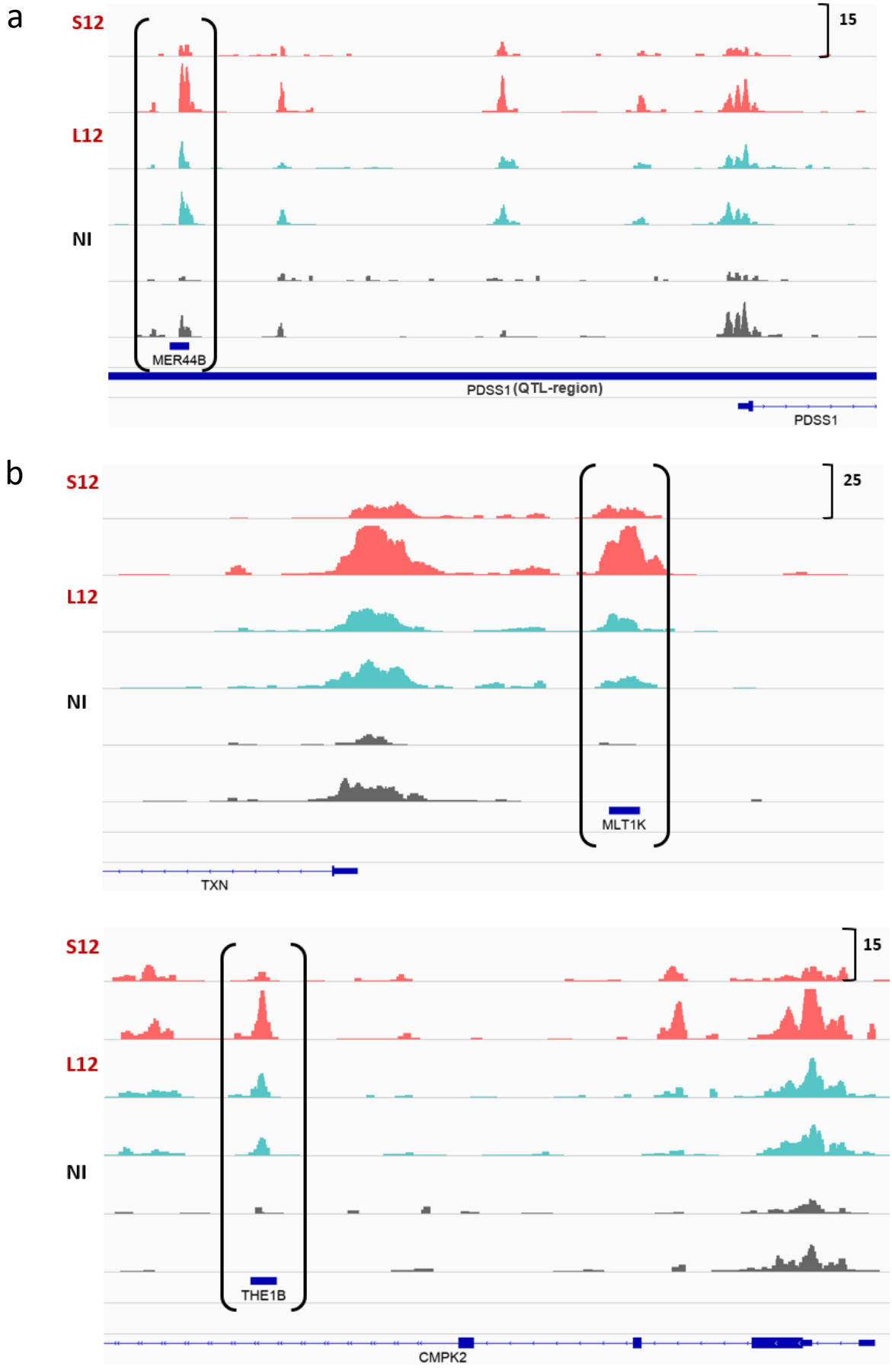


Figure S10. (a) Example of accessible PAR overlapping its corresponding QTL-region near a DEG. In this case, the MER44B instance is 15 kb upstream of the PDSS1 gene (fold-change for differential expression: 3.4 and 3.1 for S12 and L12, respectively). (b) MLT1K instance 2 kb upstream of the TXN gene (fold-change 4.0 and 3.0). (c) THE1B instance 8 kb downstream of CMPK2 transcription start site (fold-change 6.8 and 3.2). Track height is indicated at the top right and is equal across all tracks.