

Polymorphic B2 insertions shape gene expression programs in the house mice, *Mus musculus domesticus*



Laura Blanco-Berdugo and Beth Dumont
The Jackson Laboratory, Bar Harbor, Maine 04609 USA
Laura.Blanco-Berdugo@jax.org

Introduction

- Transposable elements (TEs) are DNA sequences that generate copies of themselves.
- Comprise ~50 % of mammalian genomes.
- TE insertion polymorphisms are an important source of genomic variation.
- TEs regulate gene expression and shape the genome landscape.
- Mammalian TE-derived sequences contain gene regulatory elements such as promoters, enhancers, and transcription factor binding sites[1].

Motivating Question

What is the Impact of TEs polymorphism on the regulatory landscape of *Mus musculus domesticus*?

Approach

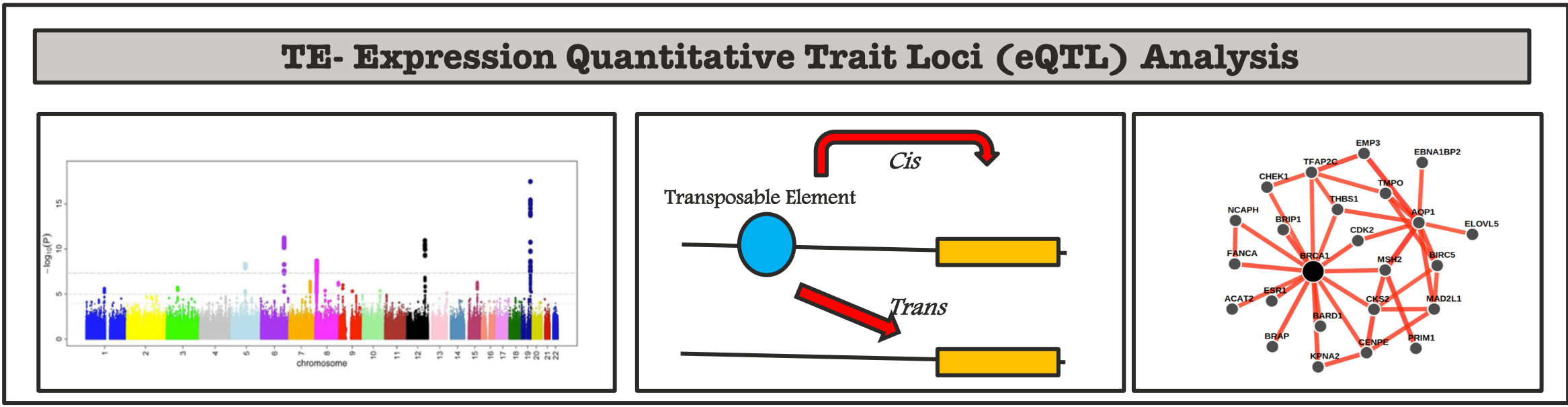
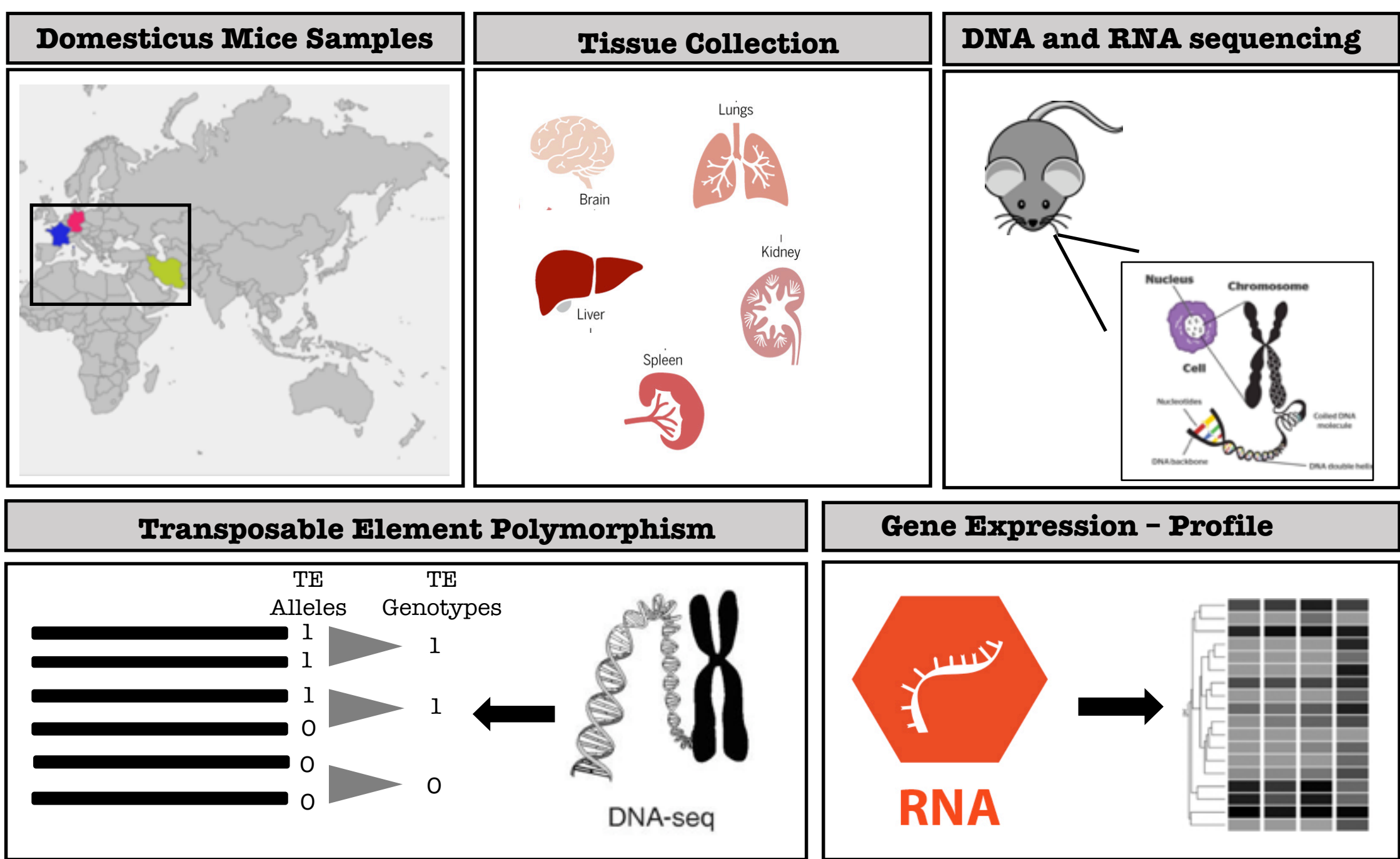


Figure 1. Polymorphic TE expression quantitative trait loci (eQTL) analysis scheme. Publicly available RNA-seq (8 tissues) and whole genome sequences were obtained for 24 wild *Mus musculus domesticus* from three populations (France, Germany and Iran) [2].

Results

Widespread polymorphic SINEs in wild house mice (*Mus musculus domesticus*)

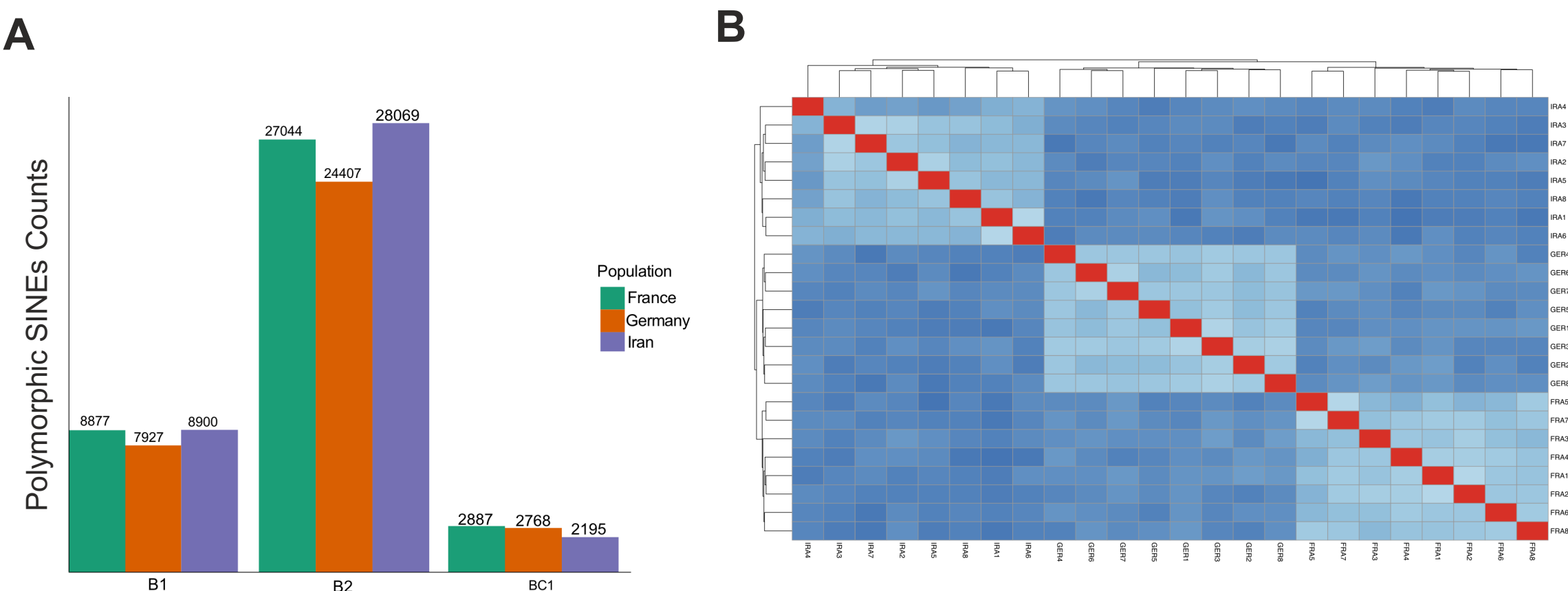


Figure 2. Distribution of polymorphic insertions for the most abundant SINEs in the *Mus musculus domesticus* genome. A) Discovery of polymorphic B1, B2 and BC1 SINEs in three different populations. B2 is the most abundant SINE in the mouse genome. Polymorphic insertions were identified using MELT [3]. (B) Clustering-based relationships among samples derived from the polymorphic B2 (poly-B2) presence/absence matrix.

Approximately 10% of polymorphic SINE B2 elements reside in annotated regulatory regions

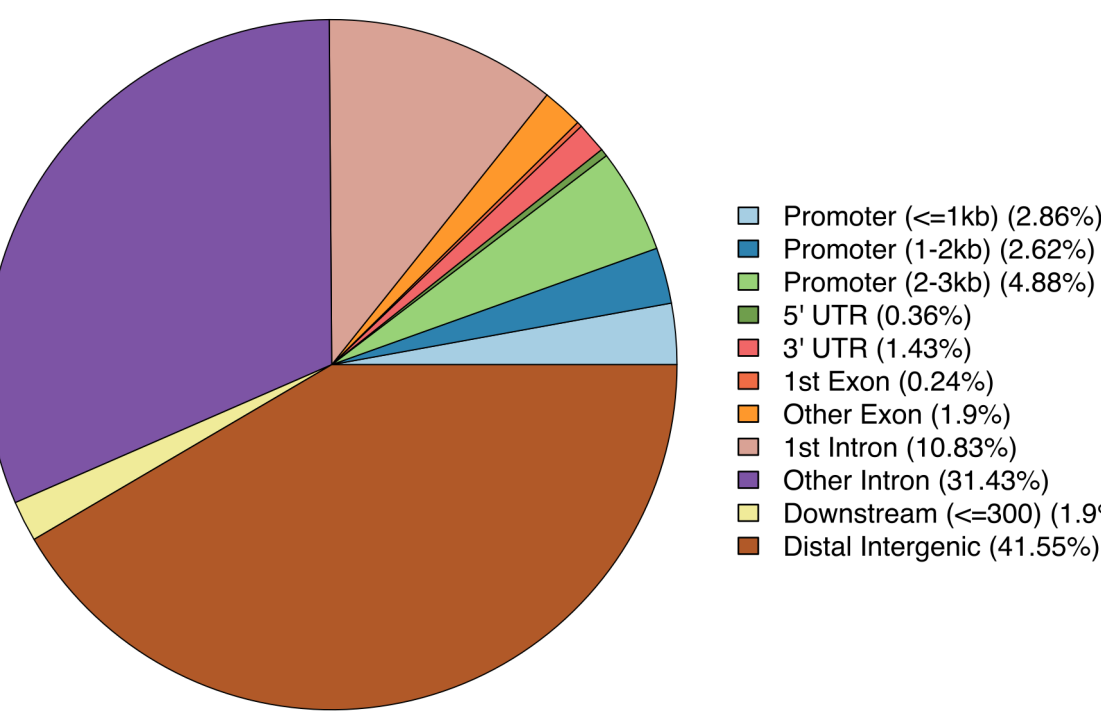


Figure 3. Genomic distribution of poly-B2s in the *M. m. domesticus* reference genome assembly (mm10).

Most poly-B2 insertions are shared between populations

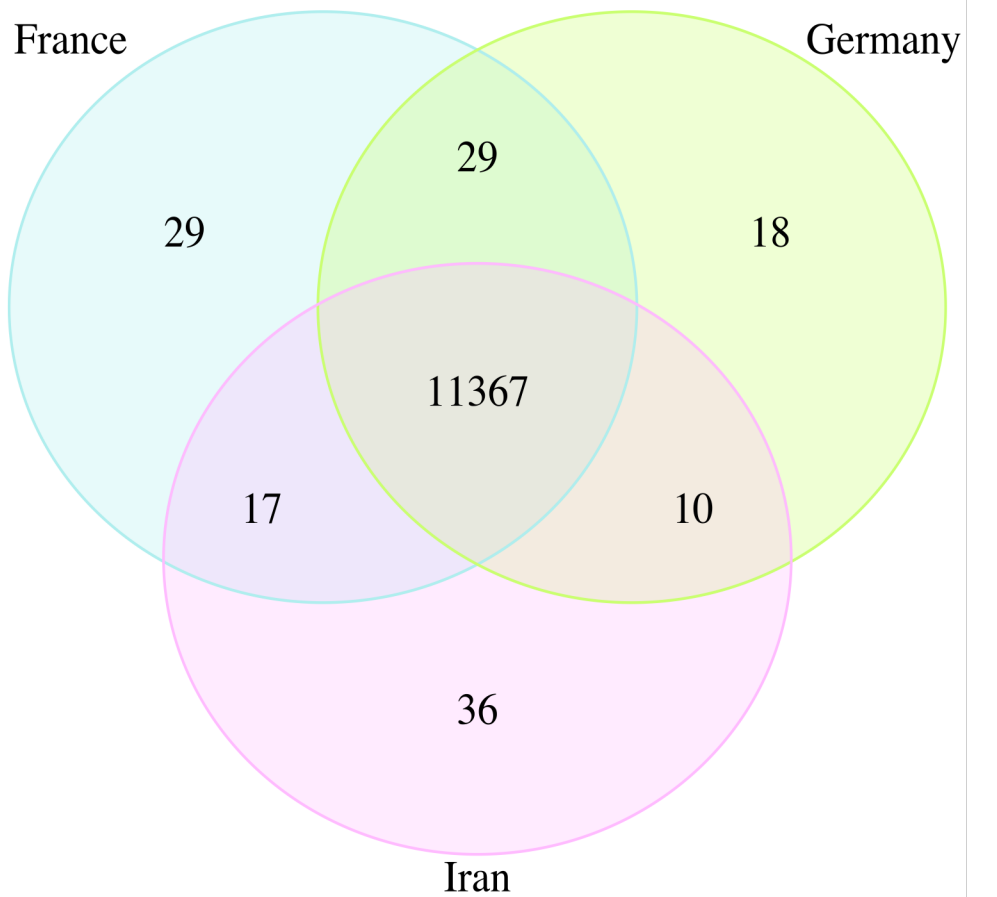


Figure 4. Venn diagram of population private and shared poly-B2 insertions.

Tissue of origin explains most variation in gene expression

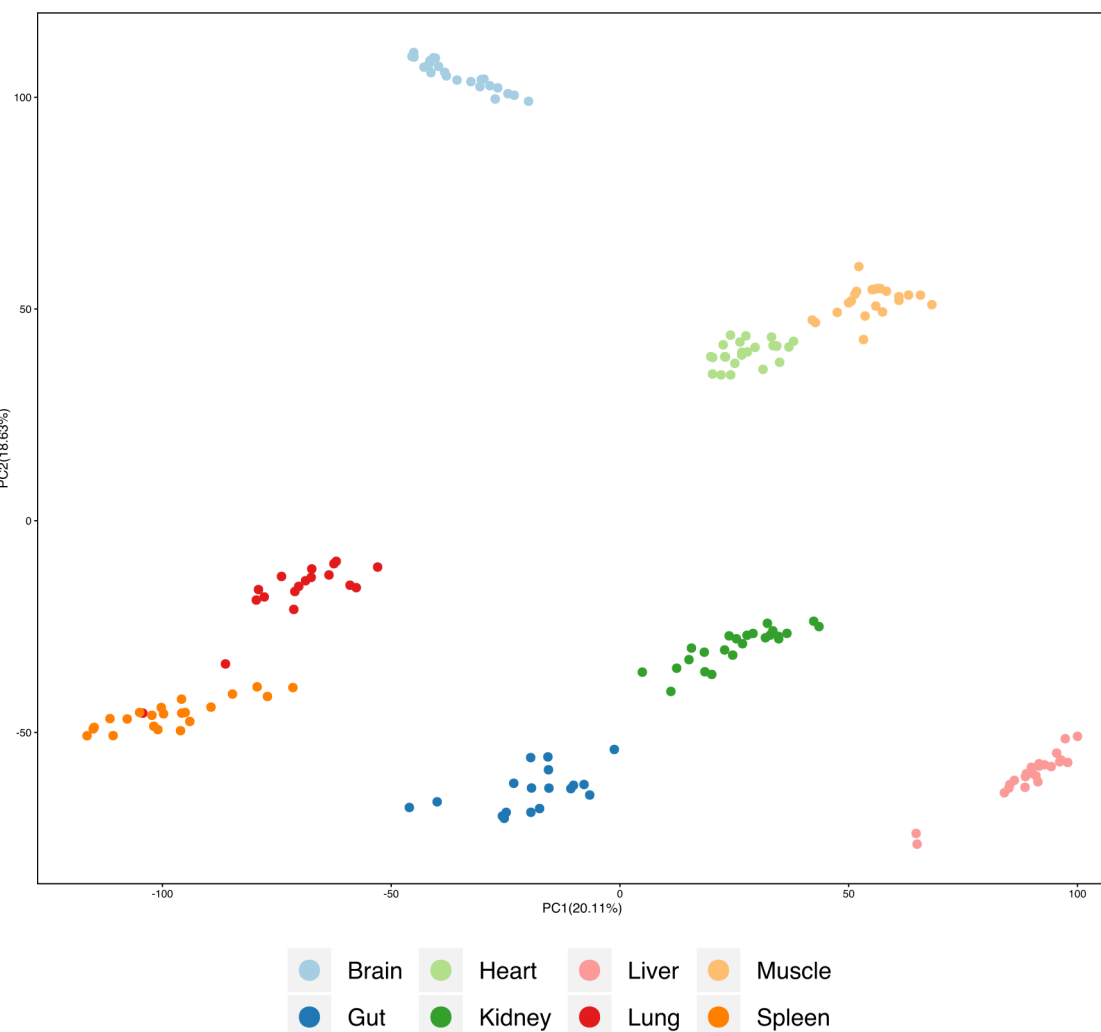


Figure 5. PCA analysis reveals that most variation in gene expression is attributable to tissue-specific effects, rather than genetic differences between populations or animals.

Significant population- and individual-level variation in the gene regulatory landscape

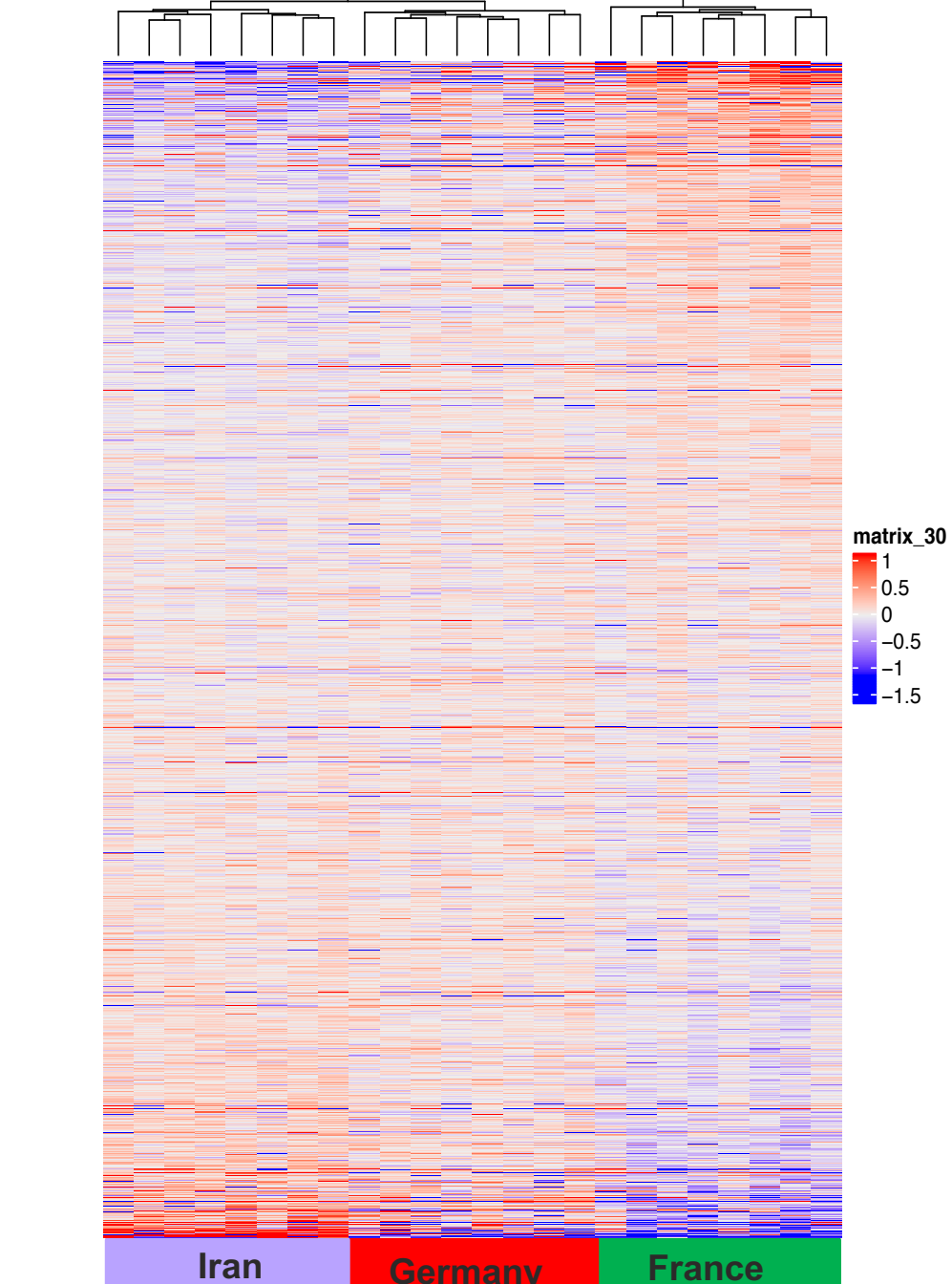


Figure 6. Heatmap displaying patterns of differential gene expression in the brain across samples from three *M. m. domesticus* populations.

eQTL analysis reveals thousands of poly-B2 elements linked with tissue specific gene regulatory programs in *trans*

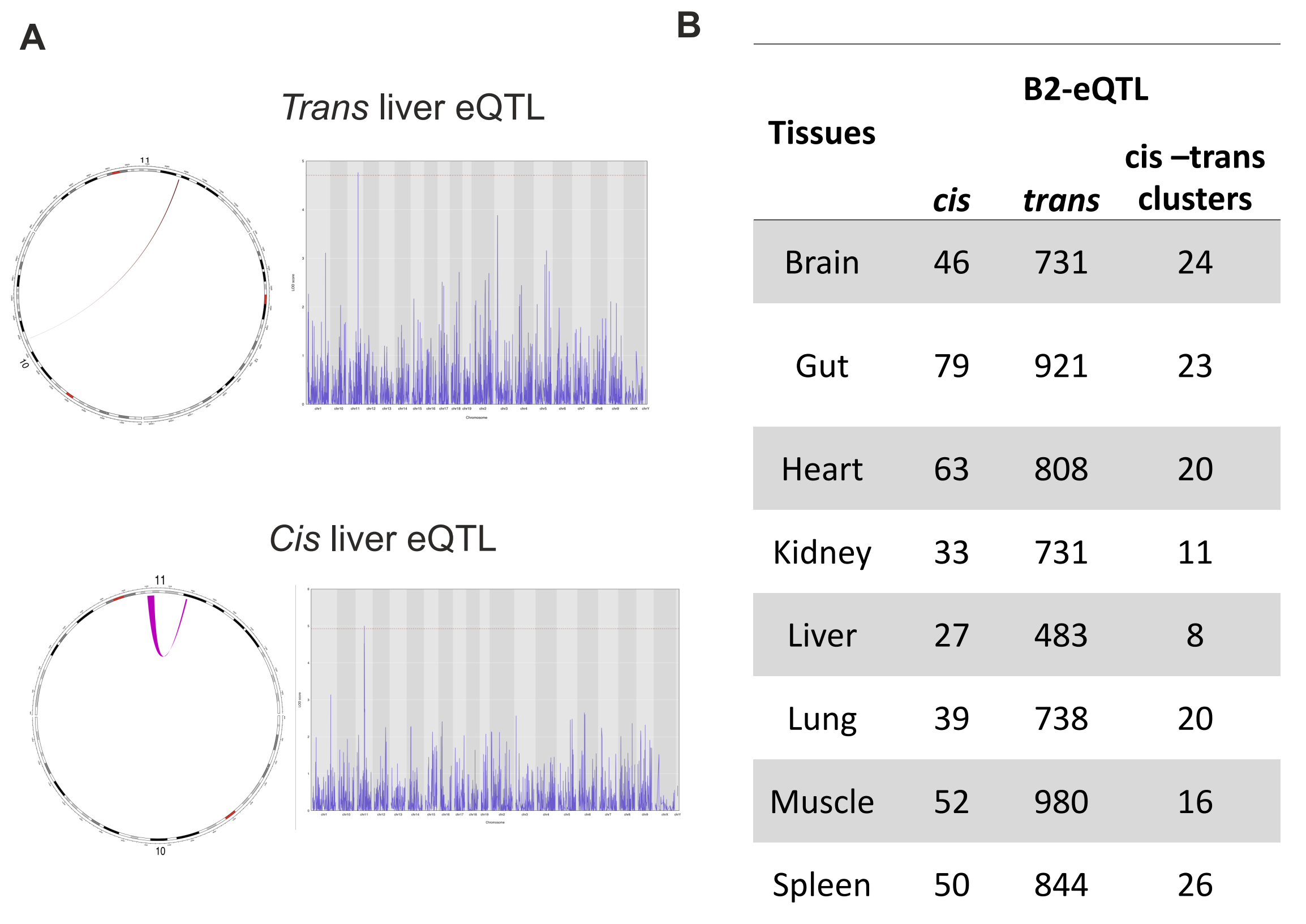


Figure 6. (A) Representative eQTL LOD plots for significant *cis* and *trans* representative. (B) The number of statistically significant (*p* value < 0.05) B2-eQTL for each tissue, broken down into *cis* and *trans* acting eQTL. Poly-B2 elements that exert both *cis* and *trans* effects on multiple genes are designated as "cis-trans clusters".

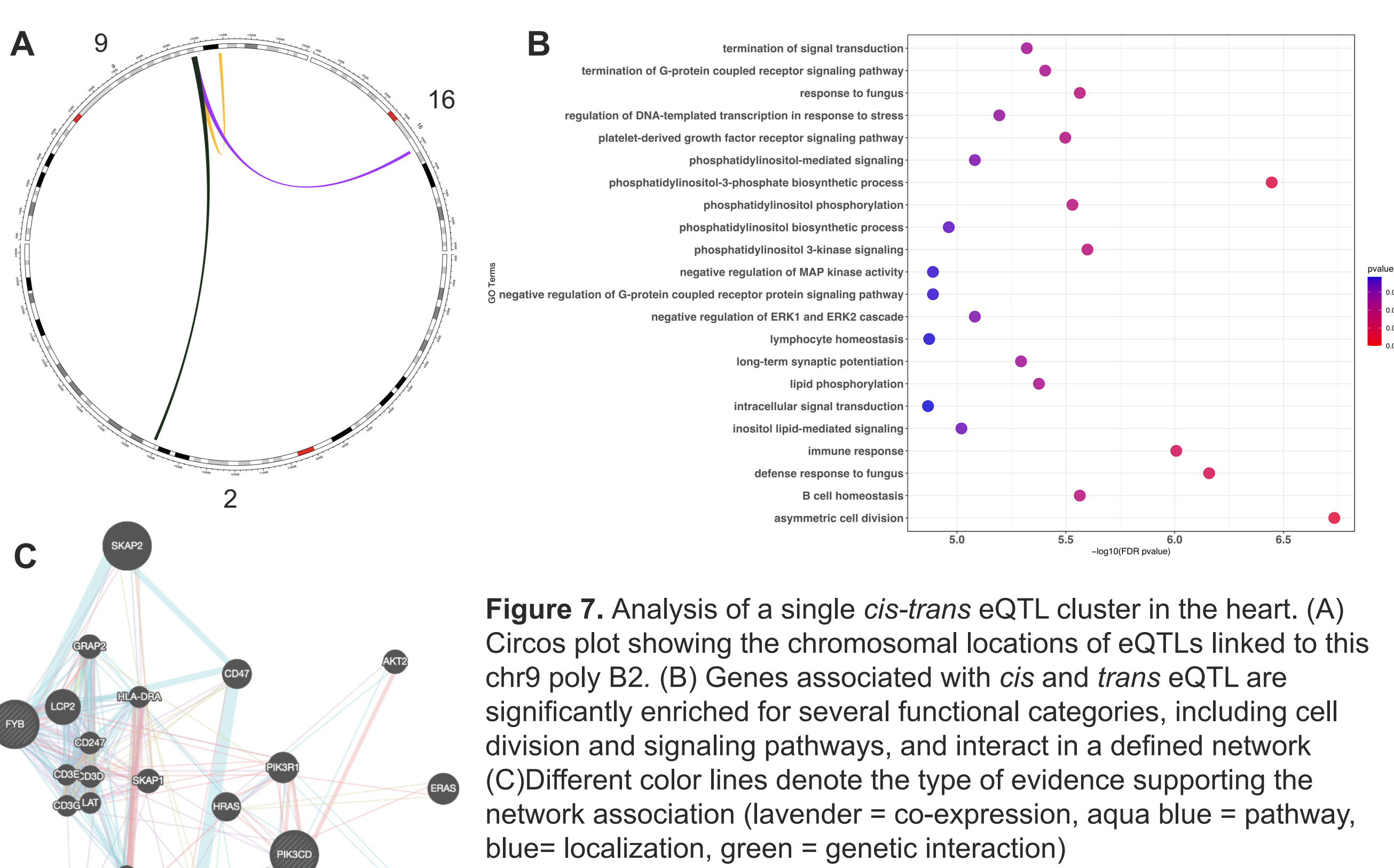


Figure 7. Analysis of a single *cis-trans* eQTL cluster in the heart. (A) Circos plot showing the chromosomal locations of eQTLs linked to this chr9 poly B2. (B) Genes associated with *cis* and *trans* eQTL are significantly enriched for several functional categories, including cell division and signaling pathways, and interact in a defined network (C) Different color lines denote the type of evidence supporting the network association (lavender = co-expression, aqua blue = pathway, blue = localization, green = genetic interaction)

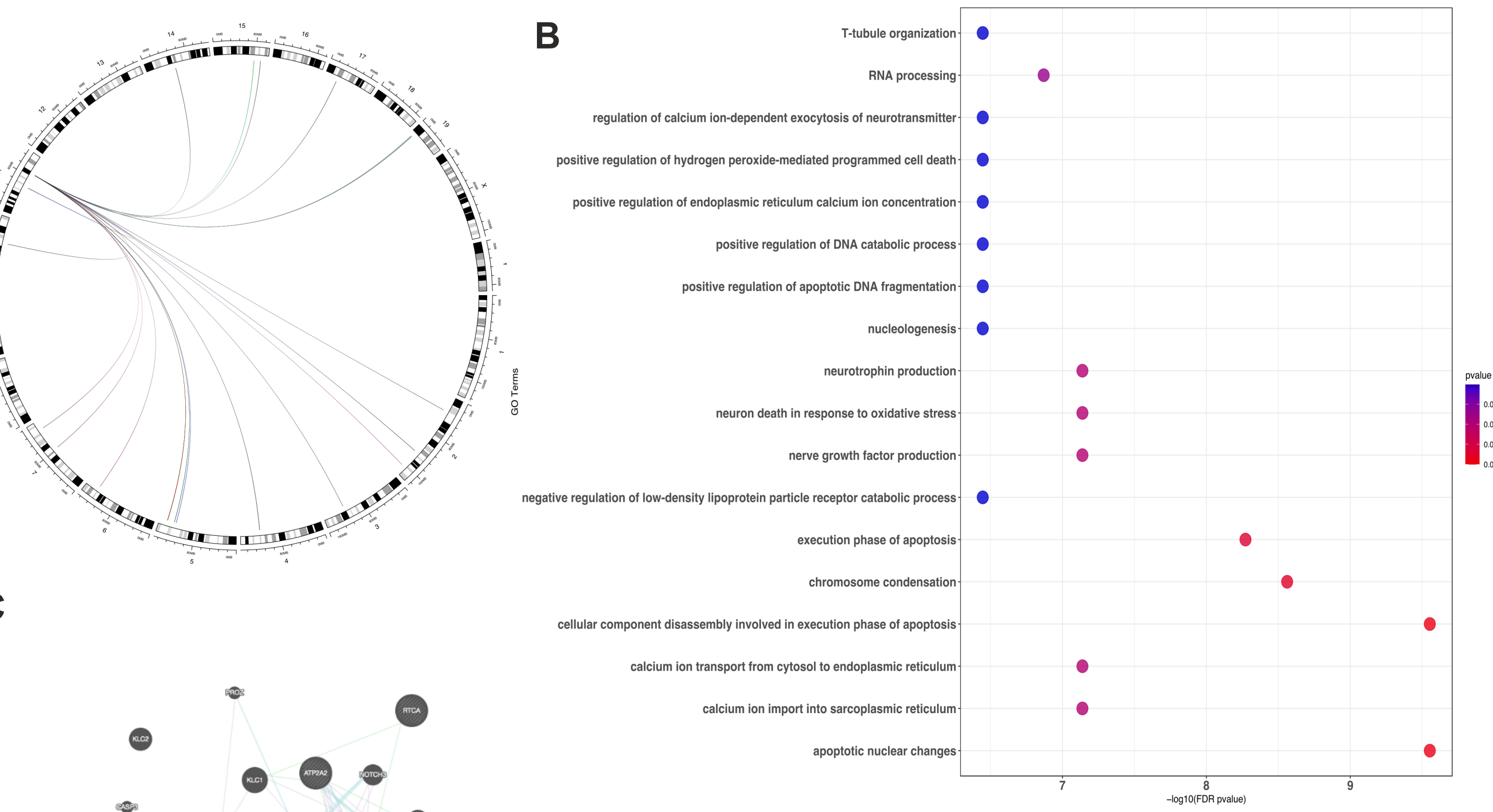


Figure 8. Analysis of a single *cis-trans* eQTL cluster in the lung (A) Circos plot showing the chromosomal locations of eQTLs linked to this chr9 poly B2. (B) Genes associated with *cis* and *trans* eQTL are significantly enriched for several functional categories, including cell death and calcium transport, and interact in a defined network (C) Different color lines denote the type of evidence supporting the network association (lavender = co-expression, aqua blue = pathway, blue = localization, green = genetic interaction)

Conclusions

- B2 SINE elements provide an important source of genetic variation in wild *M. m. domesticus* populations.
- Poly-B2s predominately shape the gene regulatory landscape through *trans* effects
- Poly-B2s exert major tissue-specific effects on gene networks, including genes with housekeeping functions.
- We hypothesize that poly-B2s with *cis* and *trans* effects may indirectly modulate gene expression by influence the regulatory of key transcription factors which in turn regulate multiple genes within a network.

Poly-B2s influence tissue-specific gene regulatory programs by modulating gene expression levels of key genes that perturb complex networks.

References

- Bourque, Guillaume. "Transposable elements in gene regulation and in the evolution of vertebrate genomes." *Current opinion in genetics & development* 19.6 (2009): 607-612.
- Hair, B., Karakoc, E., Neme, R. et al. Genomic resources for wild populations of the house mouse, *Mus musculus* and its close relative *Mus spretus*. *Sci Data* 3, 160075 (2016).
- Gardner, Eugene J., et al. "The Mobile Element Locator Tool (MELT): population-scale mobile element discovery and biology." *Genome research* 27.11 (2017): 1916-1929