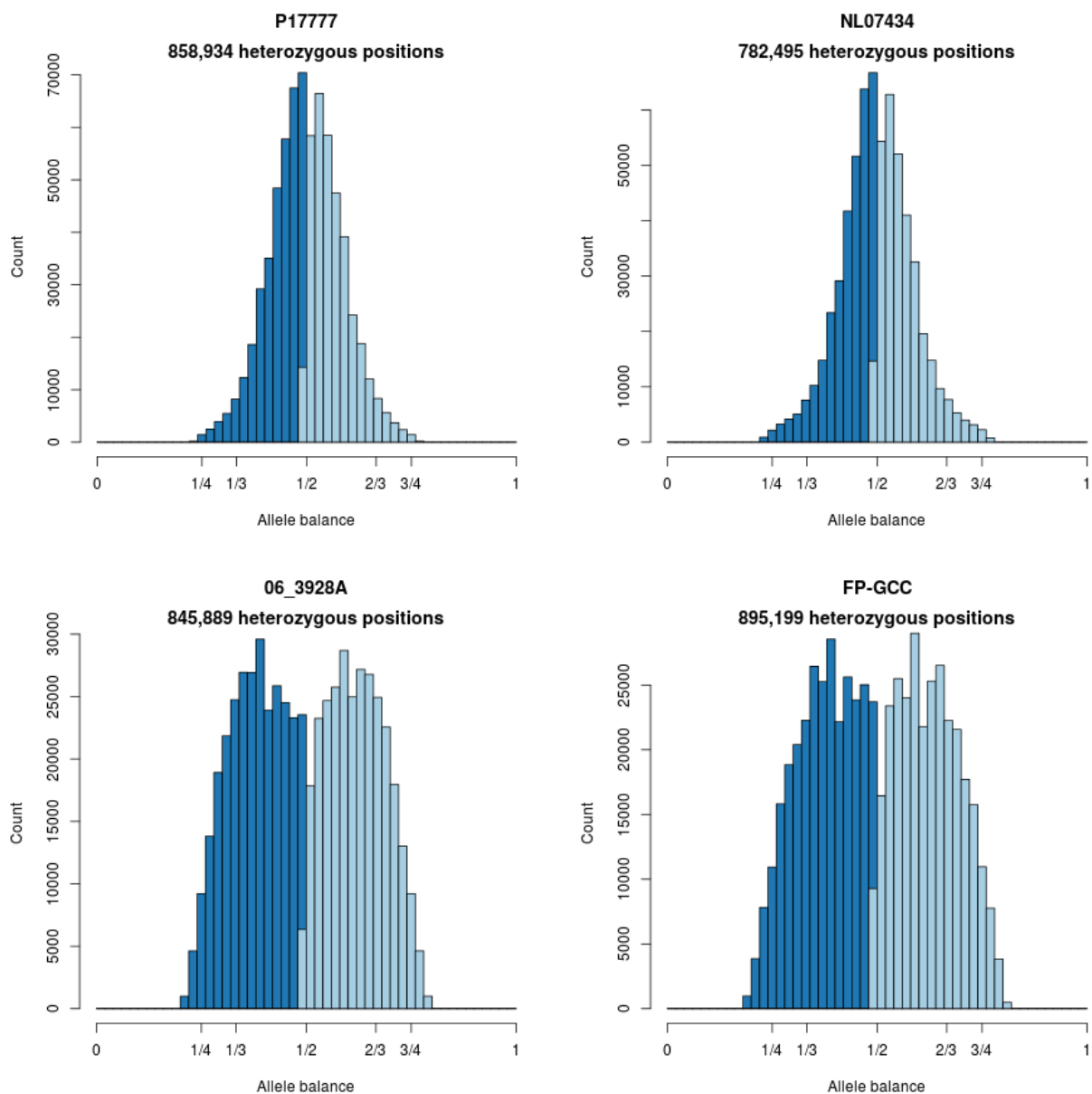
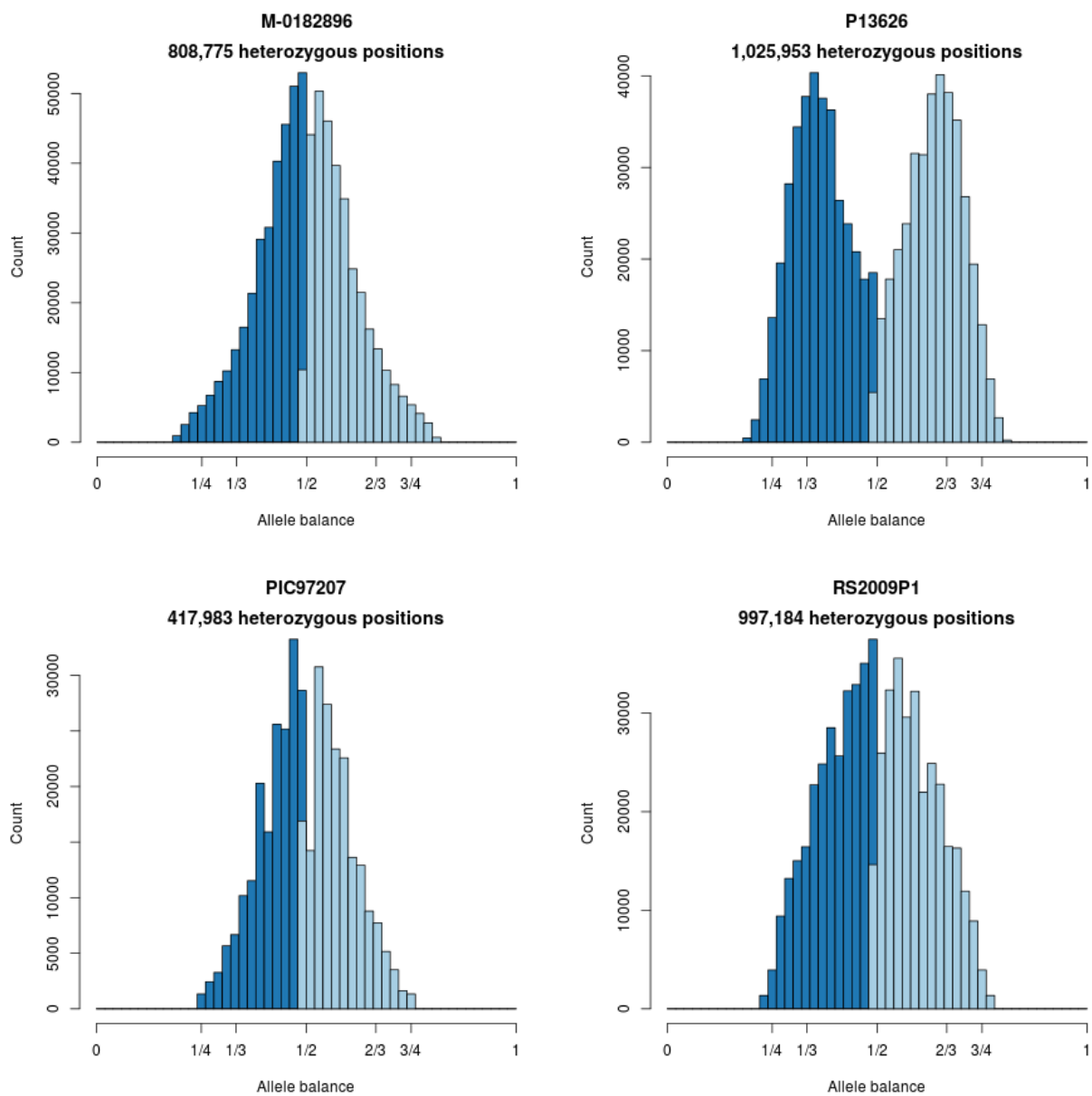


Text S6. Allele balance: *P. infestans*

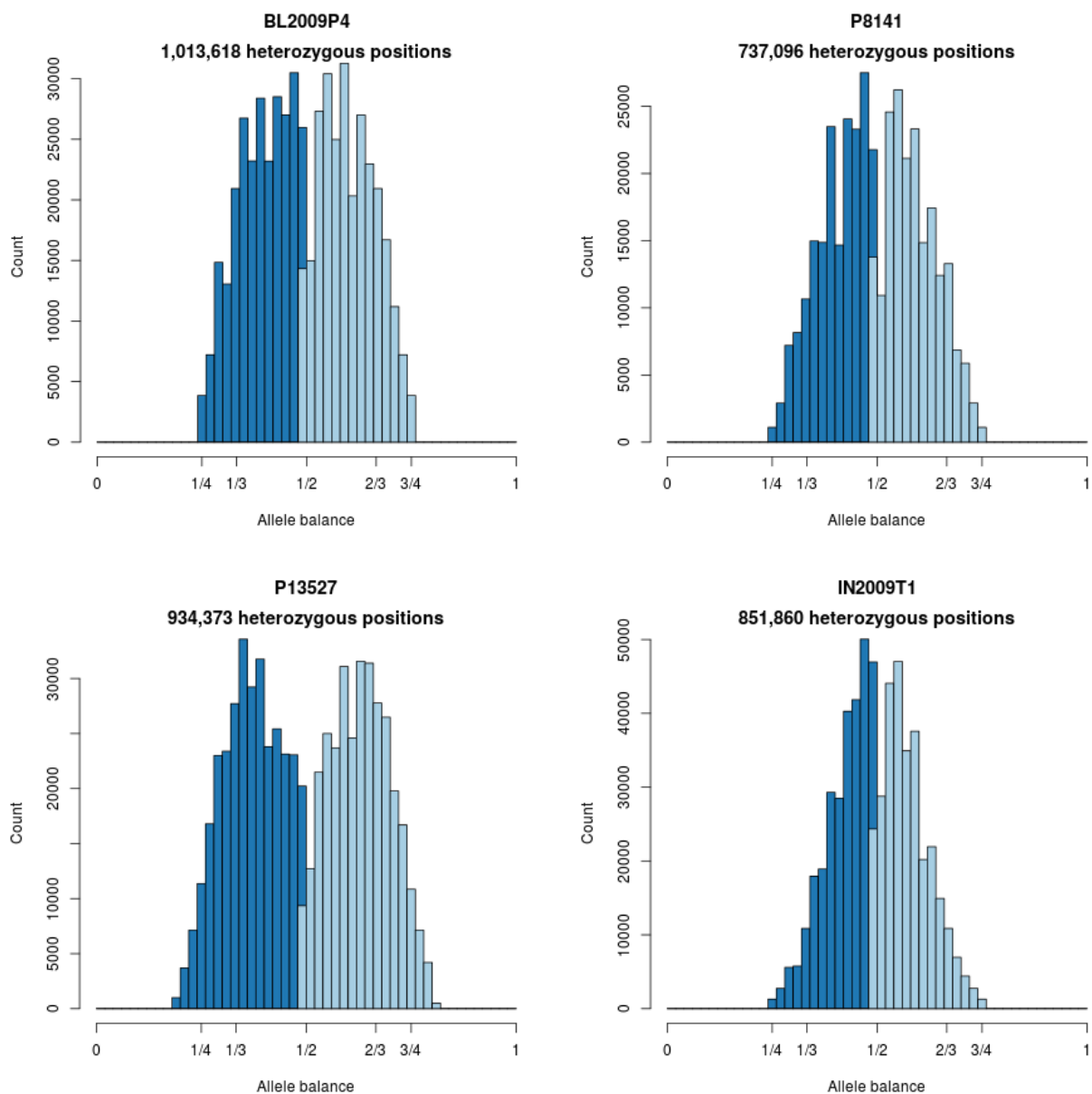
Data were processed as in the section ‘Gene copy number inference based on allele balance’ and in [3]. Allele balance, the frequency at which each allele was sequenced, was calculated for each available sample. A genomic VCF file (g.VCF) was created for each sample using GATK [4, 1]. Each g.VCF file was read into R [5] for processing with vcfR [2]. The allele depth (AD) and genotypes were extracted from the vcfR object and the first and second most abundant alleles were extracted from the allele depth. The genotype information was used to subset the allele depth information to only heterozygous positions. The 15th and 85th percentiles were calculated for each sample and each allele (the first and second most abundant alleles) and used as an inclusion threshold for depth filtering. A frequency for each allele was then calculated by dividing its allele depth by the sum of the first and second most abundant alleles. This information was used to plot each histogram. Samples are presented in the same order as in Supplemental item 1A-C.



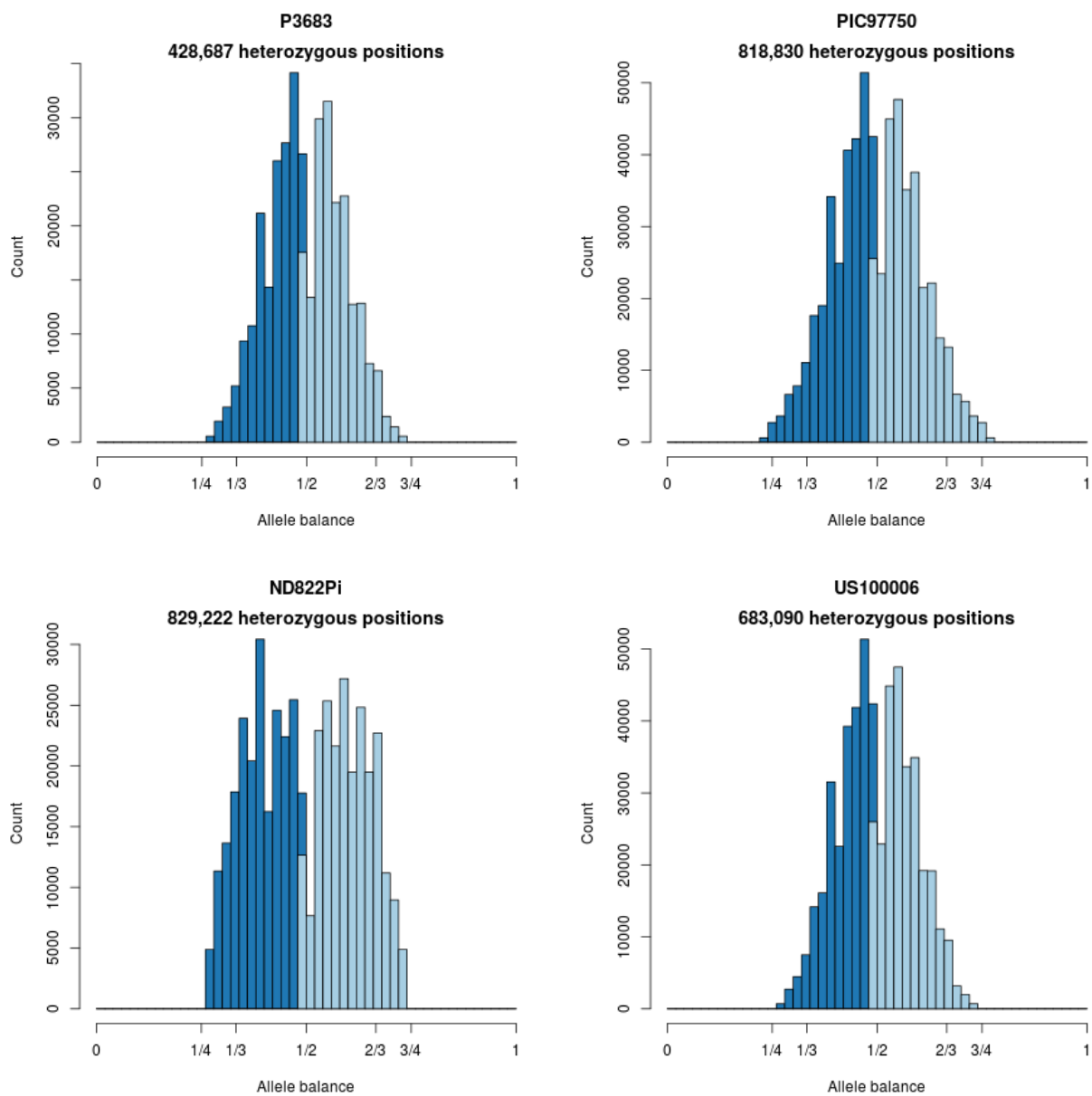
Supplemental item 3A. Histograms of allele balance for *P. infestans*.



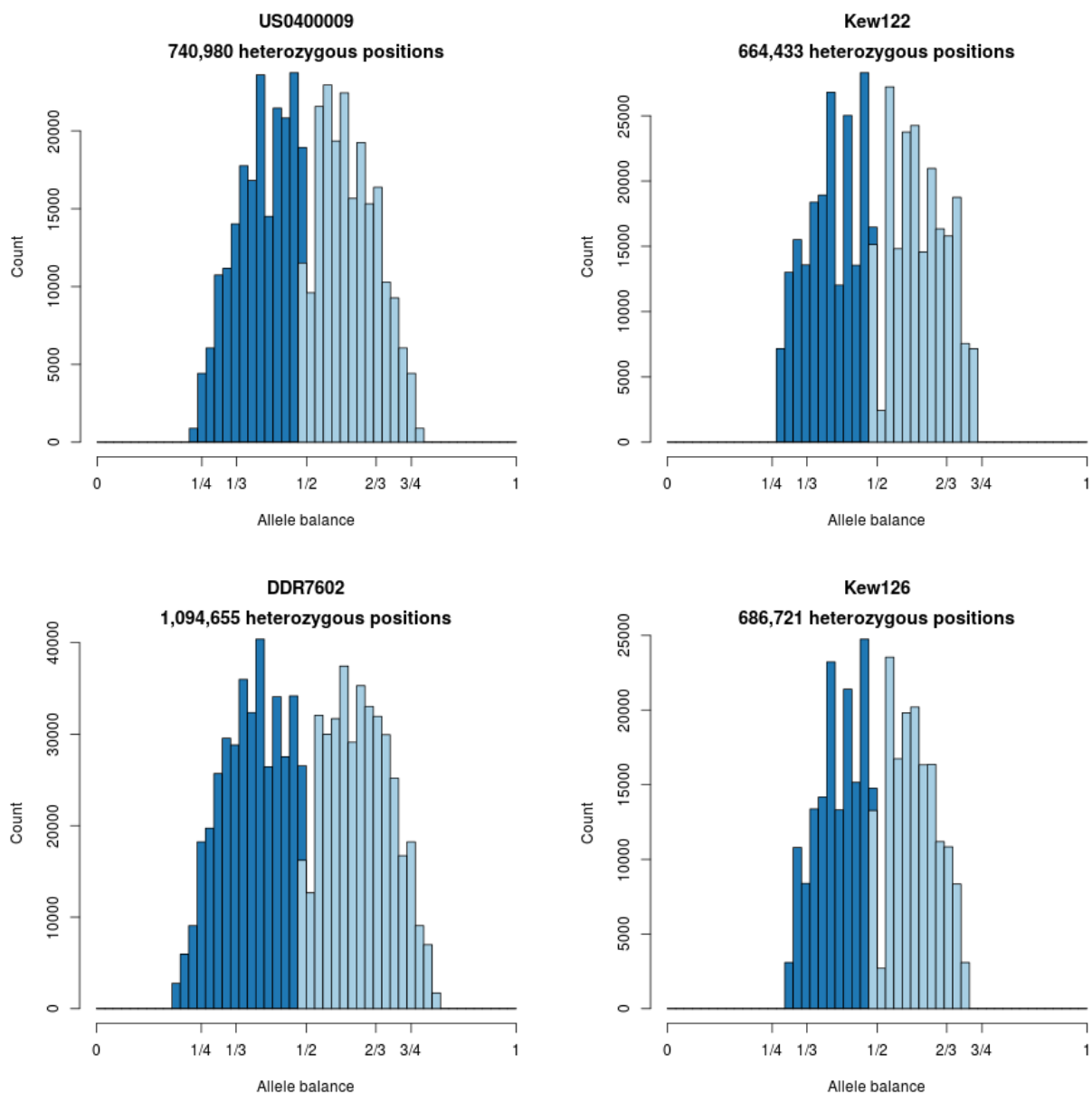
Supplemental item 3B. Histograms of allele balance for *P. infestans*.



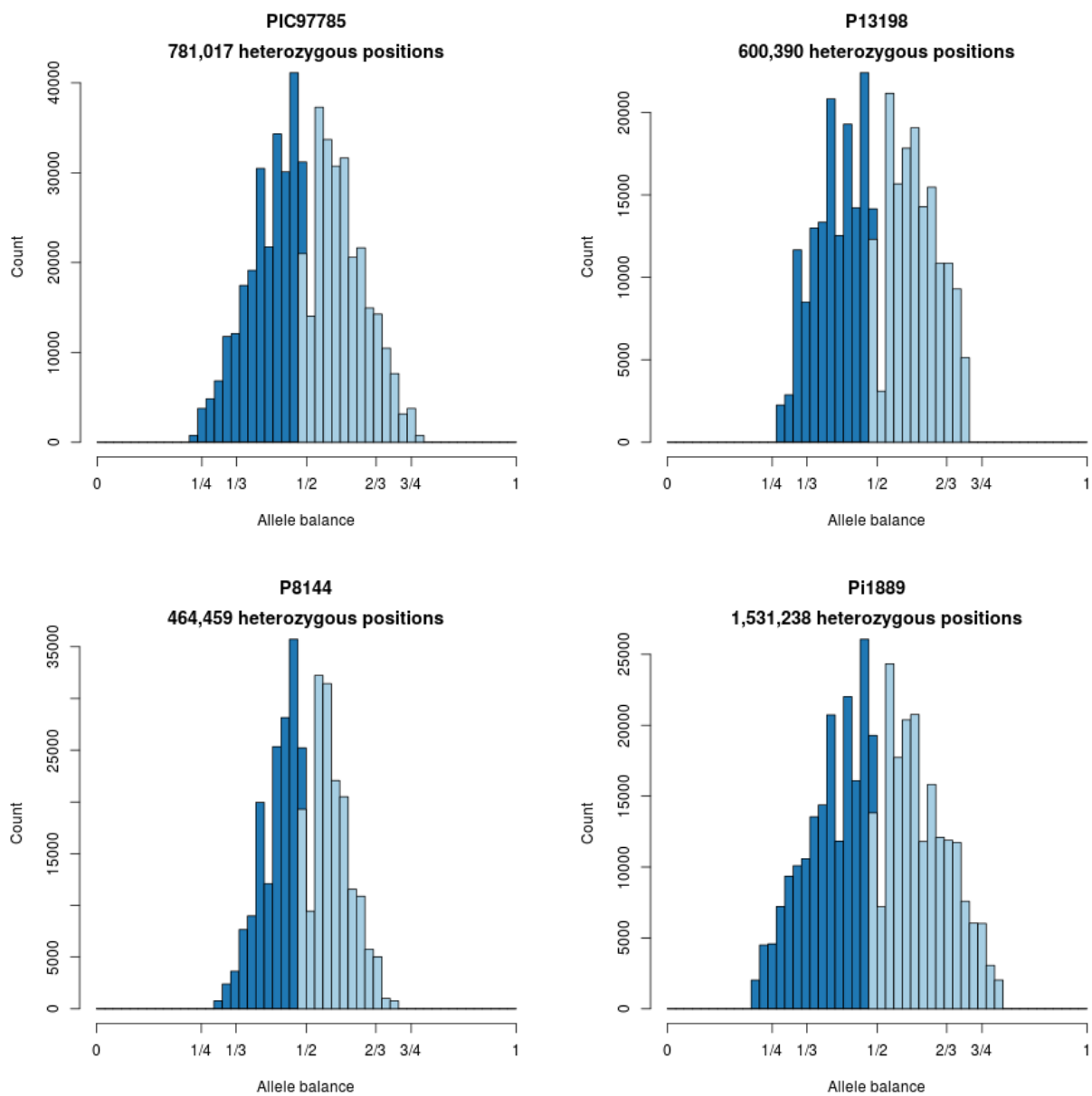
Supplemental item 3C Histograms of allele balance for *P. infestans*.



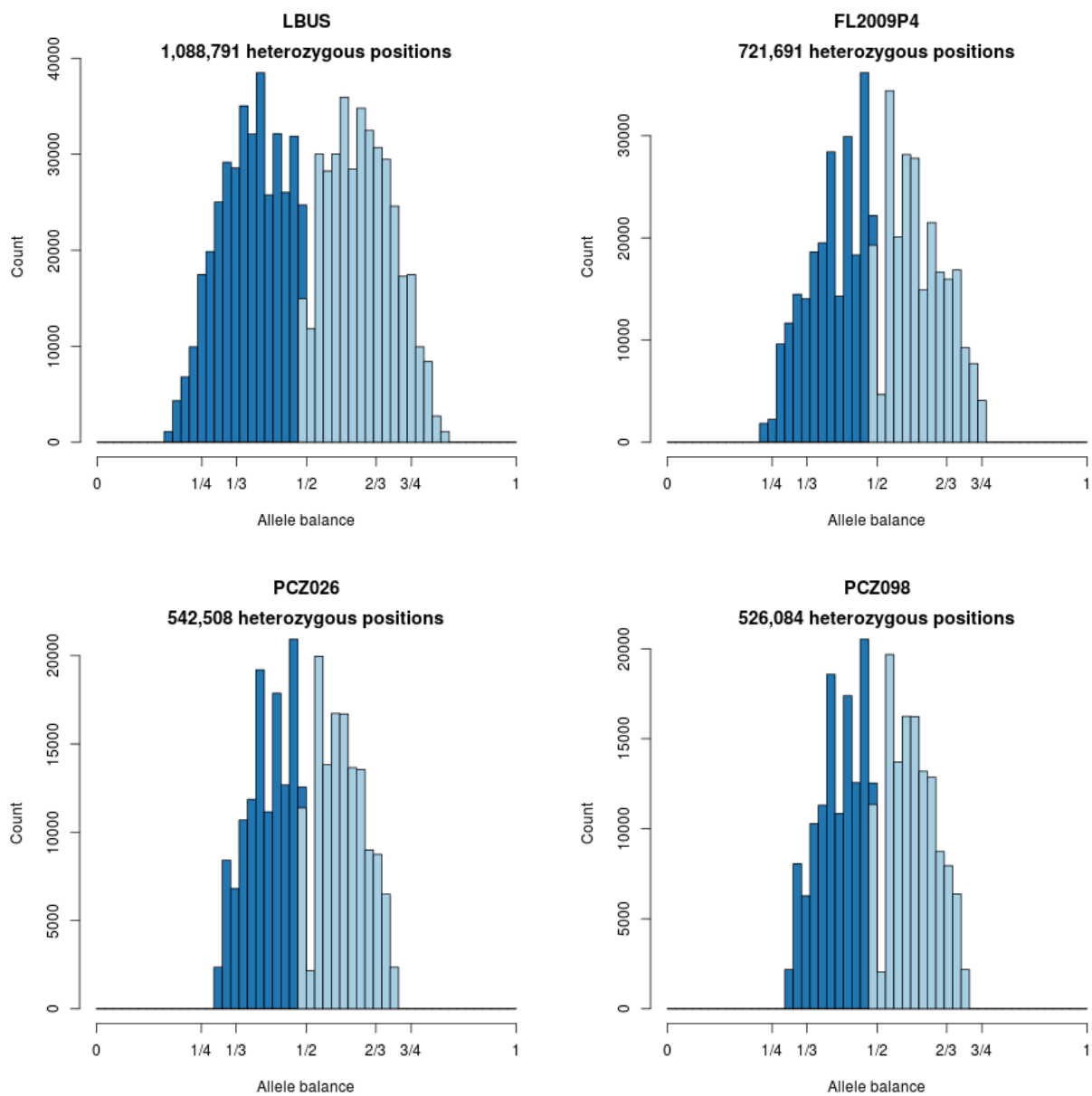
Supplemental item 3D. Histograms of allele balance for *P. infestans*.



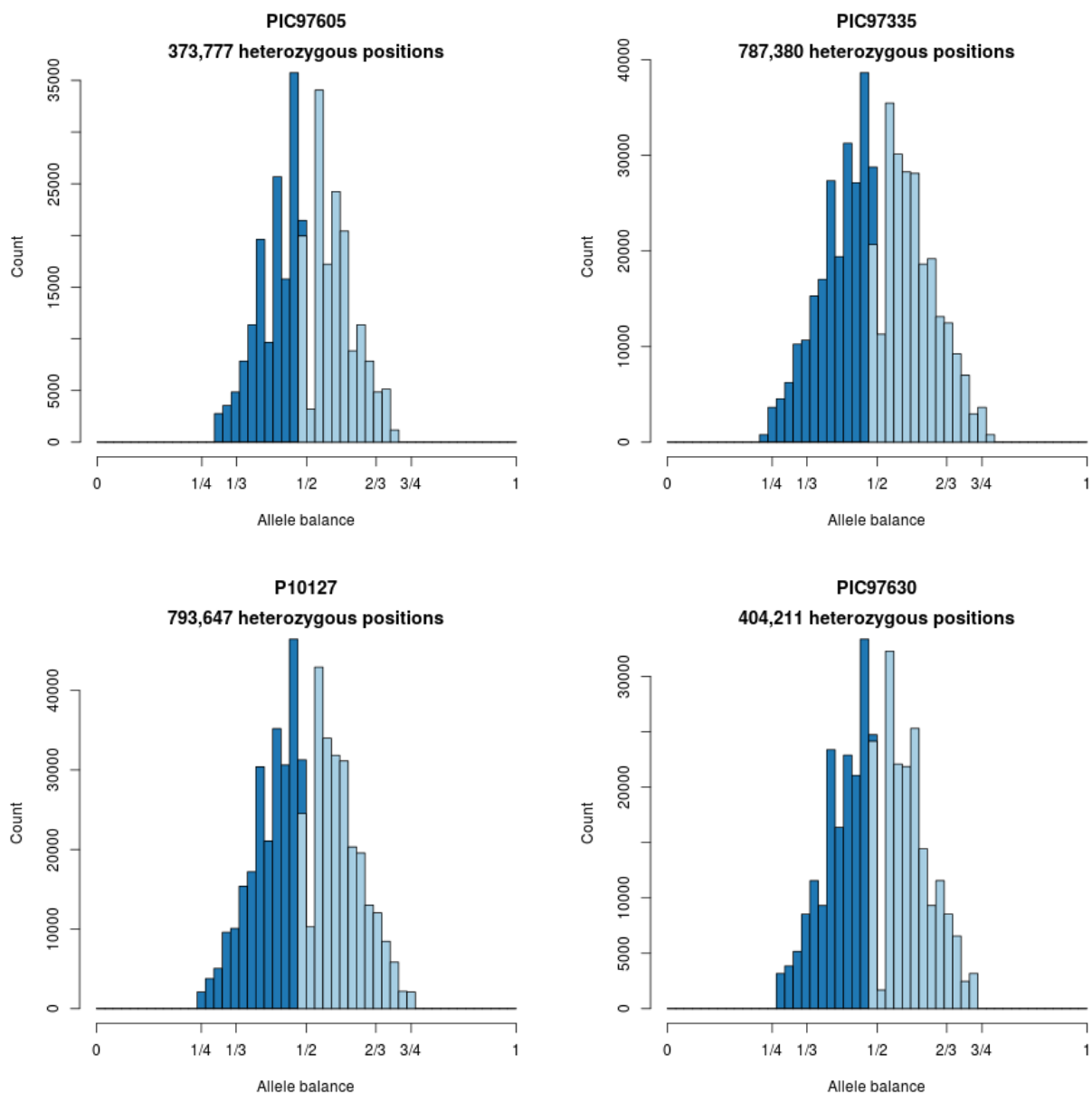
Supplemental item 3E. Histograms of allele balance for *P. infestans*.



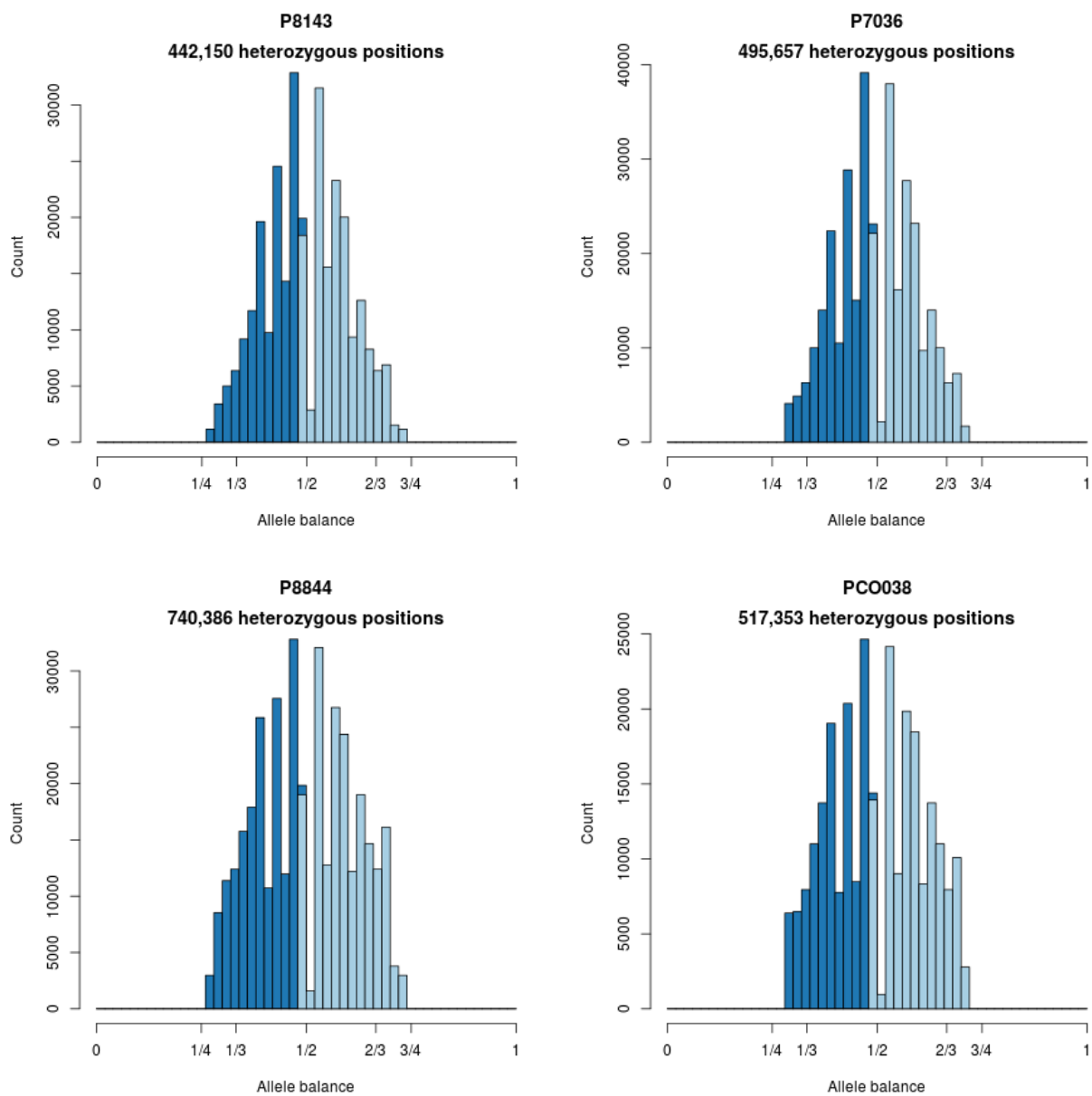
Supplemental item 3F. Histograms of allele balance for *P. infestans*.



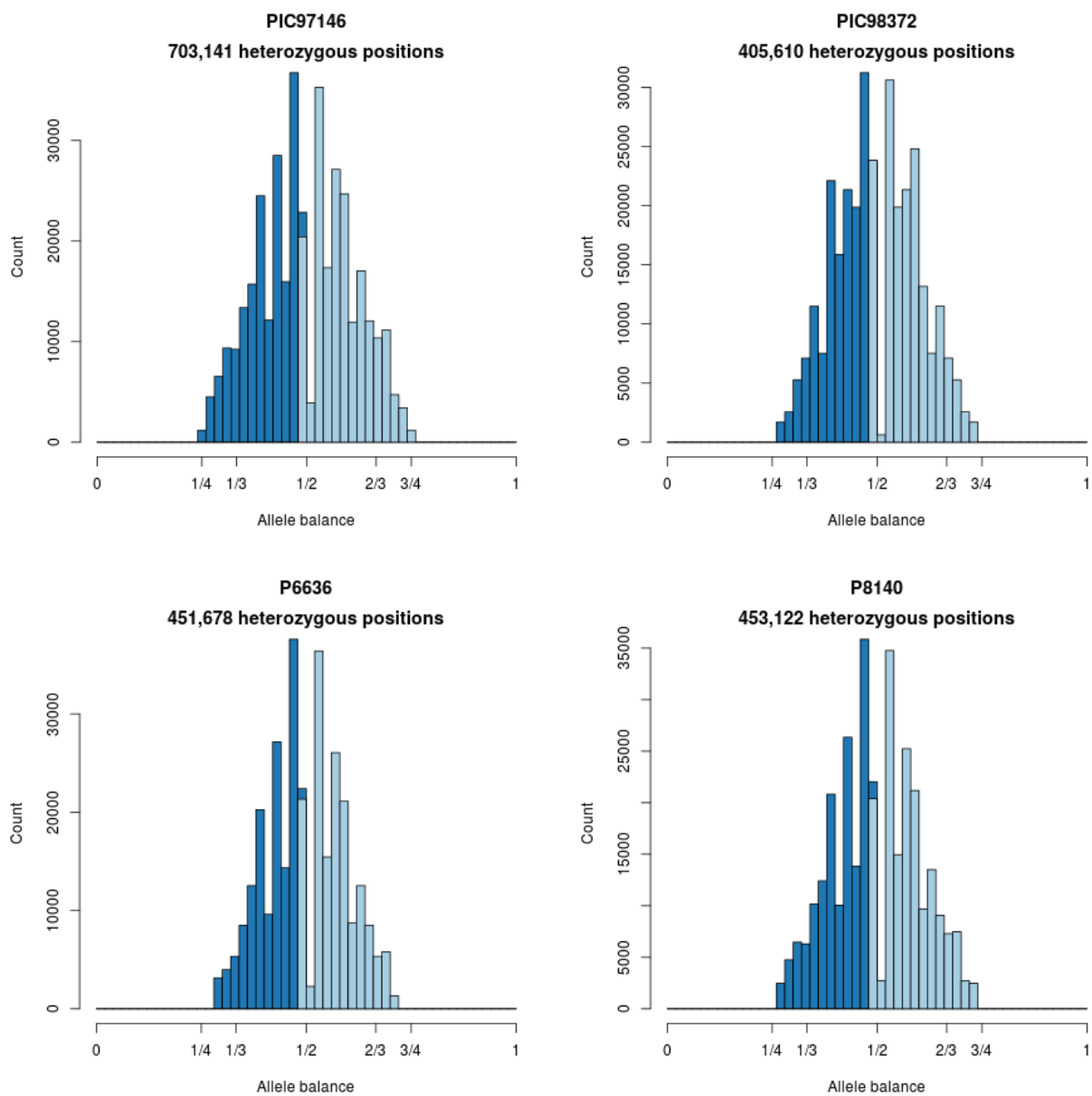
Supplemental item 3G. Histograms of allele balance for *P. infestans*.



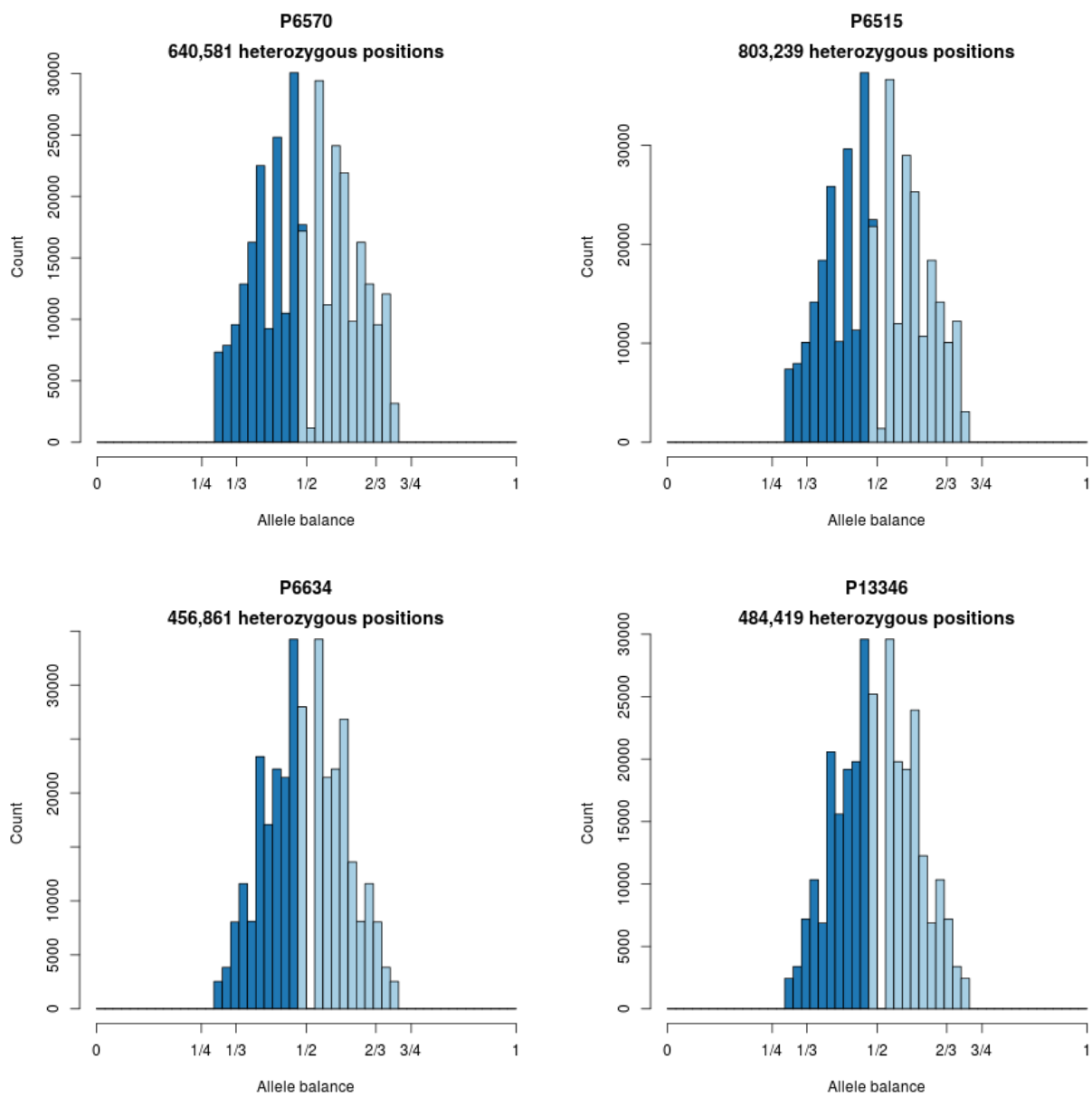
Supplemental item 3H. Histograms of allele balance for *P. infestans*.



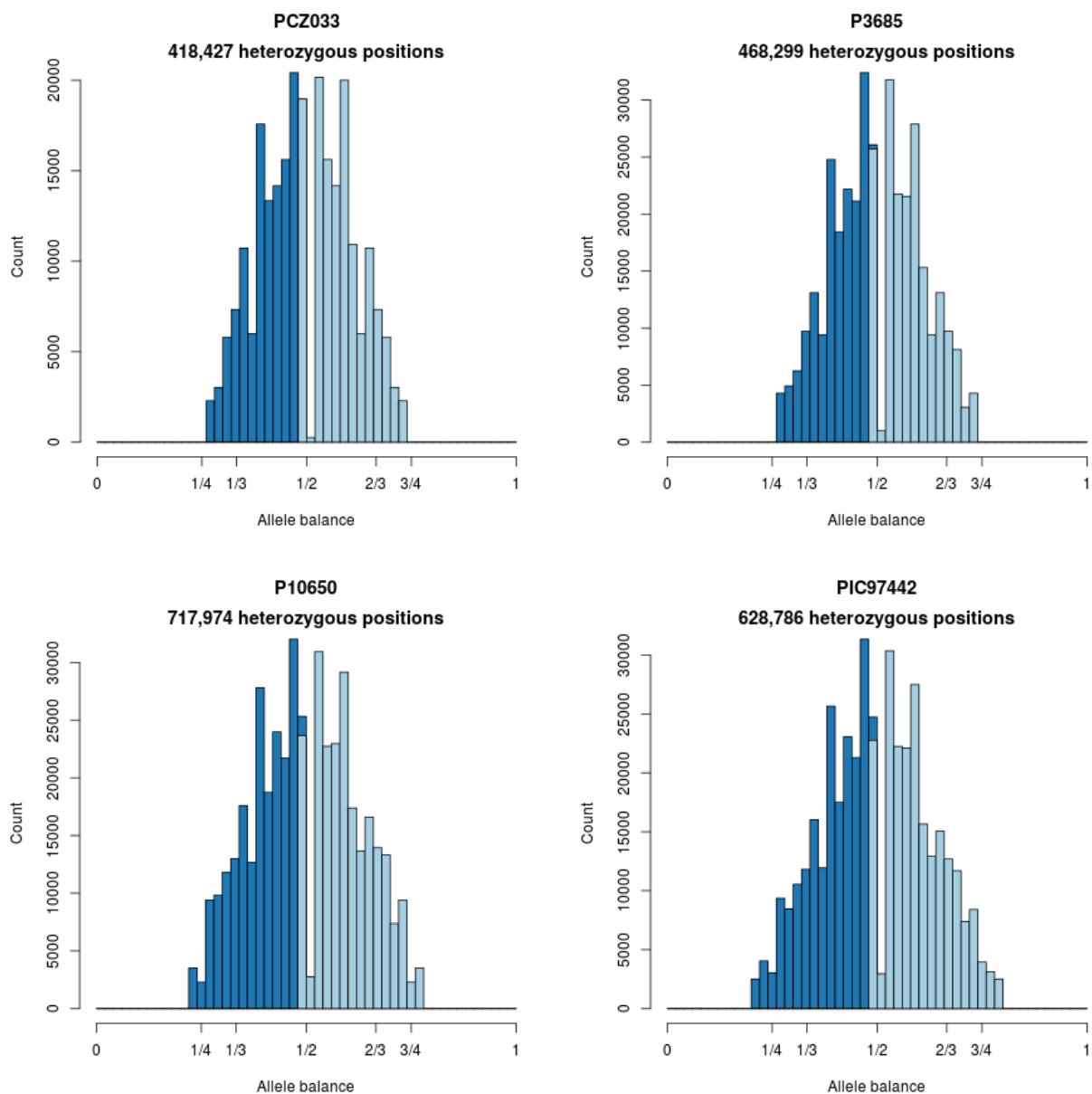
Supplemental item 3I. Histograms of allele balance for *P. infestans*.



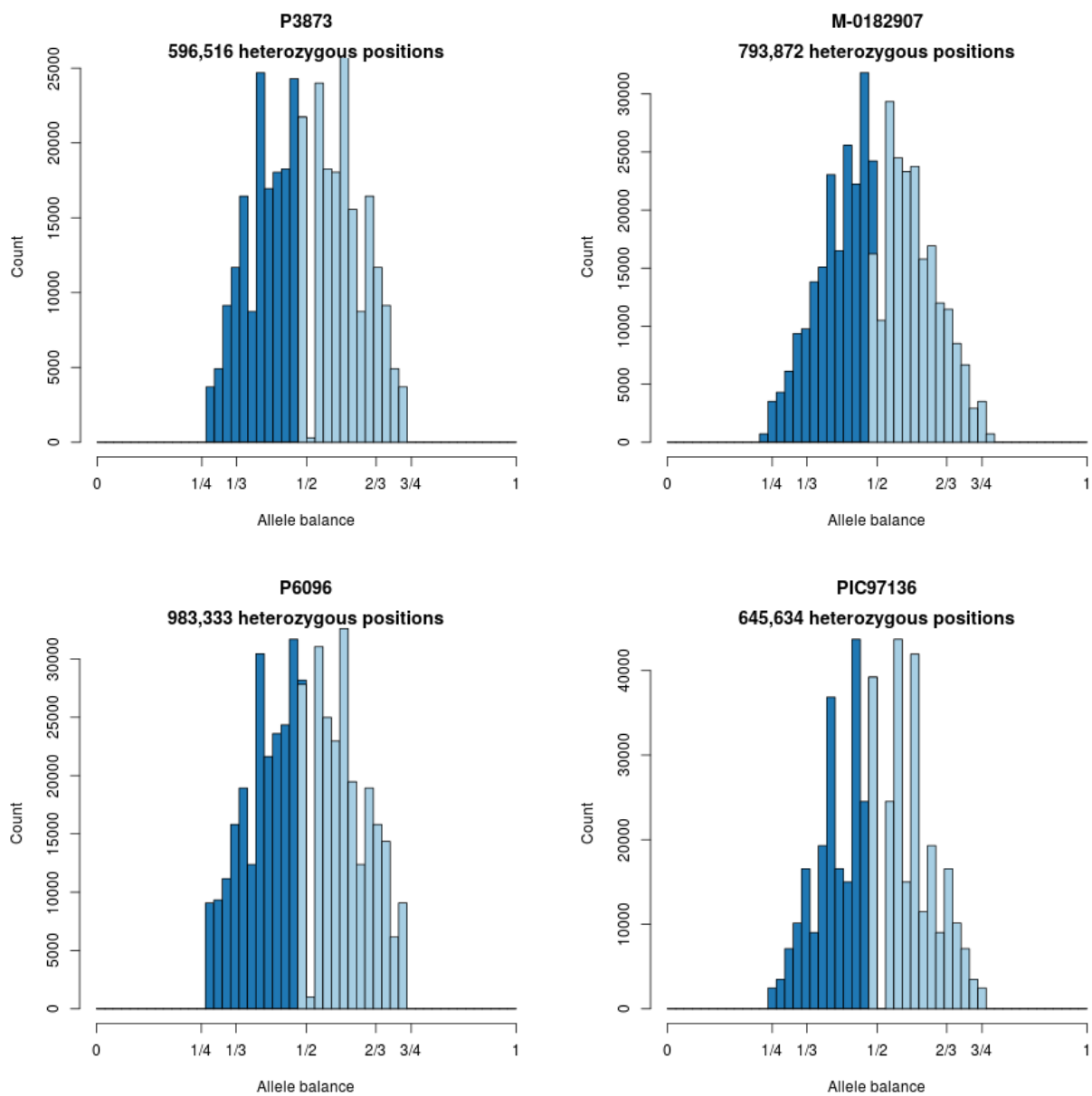
Supplemental item 3J. Histograms of allele balance for *P. infestans*.



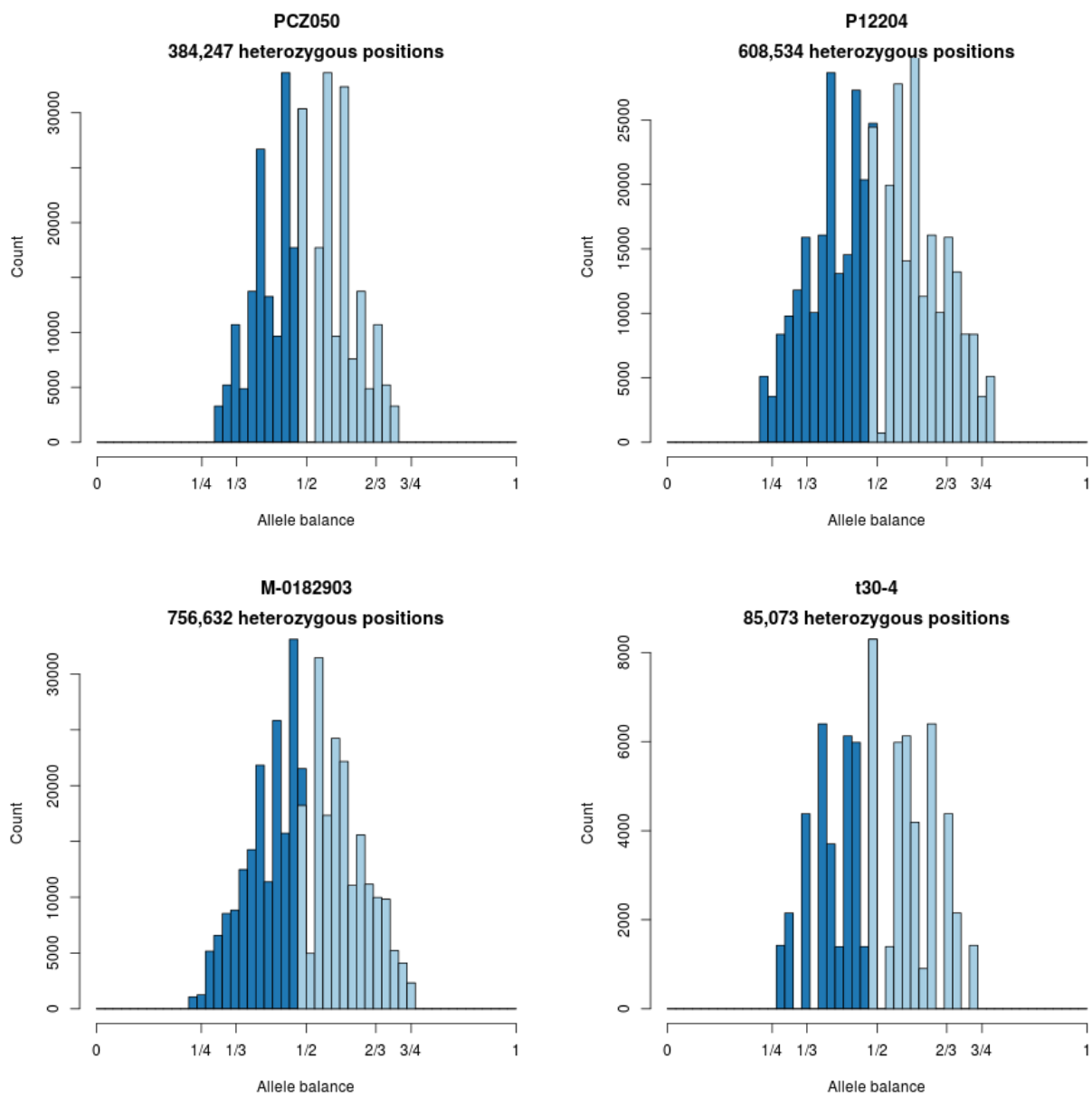
Supplemental item 3K. Histograms of allele balance for *P. infestans*.



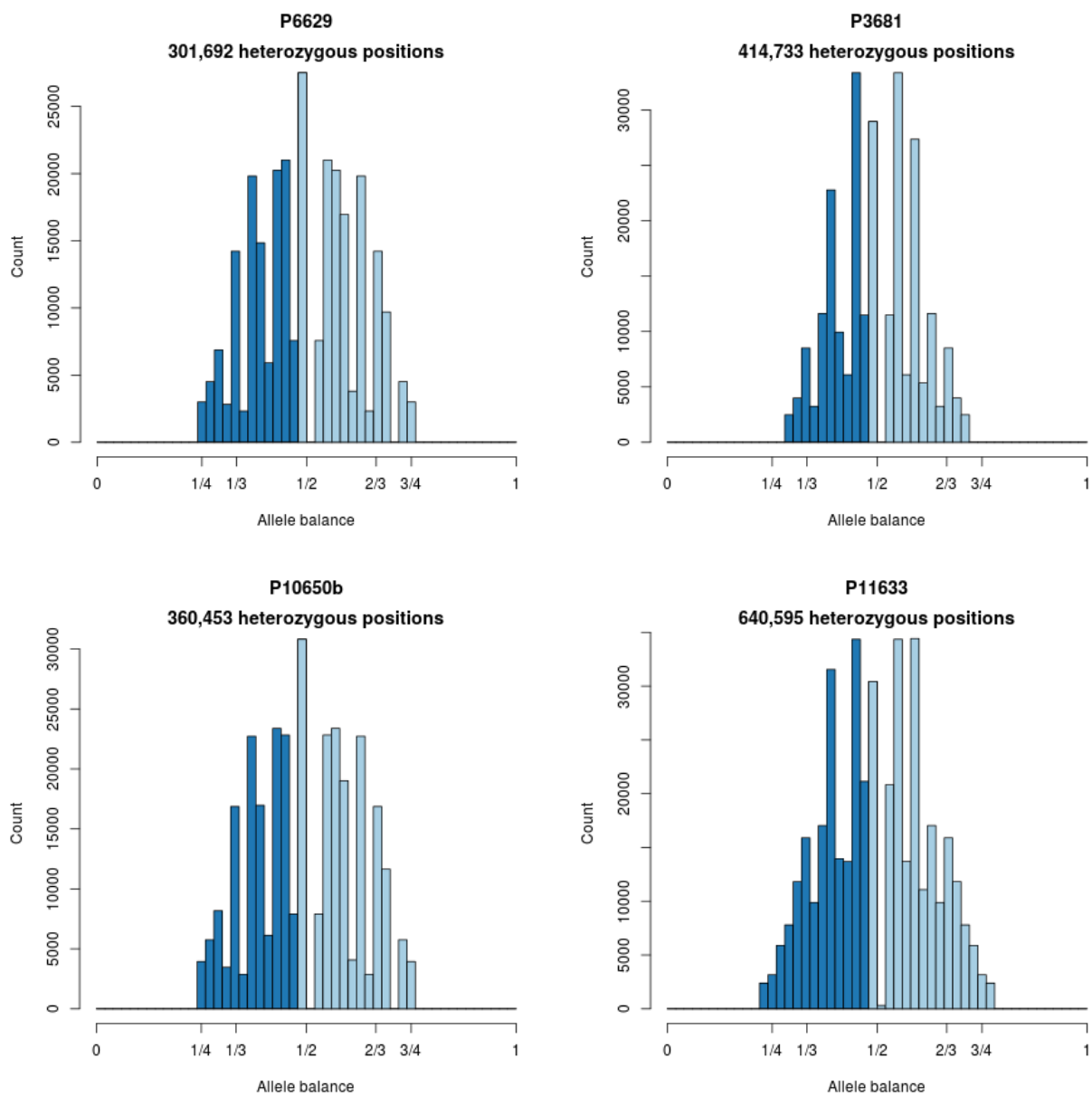
Supplemental item 3L. Histograms of allele balance for *P. infestans*.



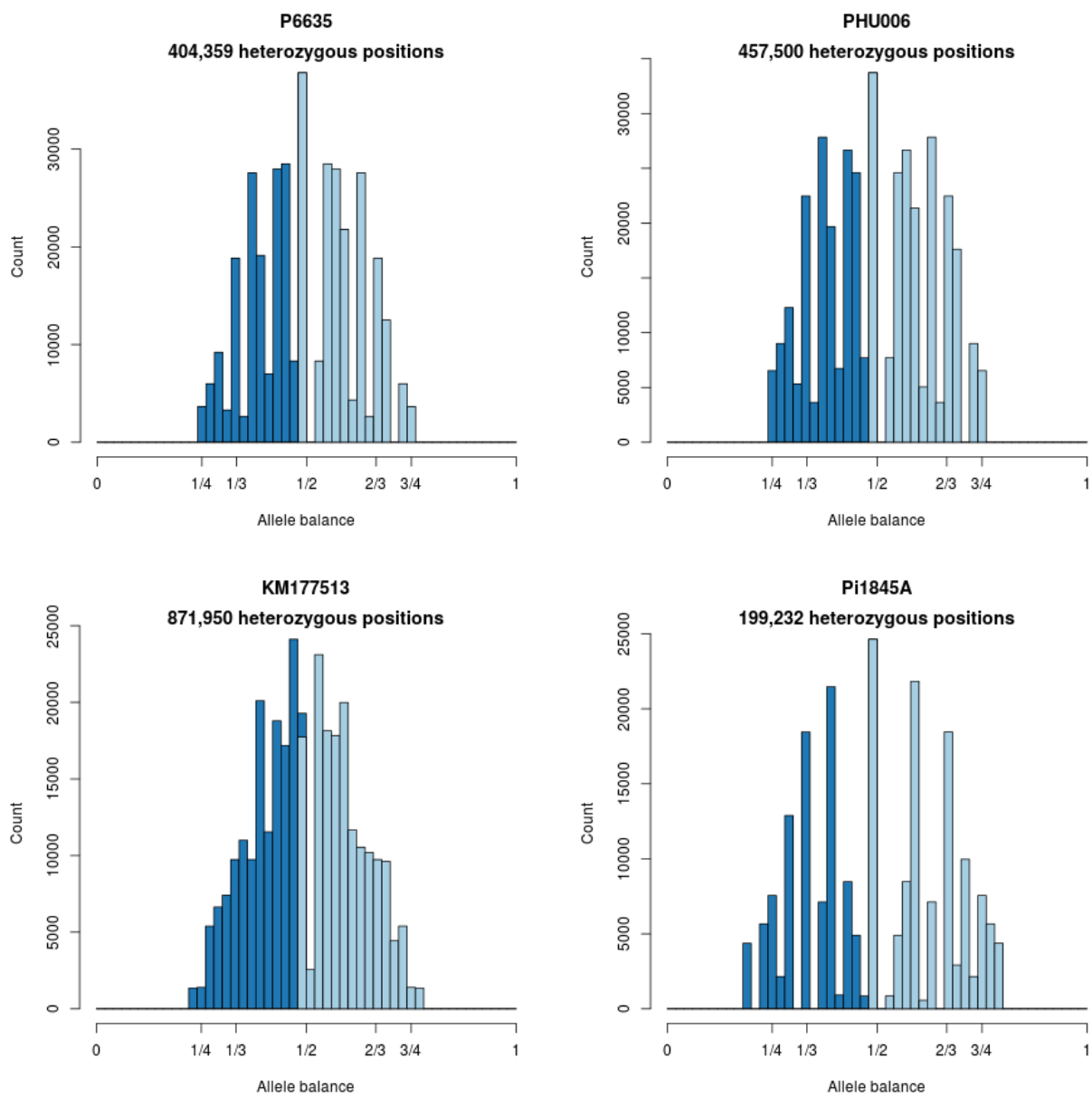
Supplemental item 3M. Histograms of allele balance for *P. infestans*.



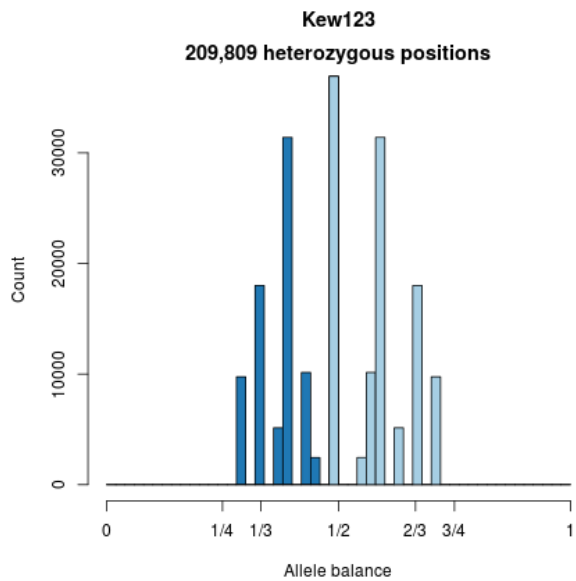
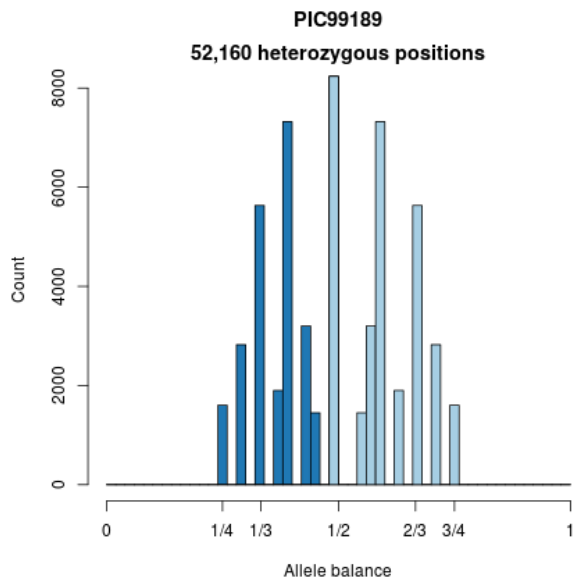
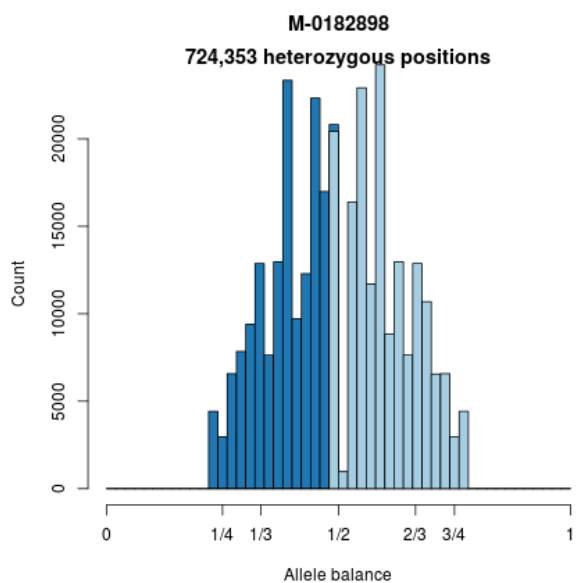
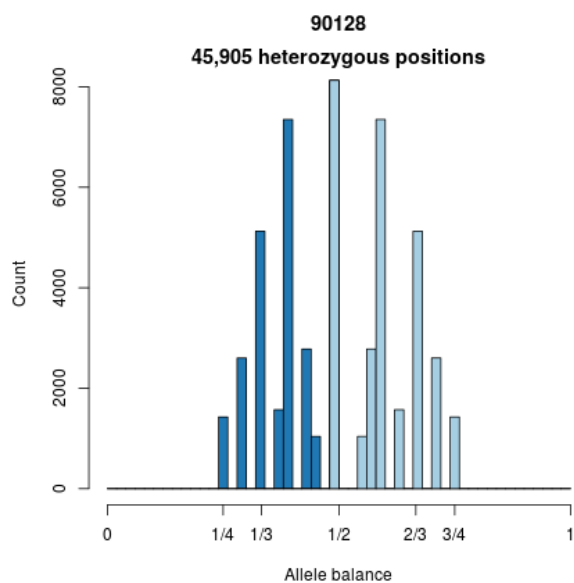
Supplemental item 3N. Histograms of allele balance for *P. infestans*.



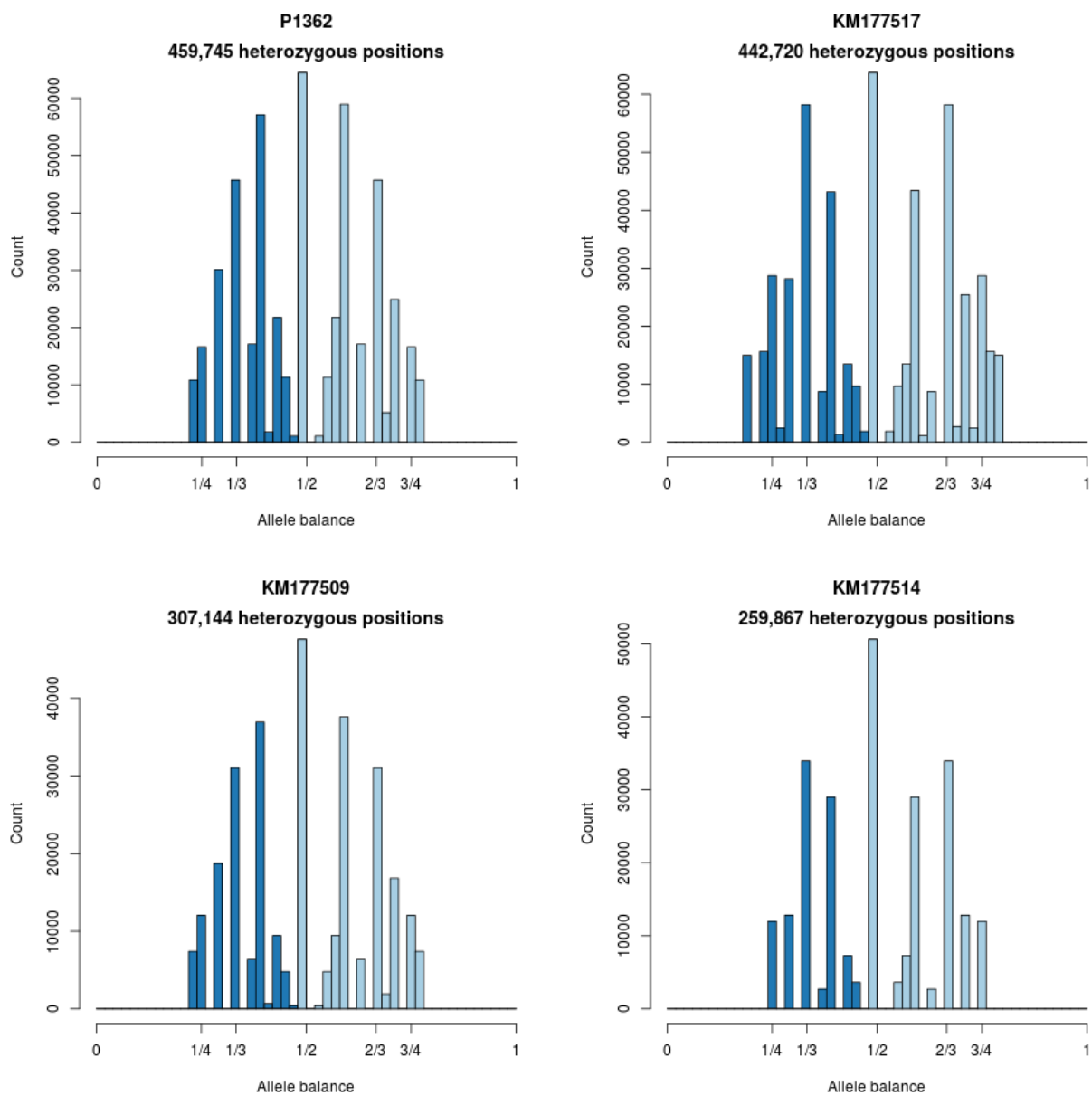
Supplemental item 3O. Histograms of allele balance for *P. infestans*.



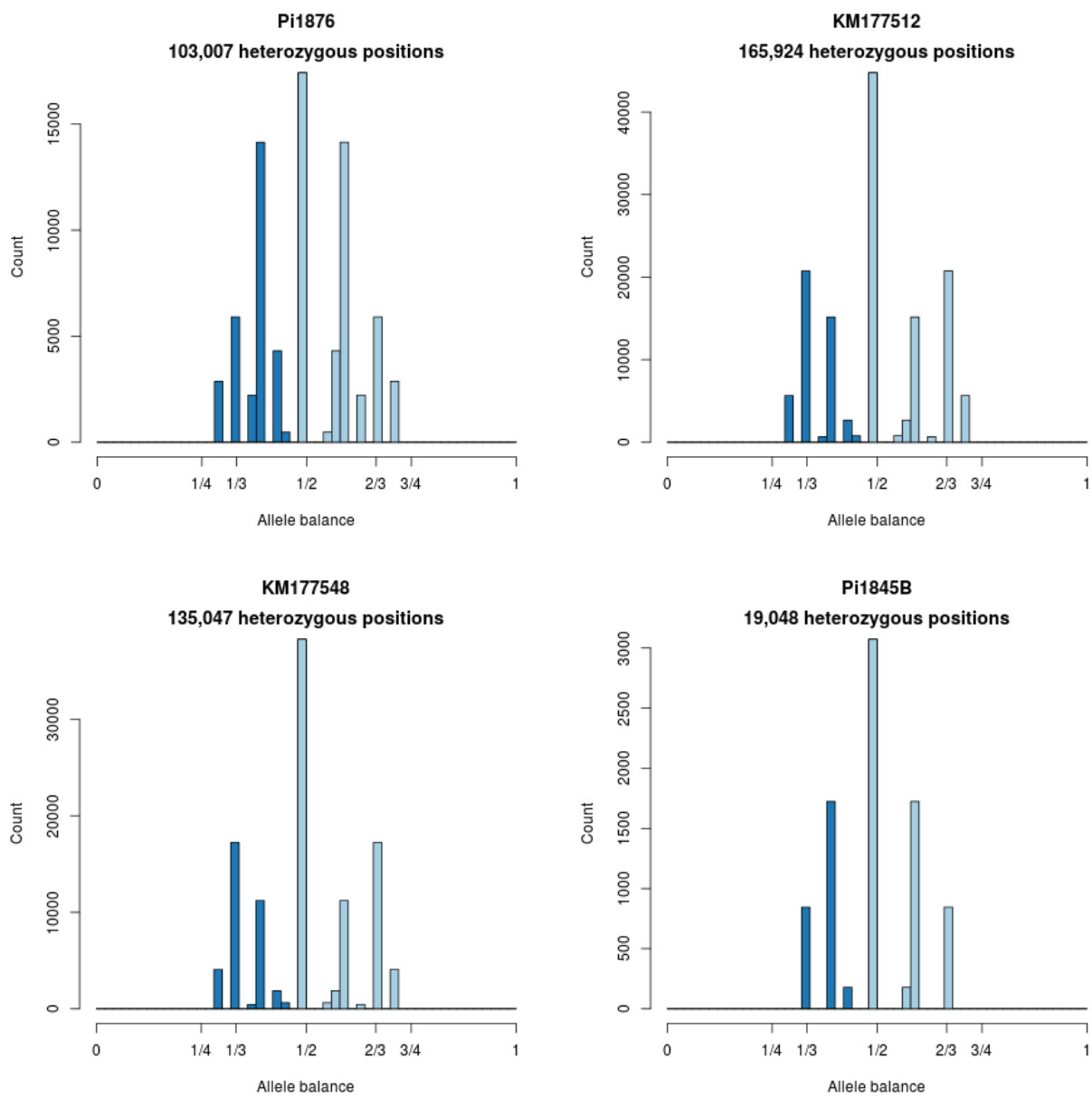
Supplemental item 3P. Histograms of allele balance for *P. infestans*.



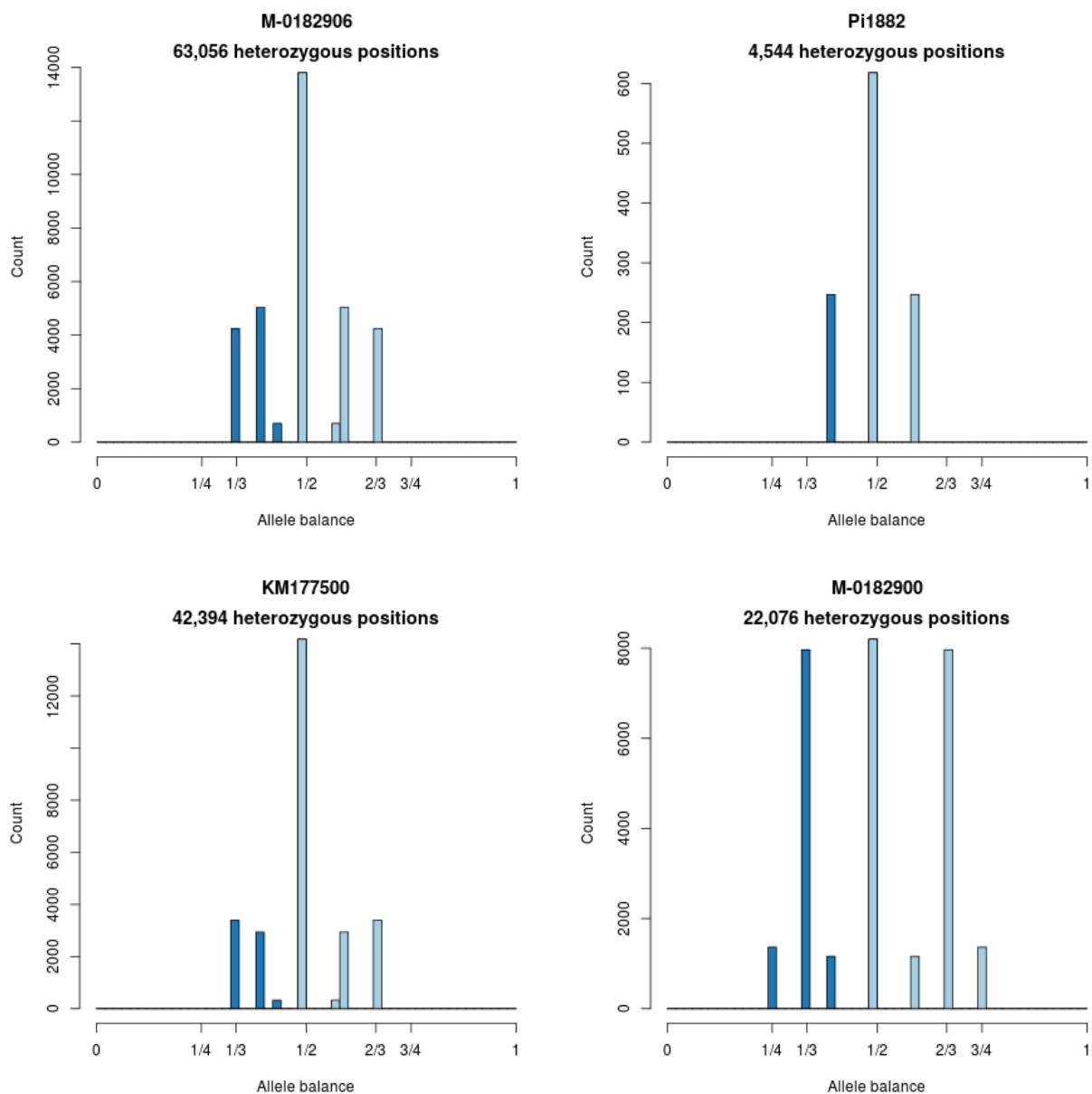
Supplemental item 3Q. Histograms of allele balance for *P. infestans*.



Supplemental item 3R. Histograms of allele balance for *P. infestans*.



Supplemental item 3S. Histograms of allele balance for *P. infestans*.



Supplemental item 3T. Histograms of allele balance for *P. infestans*.

References

- [1] MA DePristo, E Banks, R Poplin, KV Garimella, JR Maguire, C Hartl, AA Philippakis, G Del Angel, MA Rivas, M Hanna, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, 43(5):491–498, 2011.
- [2] BJ Knaus and NJ Grünwald. vcfR: a package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources*, 17(1):44–53, 2017.
- [3] BJ Knaus and NJ Grünwald. Inferring variation in copy number using high throughput sequencing data in R. *Frontiers in Genetics*, 9:123, 2018.
- [4] A McKenna, M Hanna, E Banks, A Sivachenko, K Cibulskis, A Kernytsky, K Garimella, D Altshuler, S Gabriel, M Daly, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9):1297–1303, 2010.
- [5] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2017.