

Haplotype Analysis of a Maize Multi-parent Population to Increase Power in Finding QTL

Sarah G. Odell¹, Jeffrey Ross-Ibarra^{1,2,3}, and Daniel Runcie¹

¹Department of Plant Sciences, University of California Davis; ² The Center for Population Biology, University of California Davis; ³ UC Davis Genome Center

RUNCIElab



Abstract

Maize is one of the most important crops in the world, as well as a widely-used model organism for understanding plant genetics and biology. The search for quantitative trait loci (QTL) that explain complex traits such as drought tolerance has been ongoing in many species. Methods such as biparental QTL mapping and genome wide association studies (GWAS) each have their own advantages and limitations. Multiparent advanced generation intercrossing (MAGIC) populations contain more recombination events and genetic diversity than biparental mapping populations and reduce the confounding factor of population structure that is an issue in association mapping populations.

Presented here is current progress in using a MAGIC population of double haploid maize lines created from 16 diverse founders to identify alleles associated with variation in drought response, as well as future plans to compare the QTL found to expression QTL identified from RNASeq data of the population.

Genotype Imputation & Validation

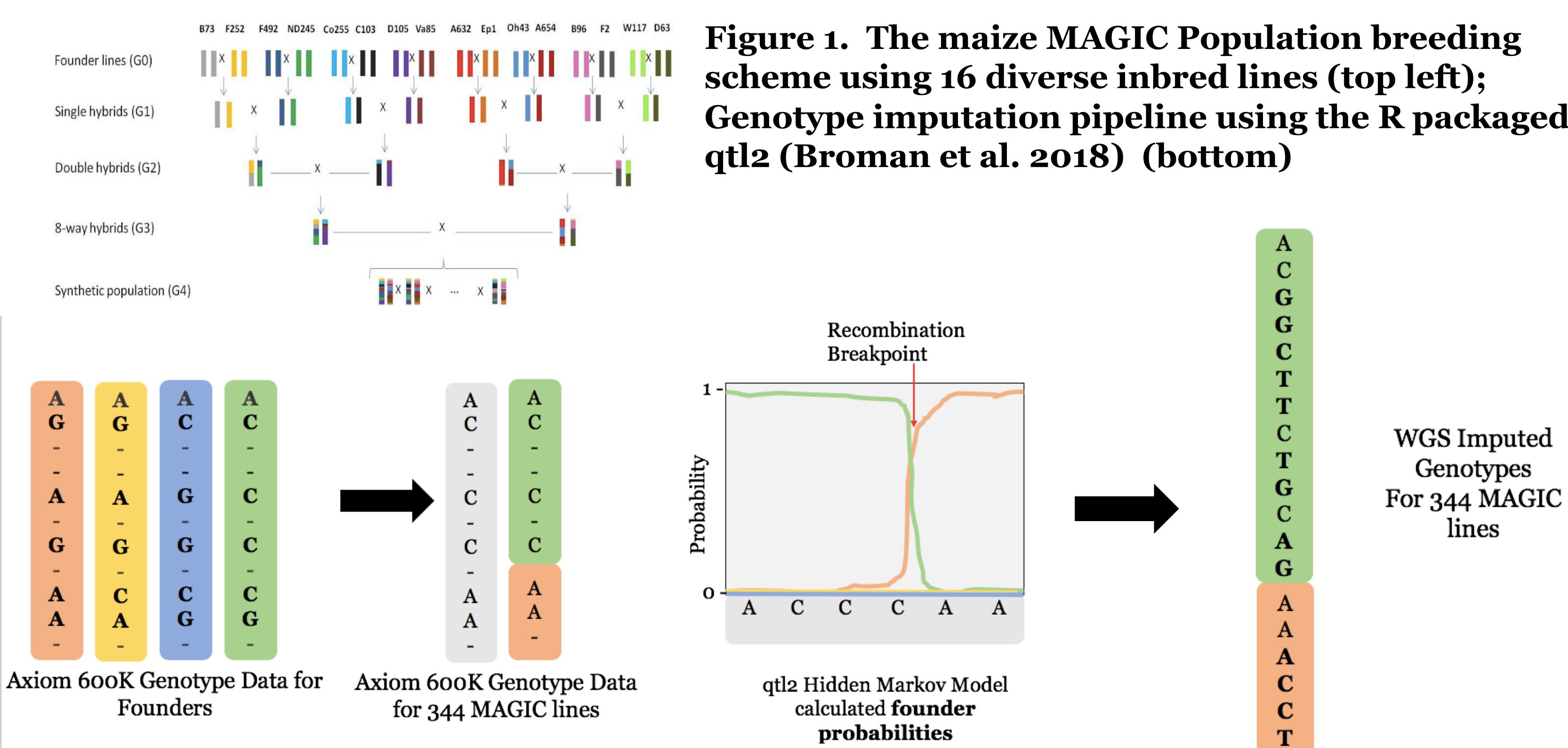


Figure 1. The maize MAGIC Population breeding scheme using 16 diverse inbred lines (top left); Genotype imputation pipeline using the R packaged qtL2 (Broman et al. 2018) (bottom)

We used the package qtL2 (Broman et al. 2018) to infer the founder that donated particular regions of DH chromosomes (**Figure 1**). From these founder probabilities, whole-genome sequence data from the 16 founders was used to reconstruct WGS data for the 344 MAGIC double haploid lines in the population.

Simulation showed that qtL2 had very high accuracy in terms of allele assignment (**Figure 2**).

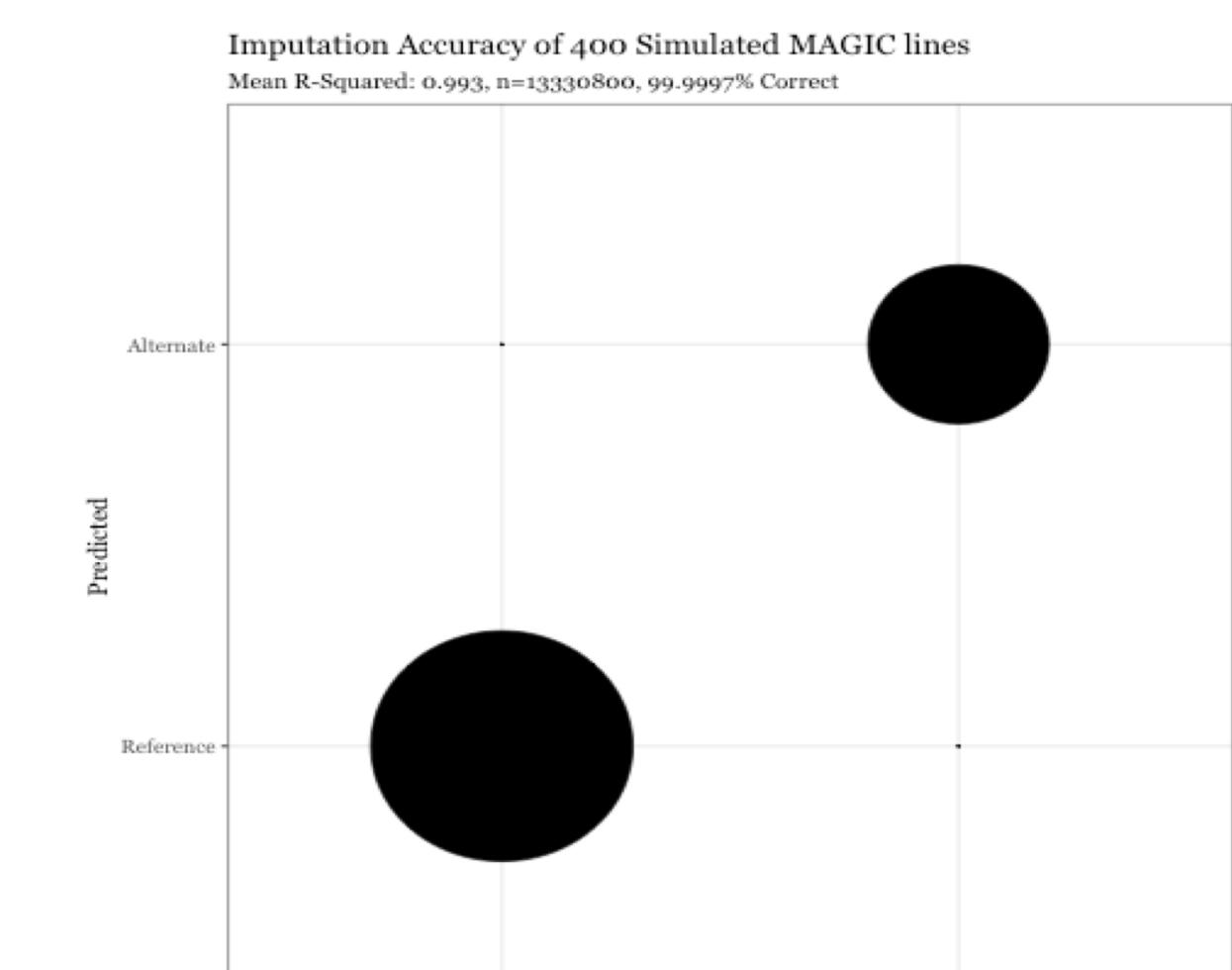


Figure 2. Assessment of qtL2 imputation accuracy using 400 simulated MAGIC lines. Actual versus predicted SNPs for reference and alternate alleles, with 99.99% of sites assigned to the correct allele for chromosome 10

Haplotype Analysis

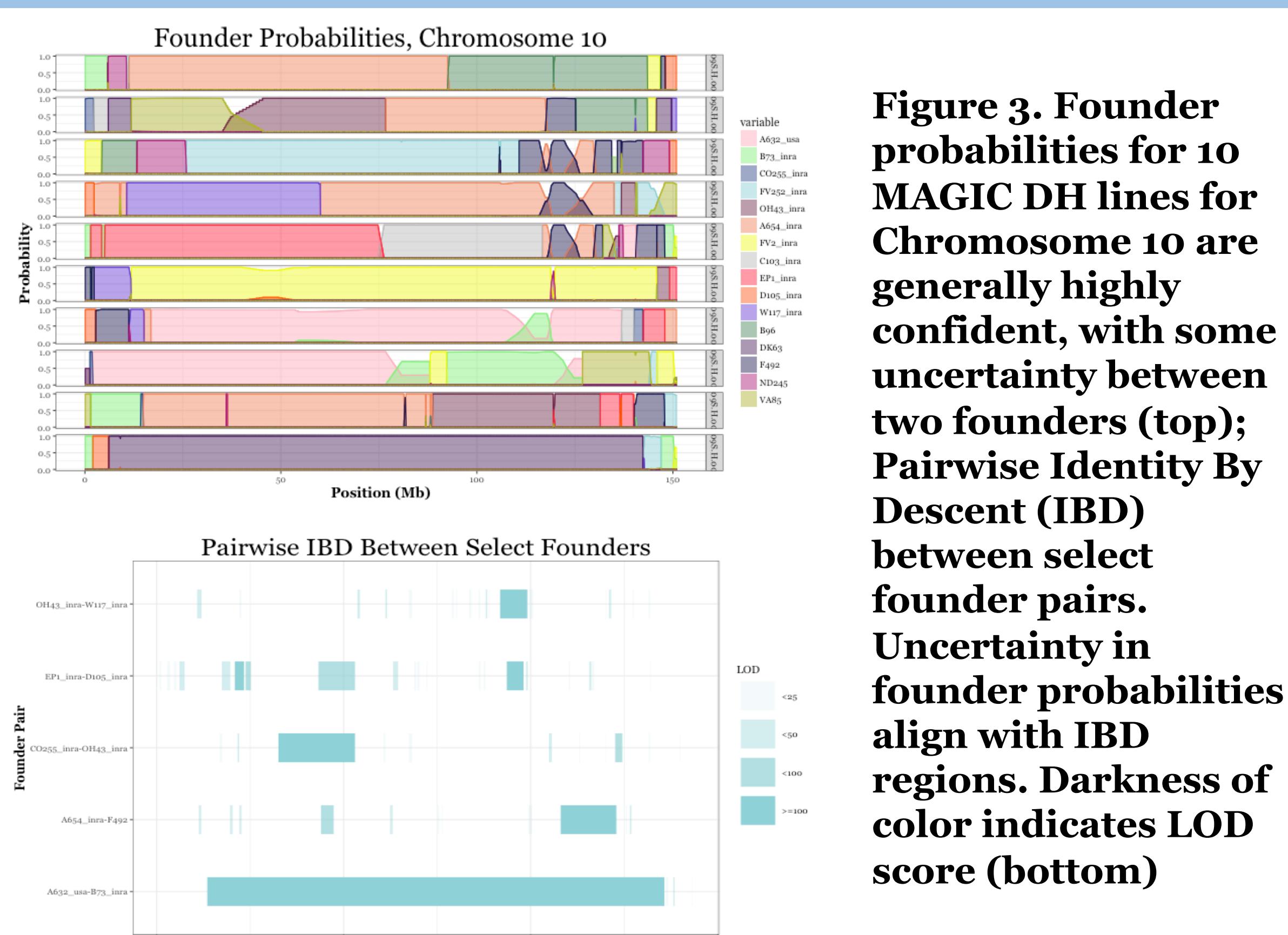


Figure 3. Founder probabilities for 10 MAGIC DH lines for Chromosome 10 are generally highly confident, with some uncertainty between two founders (top); Pairwise Identity By Descent (IBD) between select founder pairs. Uncertainty in founder probabilities align with IBD regions. Darkness of color indicates LOD score (bottom)

Identification of Identity By Descent (IBD) regions between the population founders showed that some founder pairs were in IBD, and therefore, genetically similar (**Figure 3**).

By separating chromosomes into blocks with distinct IBD patterns, we can group blocks by haplotype rather than by founder (**Figure 4**). By doing this, we hope to increase the power of GWAS and QTL mapping

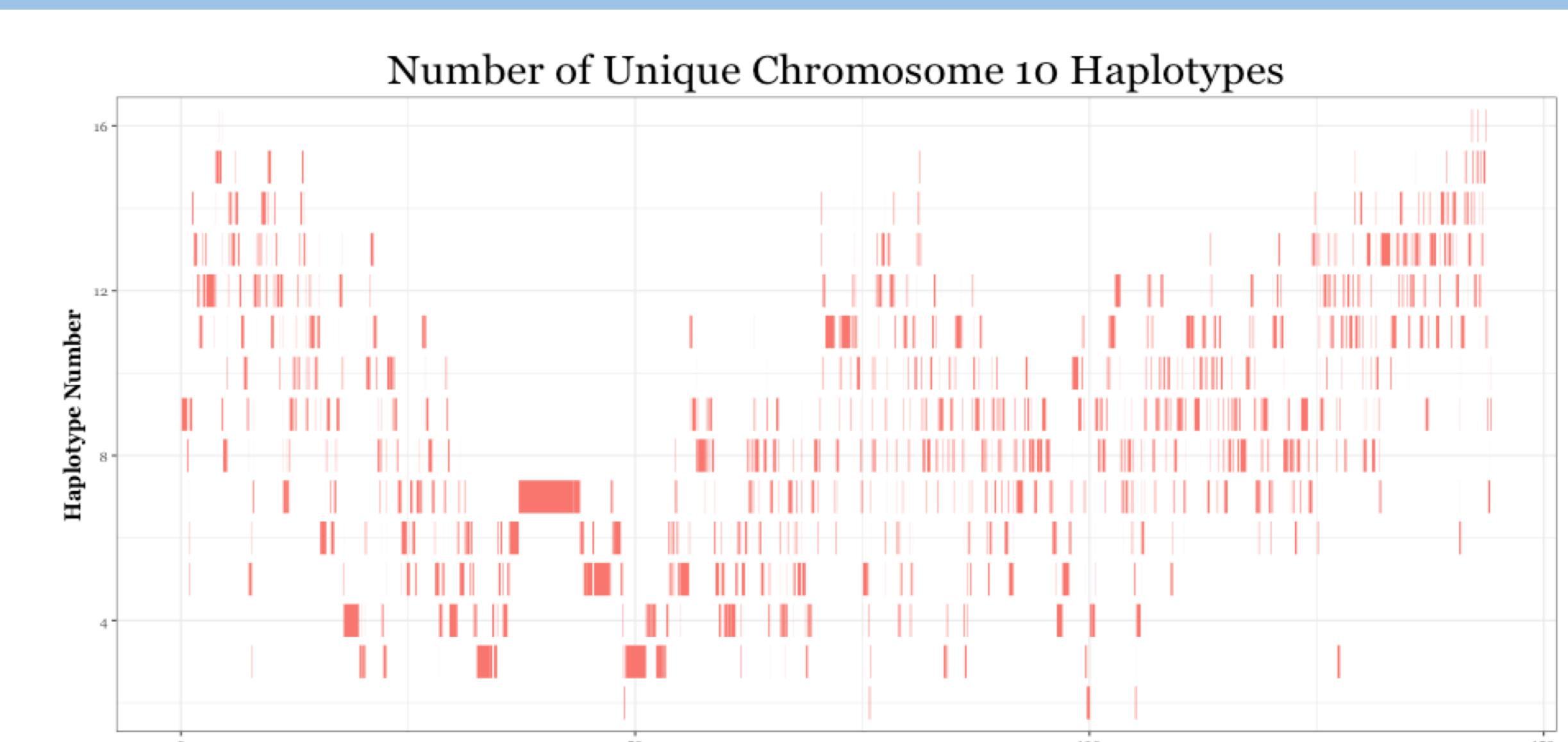
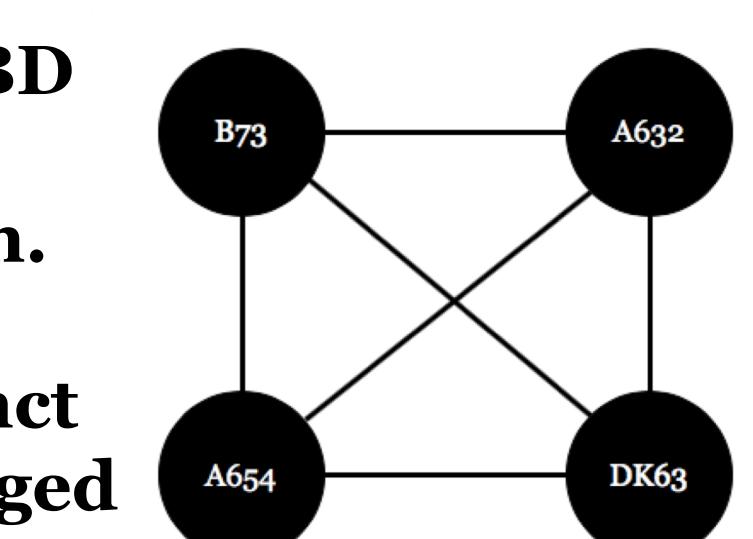
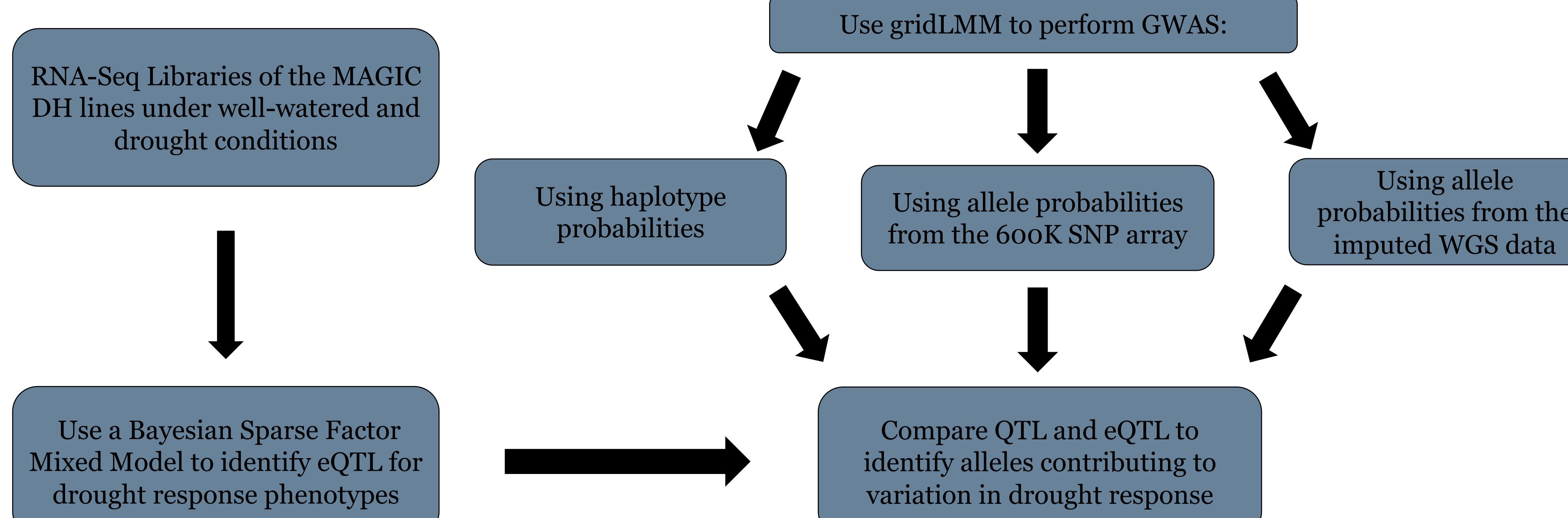


Figure 4. Number of unique haplotype per IBD block for chromosome 10 of the 16 MAGIC founder lines (top); Example of an IBD graph. Such an IBD relationship in a chromosomal region would indicate that there are 13 distinct haplotype, rather than 16. This can be leveraged to increase statistical power



Future Work



References

- Broman, K., Gatti, D., Simecek, P., Furlotte, N., Prins, P., Sen, S., Yandell, B., Churhcell, B. (2018). R/qtL2: Software for Mapping Quantitative Trait Loci with High-Dimensional Data and Multi-parent Populations, GENETICS Early online December 27, 2018; <https://doi.org/10.1534/genetics.118.301595>
- Runcie, D. E., & Mukherjee, S. (2013). Dissecting high-dimensional phenotypes with bayesian sparse factor analysis of genetic covariance matrices. *Genetics*, 194(3), 753-67. <https://doi.org/10.1534/genetics.113.151217>
- Runcie, D. E., & Crawford L. (2018) Fast and flexible linear mixed models for genome-wide genetics. bioRxiv 373902; doi: <https://doi.org/10.1101/373902>