Chapter 3: Using the *Eucalyptus polybractea* Genome Improved Genetic Variant Identification Compared to Using a Pseudo-Reference

Swapan Chakrabarty

College of Forest Resources and Environmental Science, Michigan Technological University,

Houghton, 49931 MI, USA

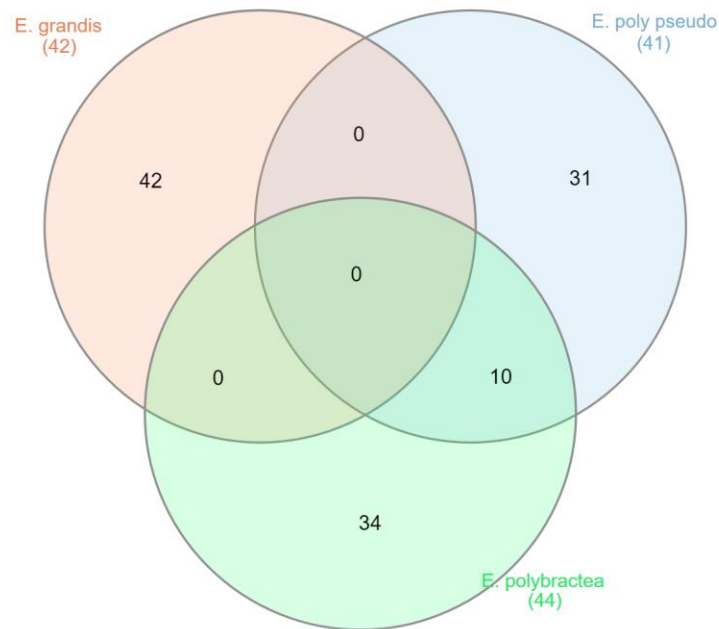## Supplementary Information 3.1

Figure S1. Venn diagram showing the number of significant associations identified from GWAS (FamCPU, BLINK, and MLMM) with *E. grandis*, *E. polybractea* pseudo and *E. polybractea* reference genome.
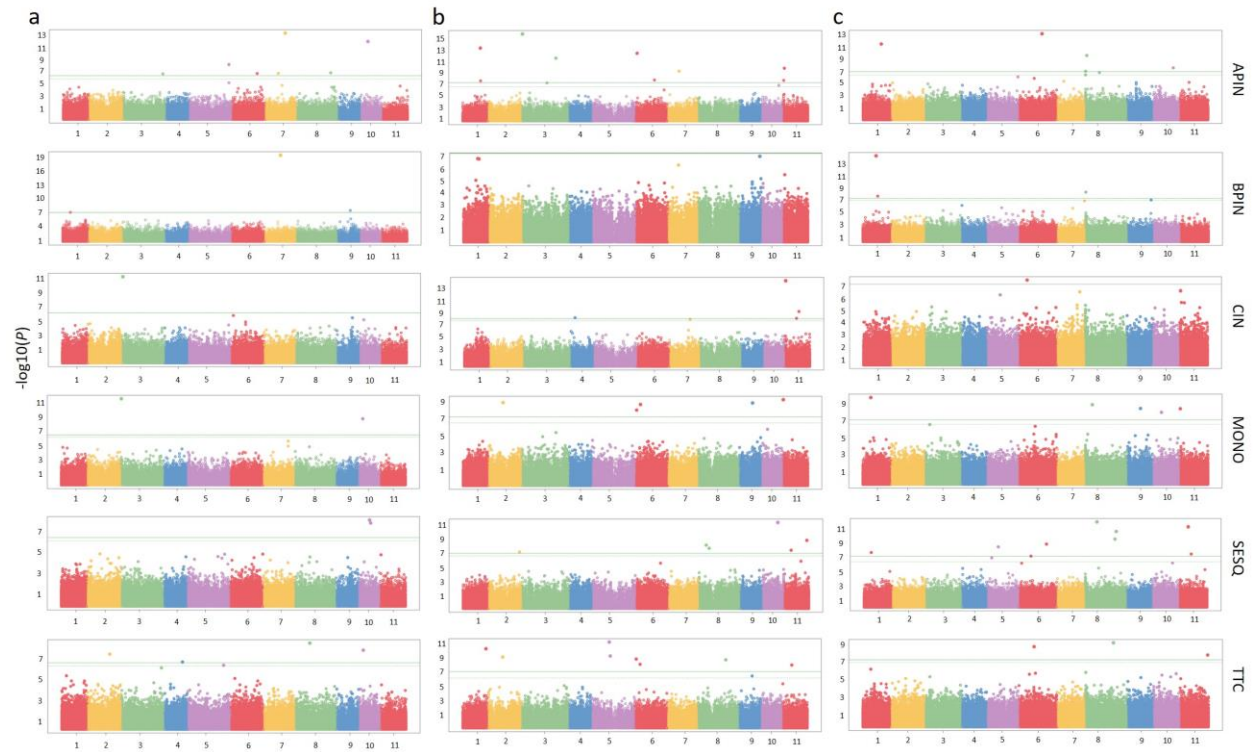
Figure S2. Manhattan plots of genome-wide association studies (GWAS) of six different terpene traits using BLINK model. (a, b, c) *E. grandis*, *E. polybractea* pseudo and *E. polybractea* reference genome mapped GWAS hits using 0.23 m, 0.51 m and 0.58 m single nucleotide polymorphisms (SNPs) in 478 *E. polybractea* samples, respectively. The x-axes showing the 11 chromosomes and y-axes show the -log$_{10}$(P) value.
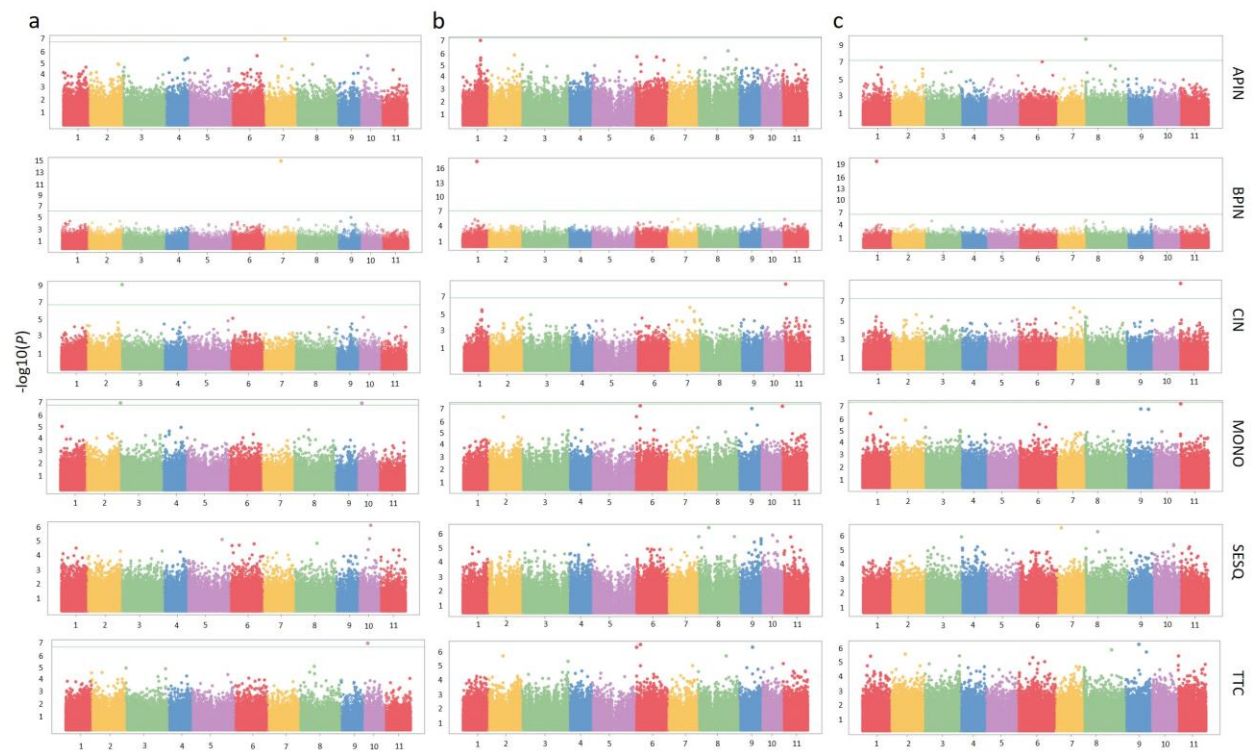
Figure S3. Manhattan plots of genome-wide association studies (GWAS) of six different terpene traits using MLMM model. (a, b, c) *E. grandis*, *E. polybractea* pseudo and *E. polybractea* reference genome mapped GWAS hits using 0.23 m, 0.51 m and 0.58 m single nucleotide polymorphisms (SNPs) in 478 *E. polybractea* samples, respectively. The x-axes showing the 11 chromosomes and y-axes show the -log₁₀(*P*) value.
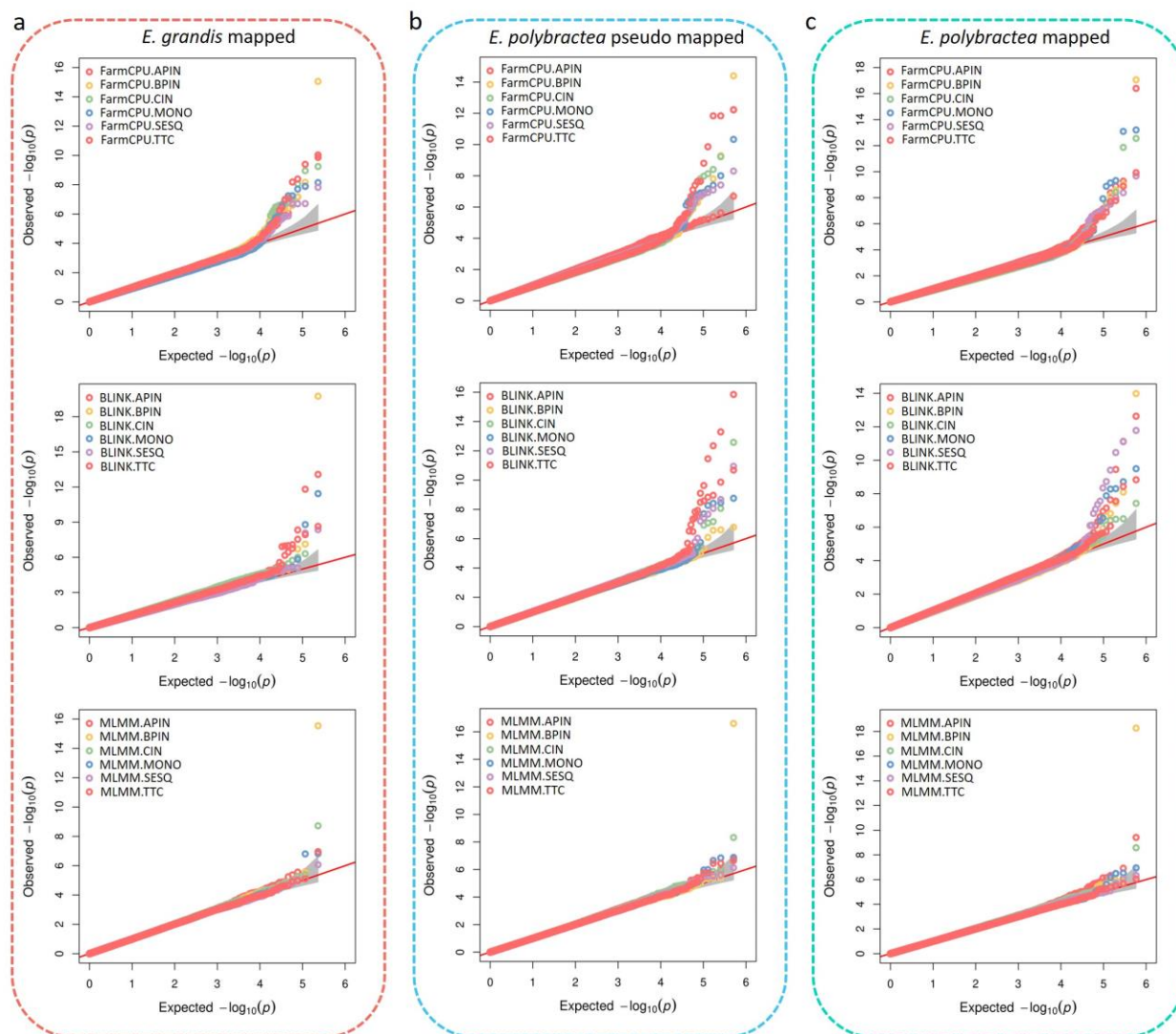
Figure S4. QQ plots of genome-wide association studies (GWAS) of six different terpene traits.
(a, b, c) QQ plots for *E. grandis*, *E. polybractea* pseudo and *E. polybractea* reference genome
mapped GWAS hits using FarmCPU, BLINK, and MLMM model.