## File S1: Captions for Figures S1-S5 and Tables S1-S3

**Genomic-environmental associations in wild cranberry (Vaccinium macrocarpon Ait.)** Jeffrey L. Neyhart, Michael B. Kantar, Juan Zalapa, and Nicholi Vorsa

# Supplemental Figure Captions

**Figure S1.** Q-Q plots comparing the distribution of p-values from the per-marker mixed-model environmental association analysis versus the distribution expected under a null model. Each panel displays the results of the association analysis with an individual environmental variable. Point shapes denote the model used.

**Figure S2.** Boxplots of variables used in the environmental association analysis along with the first 6 principal components (PCs) from a principal component analysis of those environmental variables. Individual colored points indicate the values at each of 17 locations where wild cranberry accessions were sampled.

**Figure S3.** Pairwise correlations between variables used in the environmental association analysis along with the first 6 principal components (PCs) from a principal component analysis of those environmental variables. The upper diagonal presents the estimated correlation coefficient, and the lower diagonal presents a heatmap of those coefficients, where red is a highly positive correlation, and blue is a highly negative correlation. Variables are ordered according to a hierarchical clustering based on the estimated correlation coefficients.

**Figure S4.** Summary of marker alleles from five *F*SToutliers. For each marker, a map depicts the frequency of the minor (blue) or major (orange) allele at each sampling location, where the size of each pie points is proportion to the number of accession collected at that location. If the marker was significantly associated with any environment variables, the distribution of those variables according to marker genotype classes is provided.

**Figure S5.** Summary of marker alleles from five spatial ancestry analysis (SPA)outliers. For each marker, a map depicts the frequency of the minor (blue) or major (orange) allele at each sampling location, where the size of each pie points is proportion to the number of accession collected at that location. If the marker was significantly associated with any environment variables, the distribution of those variables according to marker genotype classes is provided.

# Supplemental Table Captions

**Table S1.** Description of variables used in the environmental association analysis, including the short name, full name, class, and units of each.

**Table S2.** Summary of the principal component analysis of data from 42 environmental variables. For each of the first six principal components (PCs), the proportion of explained variance is listed along with the environmental variable with the highest magnitude loading.

**Table S3**. Summary of significant markers detected in the environmental association (EAA), *F*SToutlier, and spatial ancestry (SPA) analyses. For each significant association, the following information is provided: class of association (EAA-Geog., EAA-Prec., EAA-Temp., EAA-Soil, *F*ST, and SPA); variable full name; marker name, chromosome, and position; marker score (either the -log10 *p*-value score from the EAA, *F*ST estimate, or SPA score); minor allele frequency (MAF) in the wild cranberry germplasm; frequency of that minor allele in landrace (“Native”) and breeding (“Breeding”) germplasm; number of predicted genes within 17 kb of the marker; the name of the closest annotated gene; the distance from the marker to the closest annotated gene; and the name of the high-matching *Arabidopsis* homolog for the closest annotated gene, if any.