**Supplementary information**

Genome-Wide Linkage-Disequilibrium Mapping to the Candidate Gene Level in Melon (*Cucumis melo)*

Amit Gur[1*], Galil Tzuri[1], Ayala Meir[1], Uzi Sa'ar[1], Vitaly Portnoy[1], Nurit Katzir[1], Arthur A. Schaffer[2], Li Li[3], Joseph Burger[1] and Yaakov Tadmor[1]

[1]Plant Science Institute, Agricultural Research Organization, Newe Ya'ar Research Center, P.O. Box 1021, Ramat Yishay 3009500, Israel. [2]Plant Science Institute, Agricultural Research Organization, The Volcani Center, P.O. Box 15159, Rishon LeZiyyon 7507101, Israel. [3]Robert W Holley Center for Agriculture and Health, USDA-ARS, Plant Breeding and Genetics Section, School of Integrative Plant Science, Cornell University, Ithaca, New York 14853, USA.

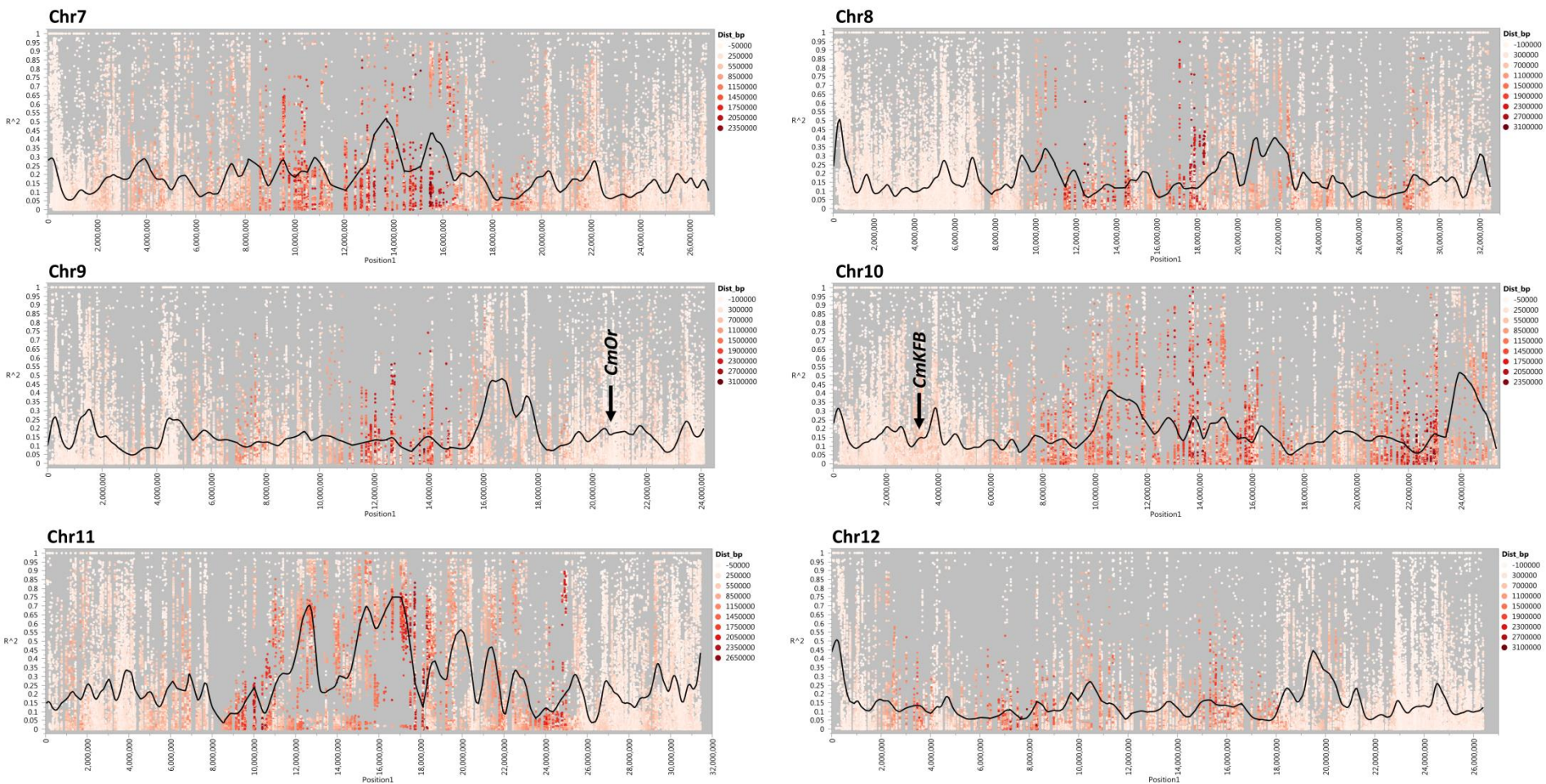*Correspondence should be addressed to A.G. (e-mail: amitgur@volcani.agri.gov.il)
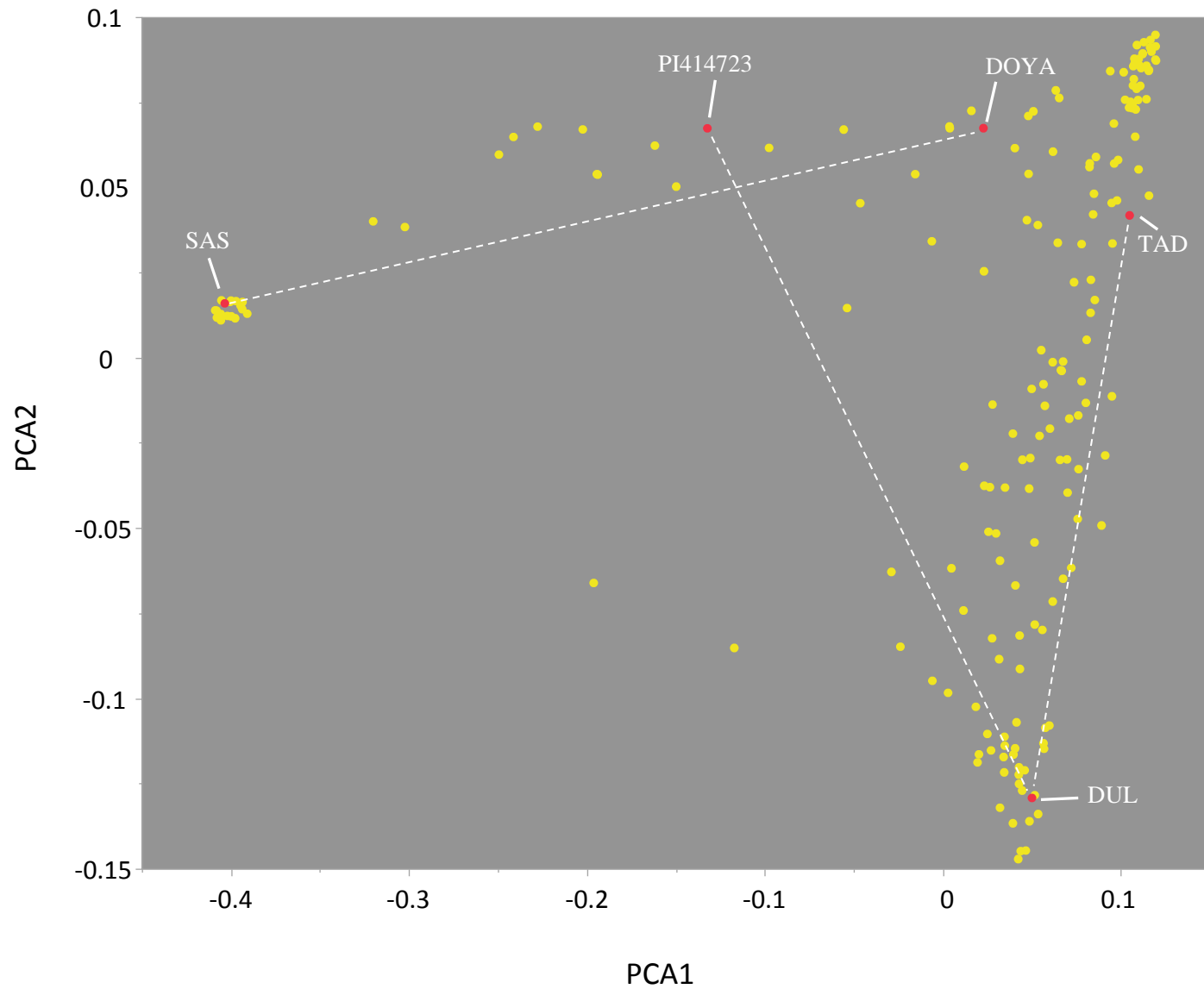
Supplementary Figure S1: **Parental lines of the 3 bi-parental populations highlighted on the genetic PCA analysis**. Crosses are shown in dashed white lines. SAS: C. *melo* ssp. *agrestis*, var *makuwa* . PI414723: C. *melo* ssp. *agrestis*, var *momordica.* DOYA: C. *melo* ssp. *melo*, var *flexuosus.* TADA*: C. melo ssp. melo, var inodorus.* DUL: C. *melo* ssp. *melo*, var *reticulatus*
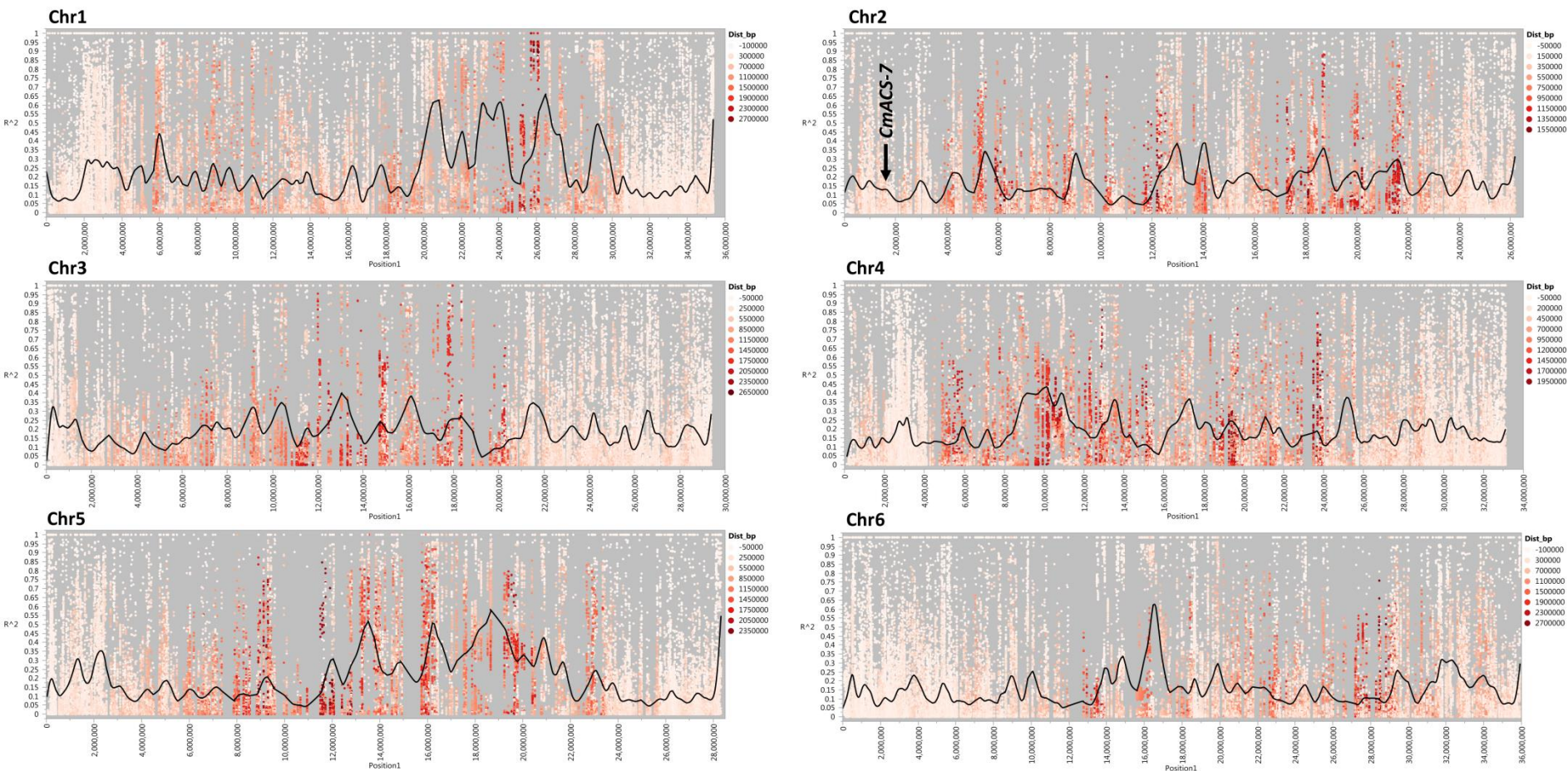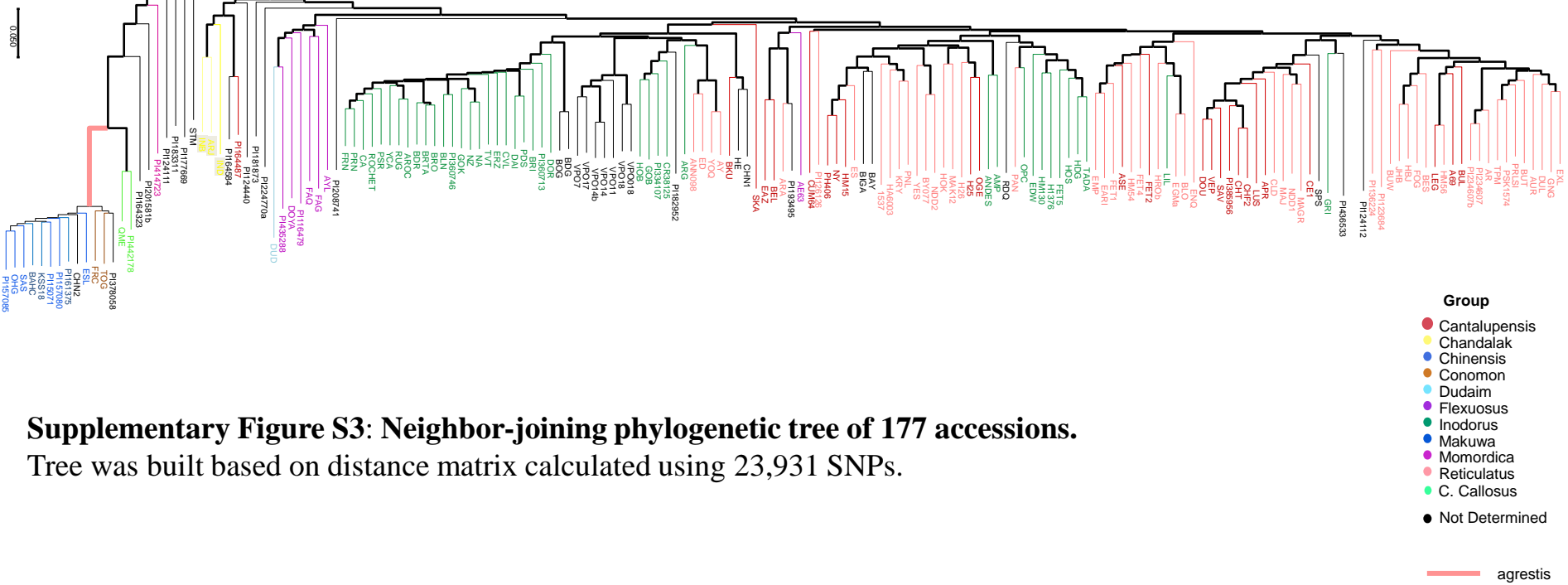
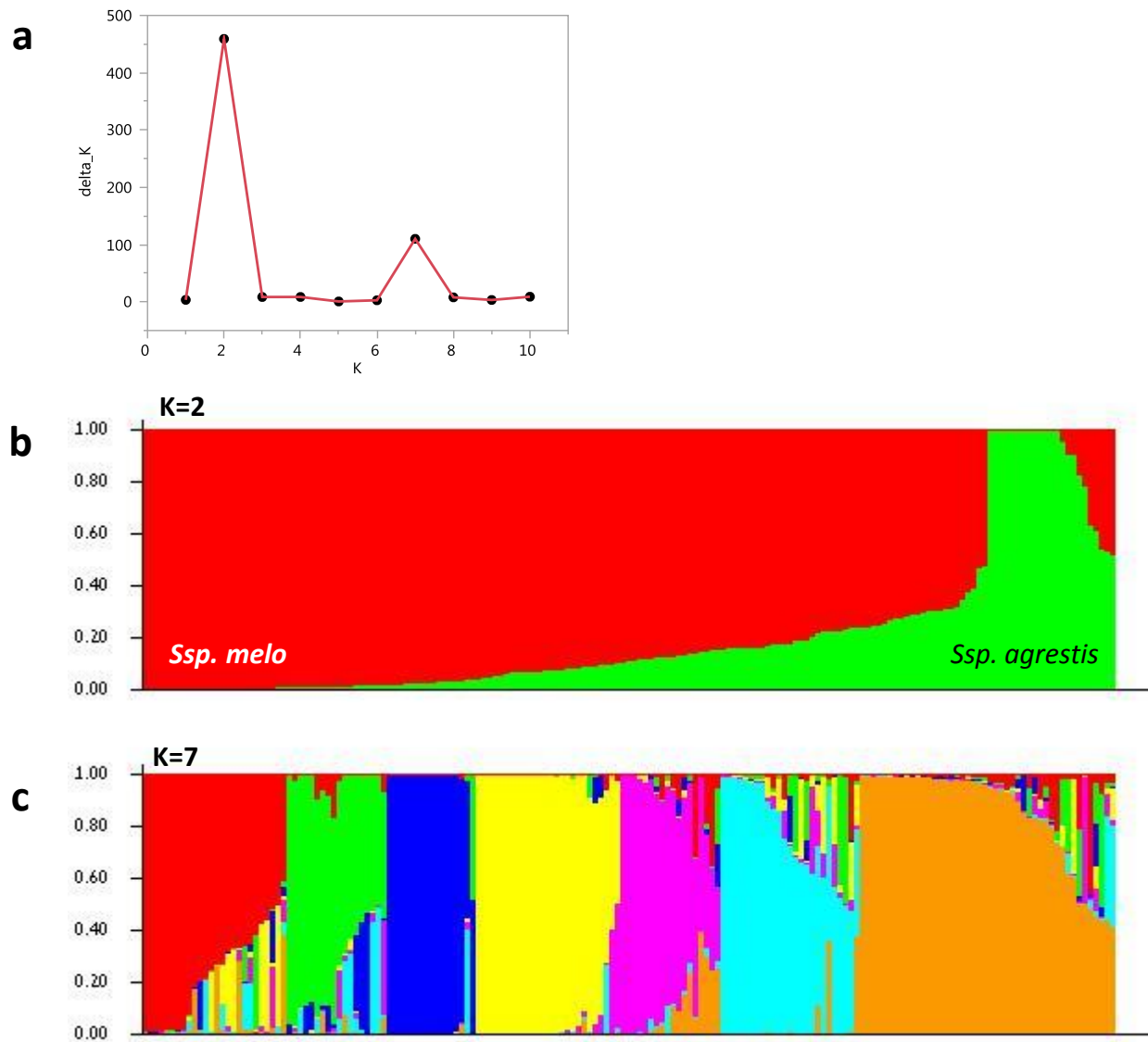**Supplementary Figure S2**: **LD (R^2) by physical position across 12 melon chromosomes.**
LD between intra-chromosomal SNP pairs is plotted (Y axis) by the physical genomic position of first SNP in each pair. Distance between SNPs in each pair is indicated using white-to-red color scale.
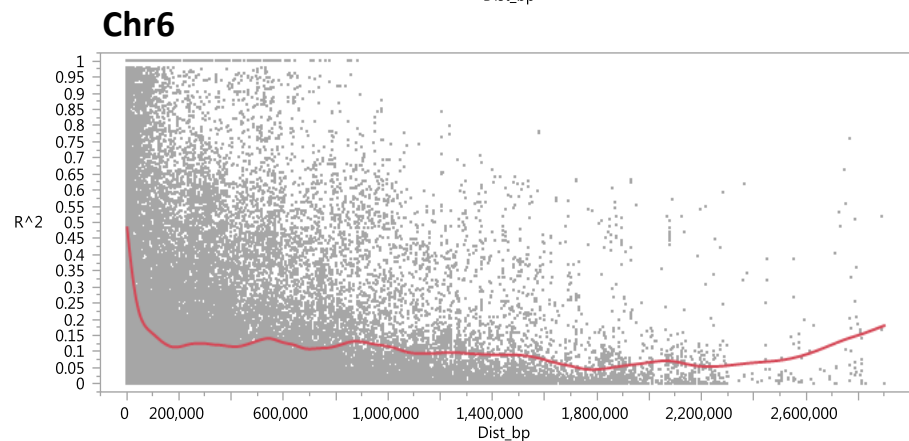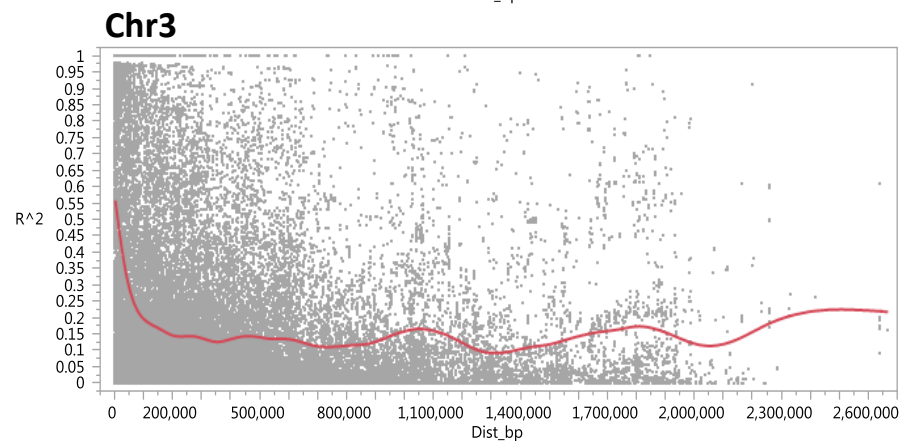
Supplementary Figure S2: **LD (R^2) by physical position across 12 melon chromosomes.**
LD between intra-chromosomal SNP pairs is plotted (Y axis) by the physical genomic position of first SNP in each pair. Distance between SNPs in each pair is indicated using white-to-red color scale.
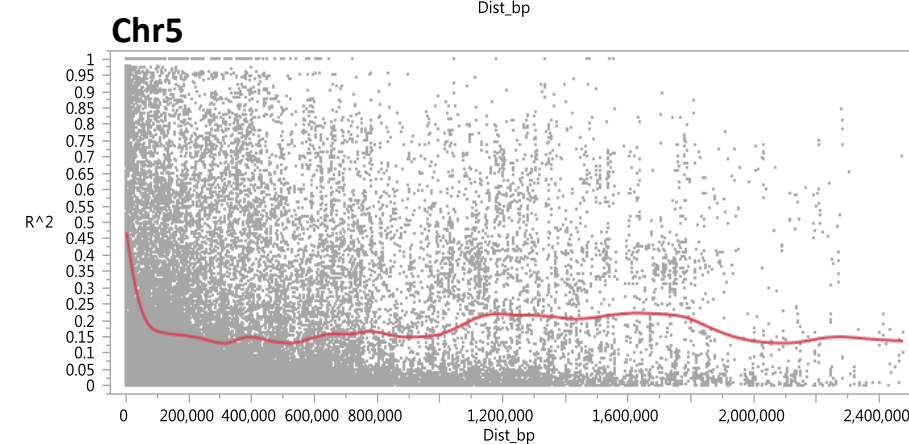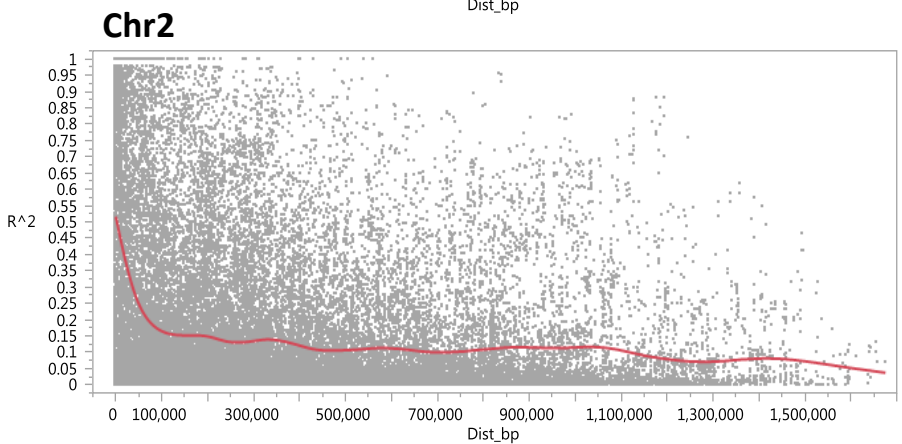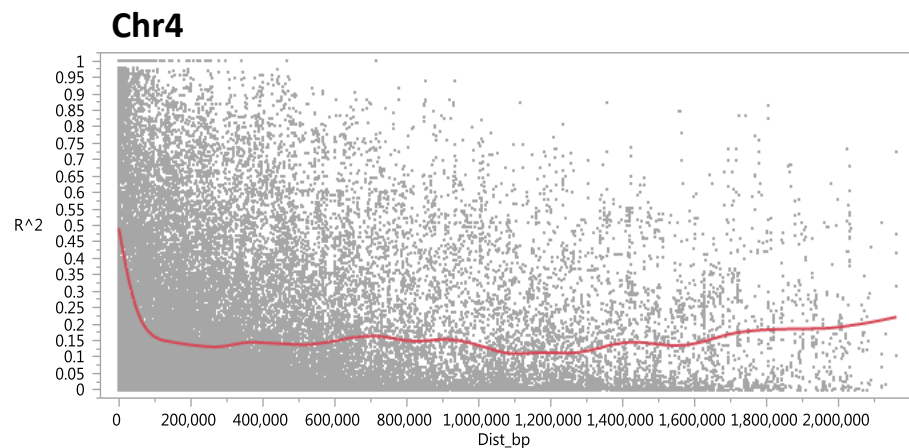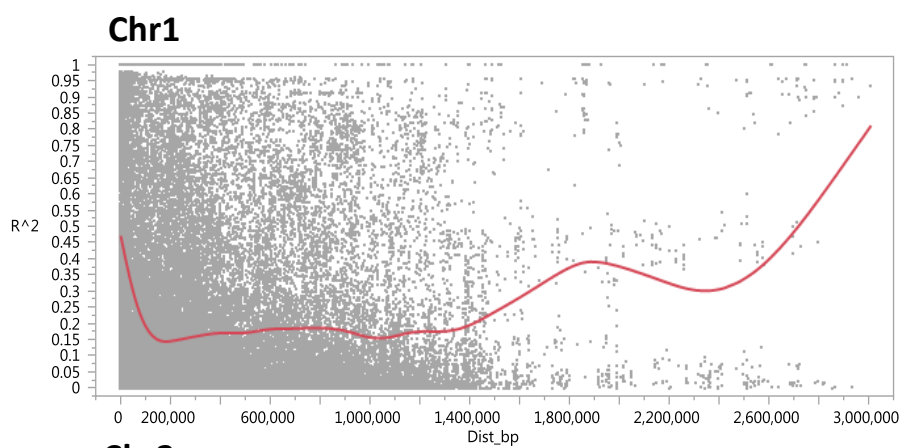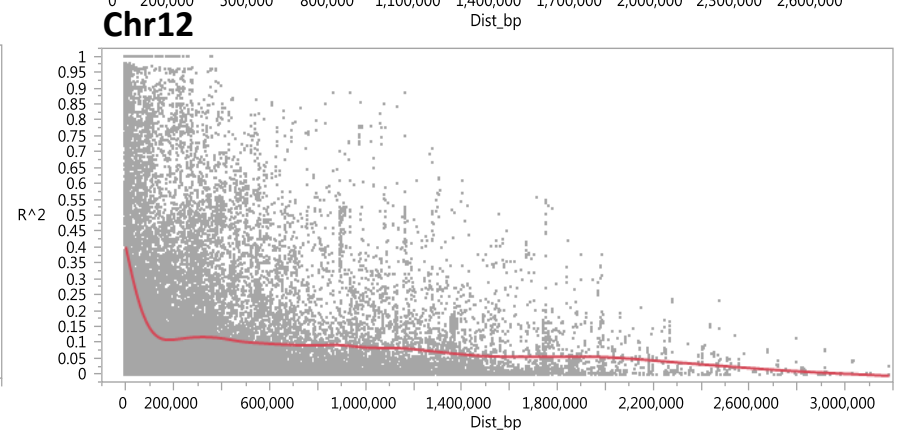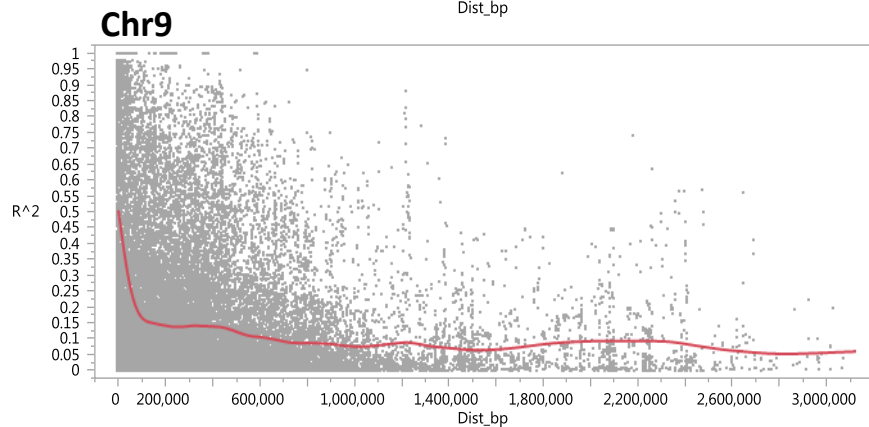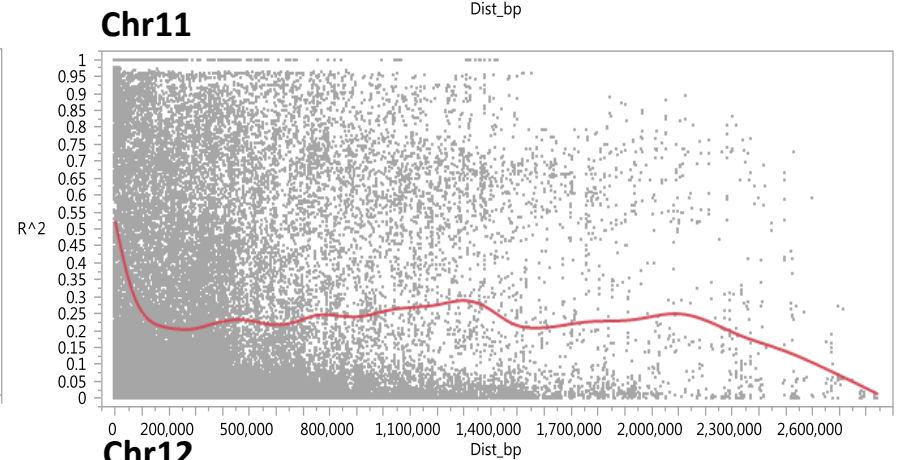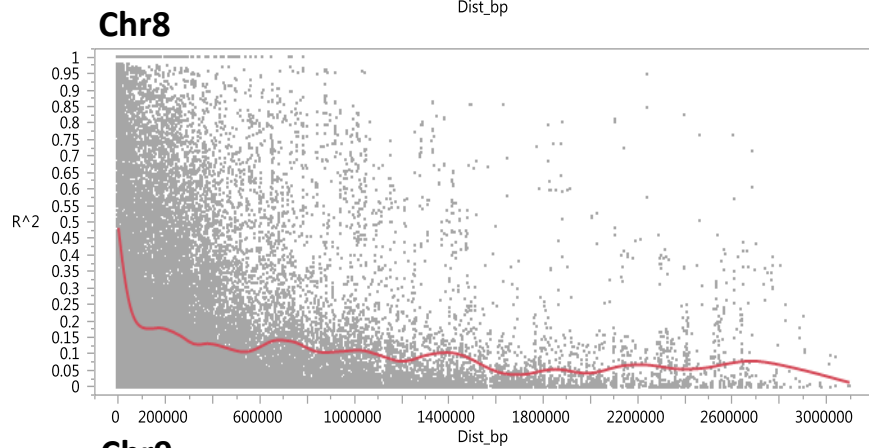
Supplementary Figure S3: **Neighbor-joining phylogenetic tree of 177 accessions.**
Tree was built based on distance matrix calculated using 23,931 SNPs.

Supplementary Figure S4: **Population structure analysis.** Results are shown for the minimum number of subpopulations (K) which sufficiently define genetic variation. Each individual is represented by a vertical line. (**a**) detection of optimal subpopulations using the delat K method (Reference 62). Plot of delta K for K = 1–10. (**b**) Population structure for K=2, in consensus with the ssp. *melo* and ssp *agrestis* division. (**c**) Population structure obtained at K=7.

Supplementary Figure S5a: **LD decay plots by chromosomes across 177 accessions**. R^2 between intra-chromosomal marker pairs plotted against the physical distance between them.

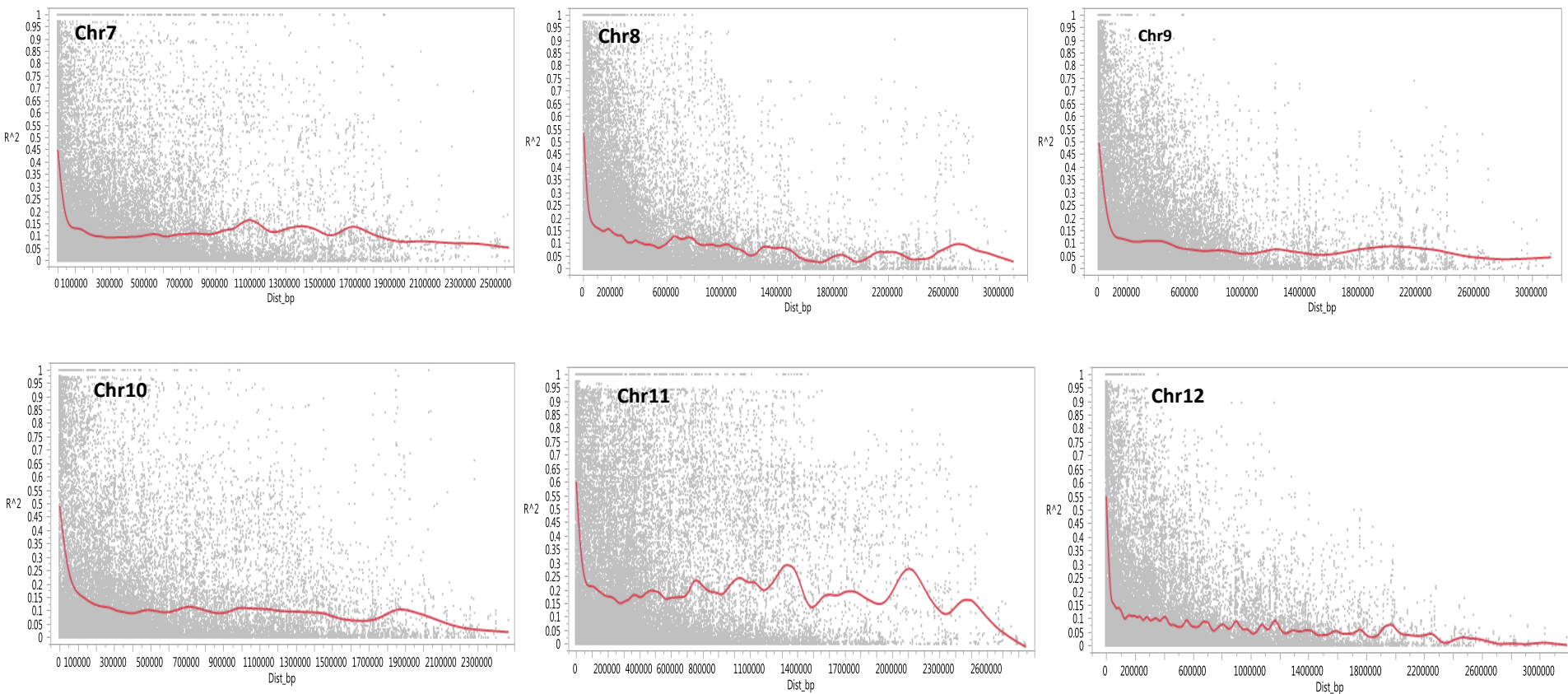**Supplementary Figure S5a**: **LD decay plots by chromosomes across 177 accessions**. R^2 between intra-chromosomal marker pairs plotted against the physical distance between them.

Supplementary Figure S5b: **LD decay plots by chromosomes within the *ssp. melo* group**. R^2 between intra-chromosomal marker pairs plotted against the physical distance between them.

Supplementary Figure S5b: **LD decay plots by chromosomes within the *ssp. melo* group.** R^2 between intra-chromosomal marker pairs plotted against the physical distance between them.

Supplementary Figure S5c: **LD decay plots by chromosomes within the *ssp. agrestis* group**. R^2 between intra-chromosomal marker pairs plotted against the physical distance between them.
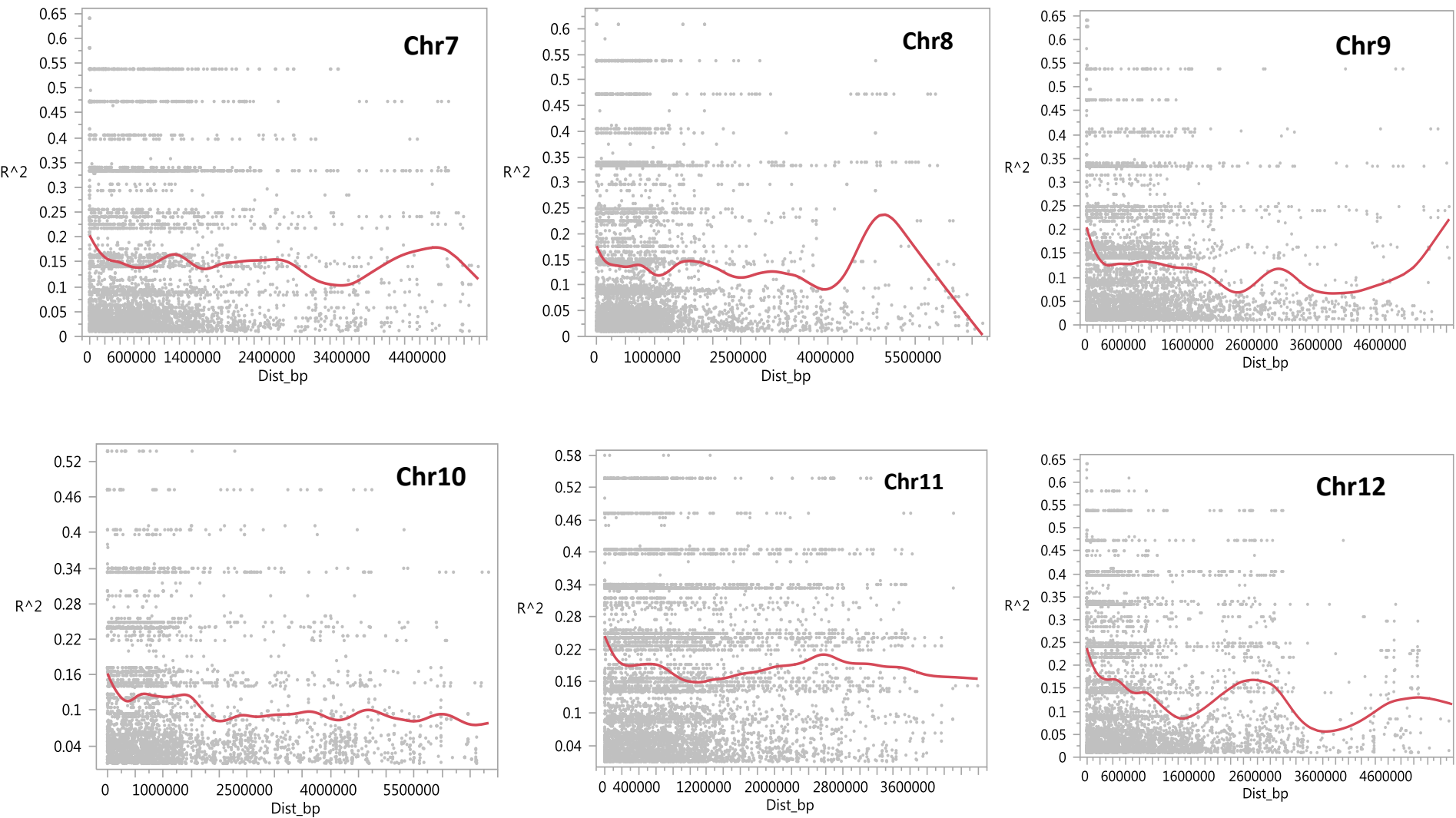
Supplementary Figure S5c: **LD decay plots by chromosomes within the *ssp. agrestis* group**. R^2 between intra-chromosomal marker pairs plotted against the physical distance between them.
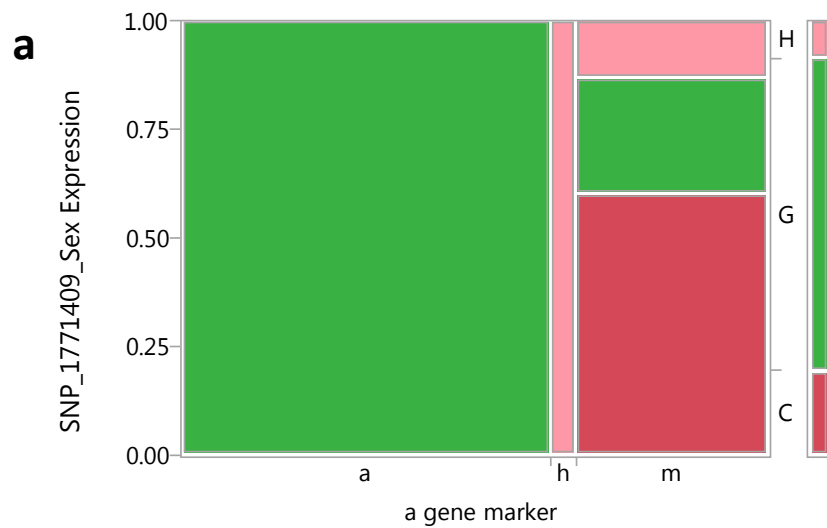
a

SNP_1771409_Sex Expression

a gene marker

SNP_1771409_Sex Expression

b

a gene marker

| Count Total % Col% Row %	C	G	H	Total
a	0	29	0	29
	0.00	63.04	0.00	63.04
	0.00	87.88	0.00	
	0.00	100.00	0.00	
h	0	0	2	2
	0.00	0.00	4.35	4.35
	0.00	0.00	50.00	
	0.00	0.00	100.00	
m	9	4	2	15
	19.57	8.70	4.35	32.61
	100.00	12.12	50.00	
	60.00	26.67	13.33	
Total	9	33	4	46
	19.57	71.74	8.70	

Figure S6: **Validation of SNP_1771409 for flower sex-expression through alignment with PCR marker at the *CmACS-7* gene** (MELO3C015444, Boualem et al. 2008). (**a**) Mosaic plot of segregation of the two markers. (**b**) contingency table of segregation data
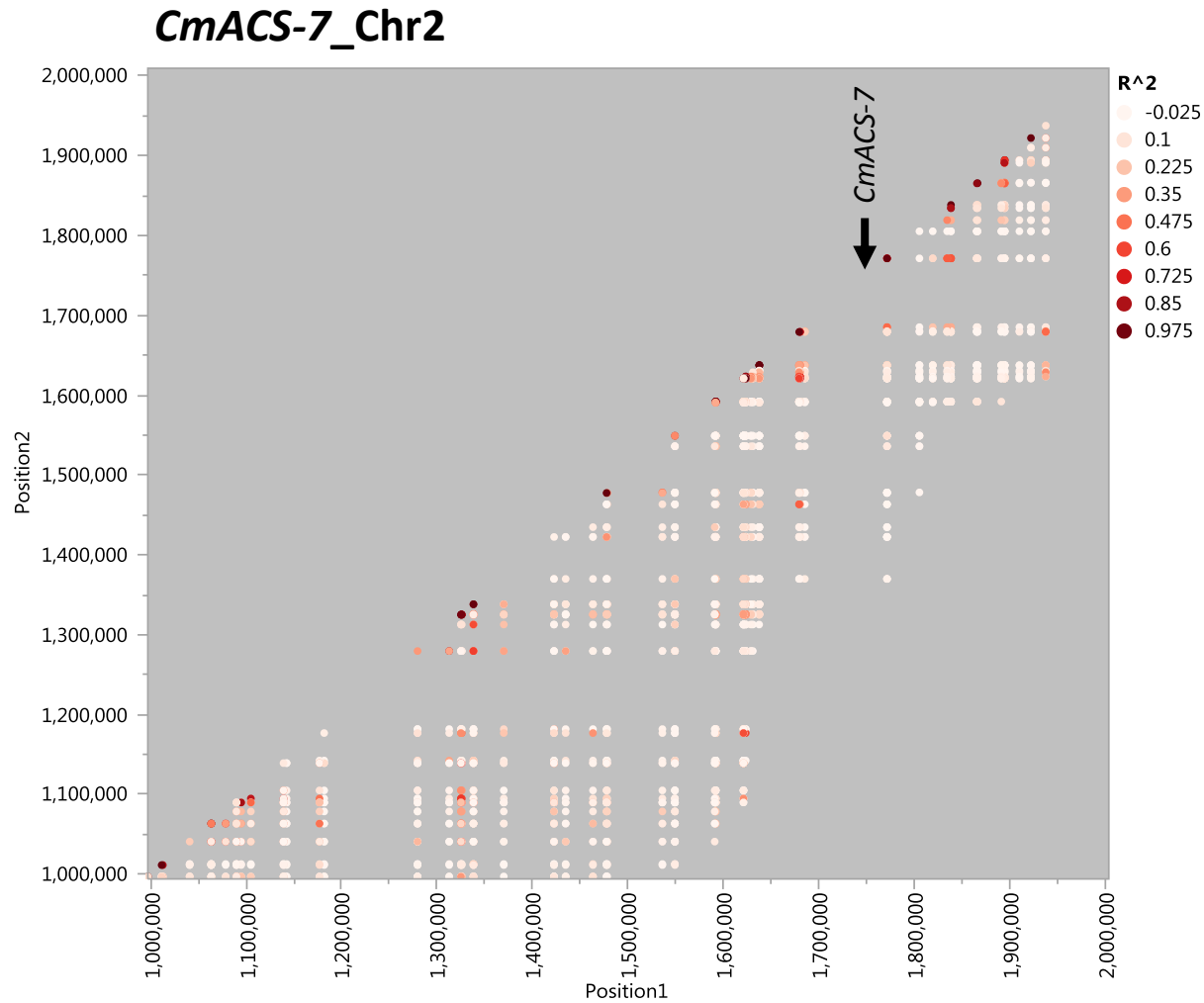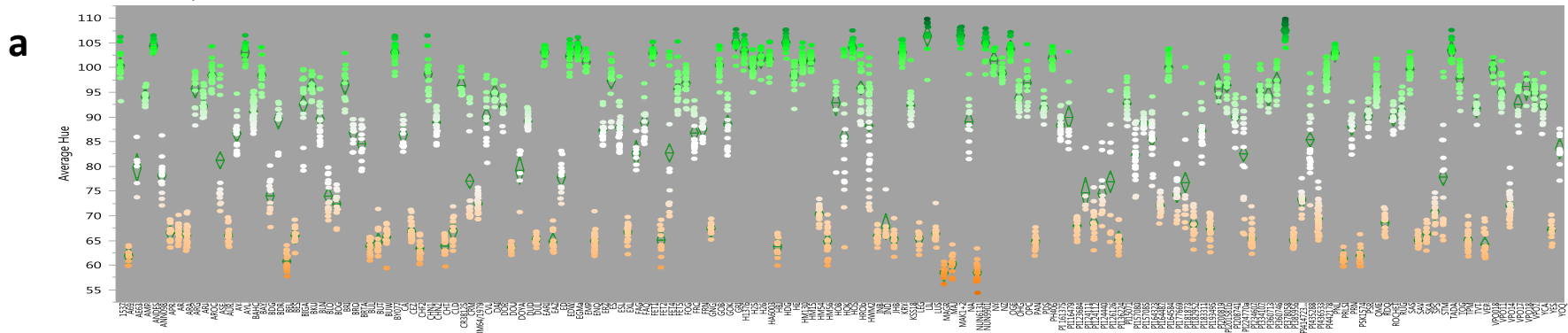
Supplementary Figure S7: **LD distribution plot in a 1,000,000 bp interval surrounding the *CmACS-7* gene on chromosome 2**.

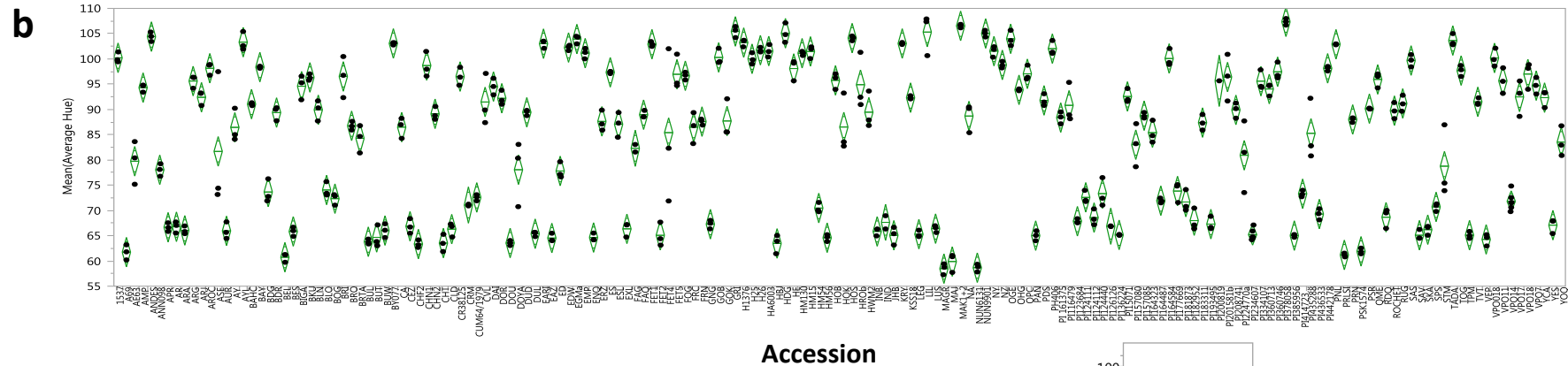LD is expressed as R2 and color-coded using white-to-Red scale.
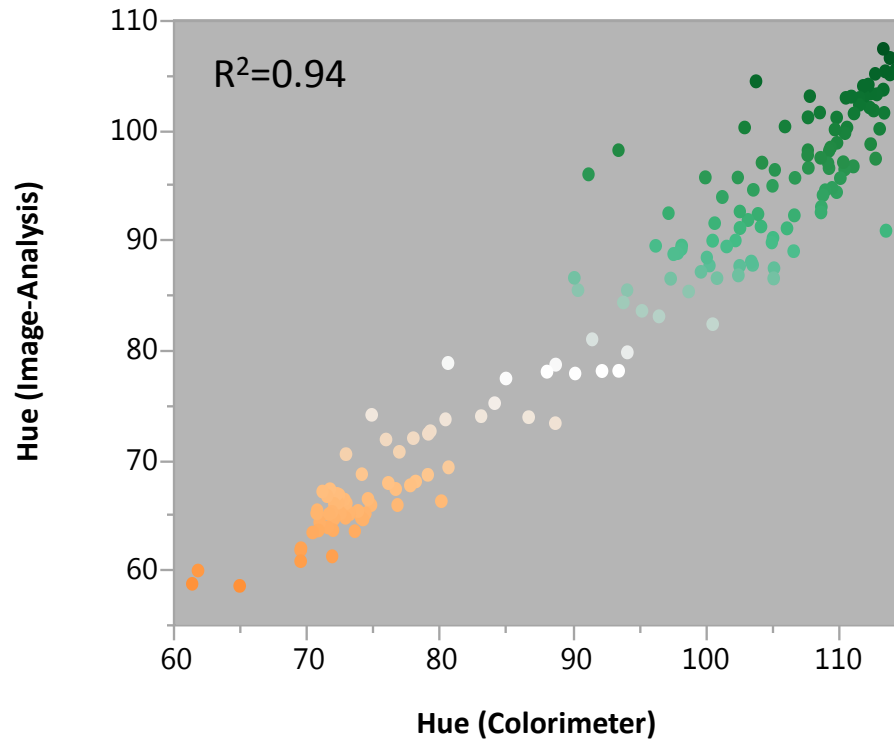
Supplementary Figure S8: **Correlation between fruit flesh hue, measured using colorimeter (X axis) and through fruit image analyses (Y axis).**

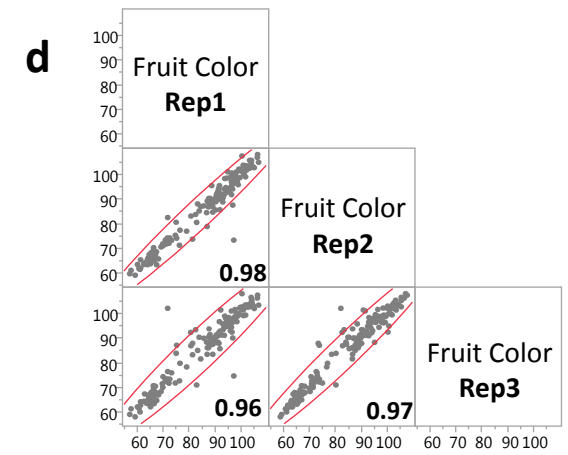**Supplementary Figure S9**: **Properties of flesh color variation .** (a) one-way ANOVA at the fruit basis. (b) one-way ANOVA at the plot basis. (c) variation components table for the fruit basis ANOVA. (d) correlation between replications for plot means.
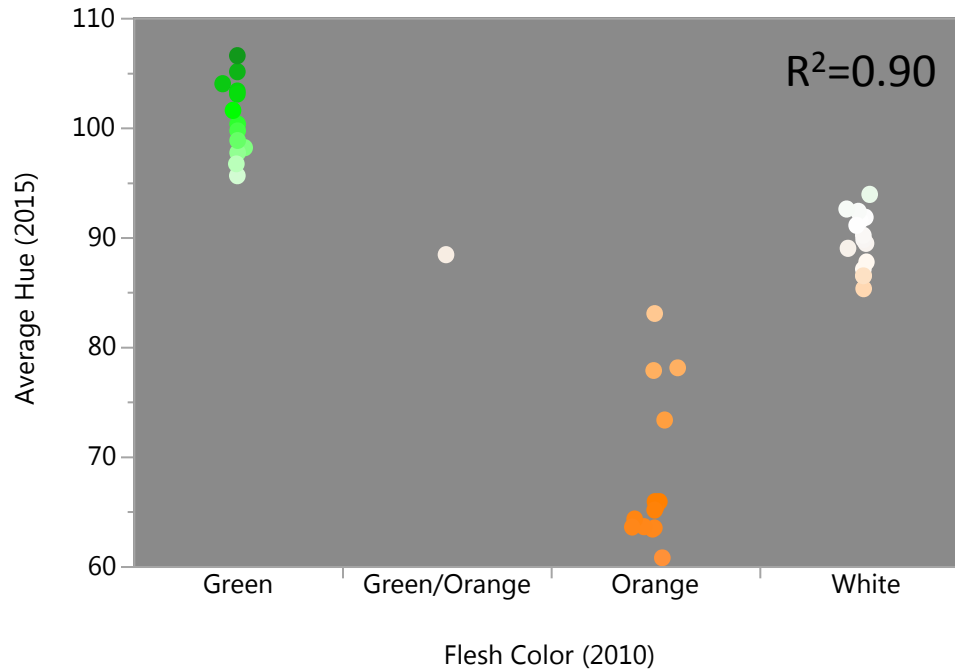
Supplementary Figure S10: **repeatability of flesh color in two different years across 43 diverse accessions.**

Each point represent accession mean. X axis are flesh color classes scores on the Summer 2010 experiment. On the Y axis are flesh color quantitative data collected from digital fruit images on Summer 2016 (see materials and methods).

a

SNP_20550439_melo3C005449 (y-axis)

CmOr SNP (resequencing) — Not Orange, Orange
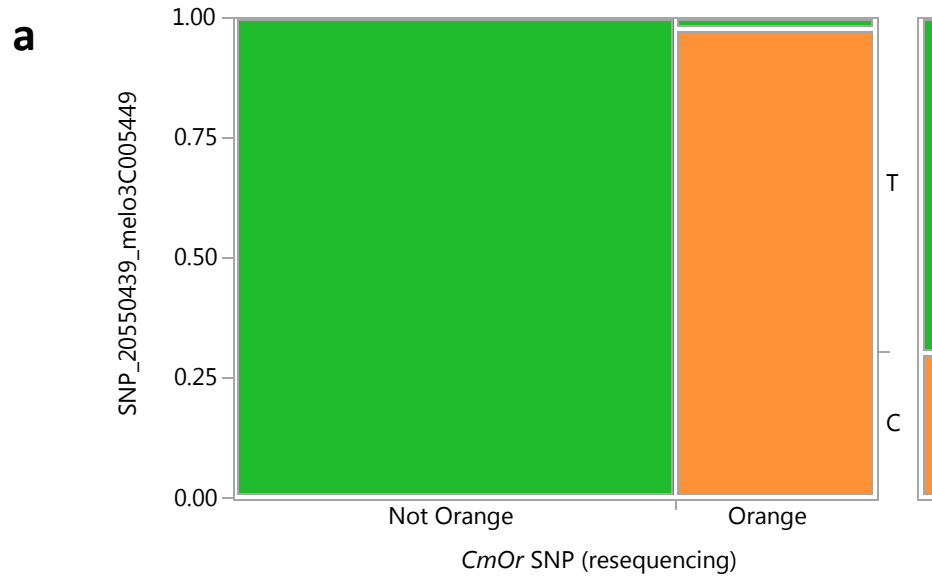
SNP_20550439_melo3C005449

b

CmOr SNP (resequencing)

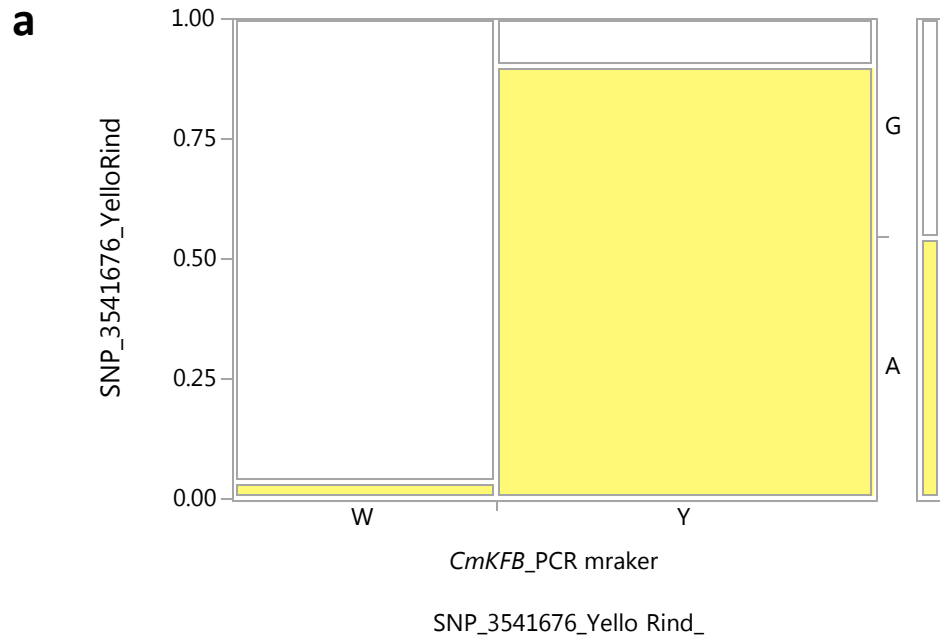| Count Total % Col % Row %	C	T	Total
Not Orange	0 0.00 0.00 0.00	91 68.94 98.91 100.00	91 68.94
Orange	40 30.30 100.00 97.56	1 0.76 1.09 2.44	41 31.06
Total	40 30.30	92 69.70	132

Figure S11: **Validation of SNP_20550439 for flesh color through alignment with resequencing-based SNP at the *CmOr* gene** (melo3C005449, Tzuri et al. 2015). (**a**) Mosaic plot of segregation of the two markers. (**b**) contingency table of segregation data
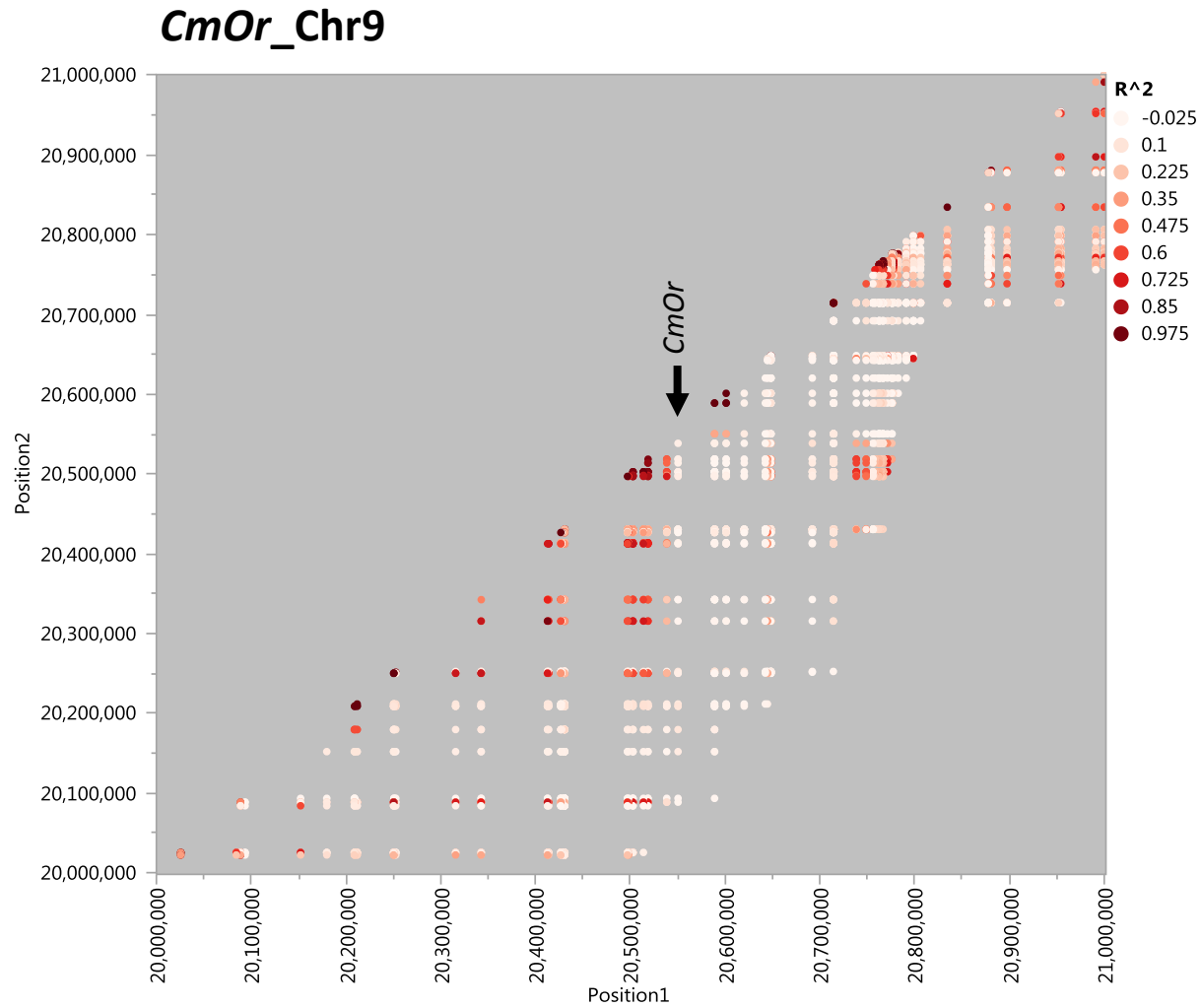
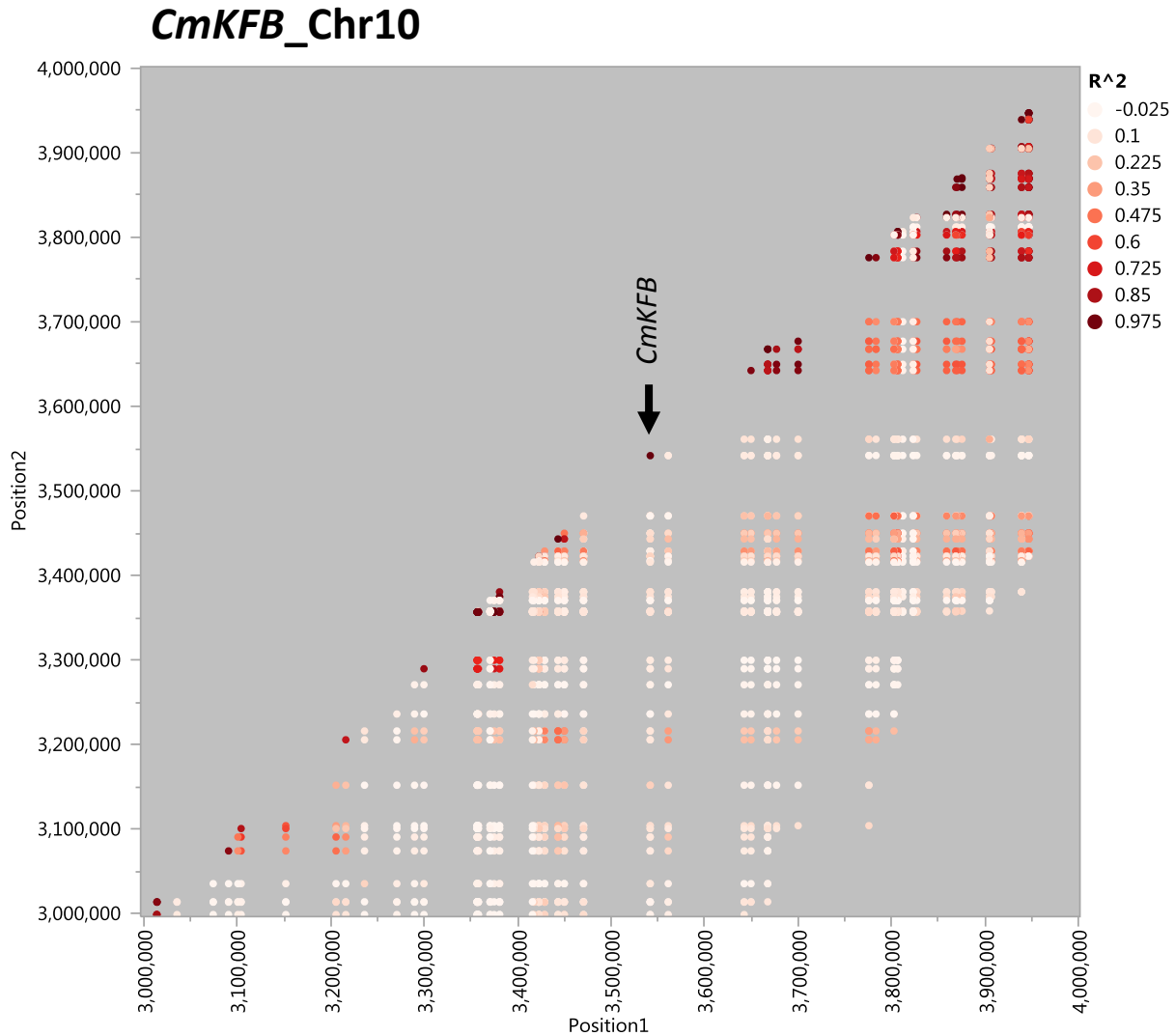**Supplementary Figure S12**: **LD distribution plot in a 1,000,000 bp interval surrounding the *CmOr* gene on chromosome 9.**
LD is expressed as R2 and color-coded using white-to-Red scale.

a

SNP_3541676_YelloRind

*CmKFB*_PCR mraker

SNP_3541676_Yello Rind_

b

Count Total % Col % Row %	A	G	Total
W	1 1.47 2.70 3.57	27 39.71 87.10 96.43	28 41.18
Y	36 52.94 97.30 90.00	4 5.88 12.90 10.00	40 58.82
Total	37 54.41	31 45.59	68

*CmKFB*_PCR Marker

Figure S13: **Validation of SNP_3541676 for flesh color through alignment with PCR marker scoring an INDEL at the *CmKFB* gene** (MELO3C11980, Feder et al. 2015). (**a**) Mosaic plot of segregation of the two markers. W=White Allele, Y=Yellow allele. (**b**) contingency table of segregation data

Supplementary Figure S14: LD distribution plot in a 1,000,000 bp interval surrounding the *CmKFB* gene on chromosome 10.
LD is expressed as R2 and color-coded using white-to-Red scale.
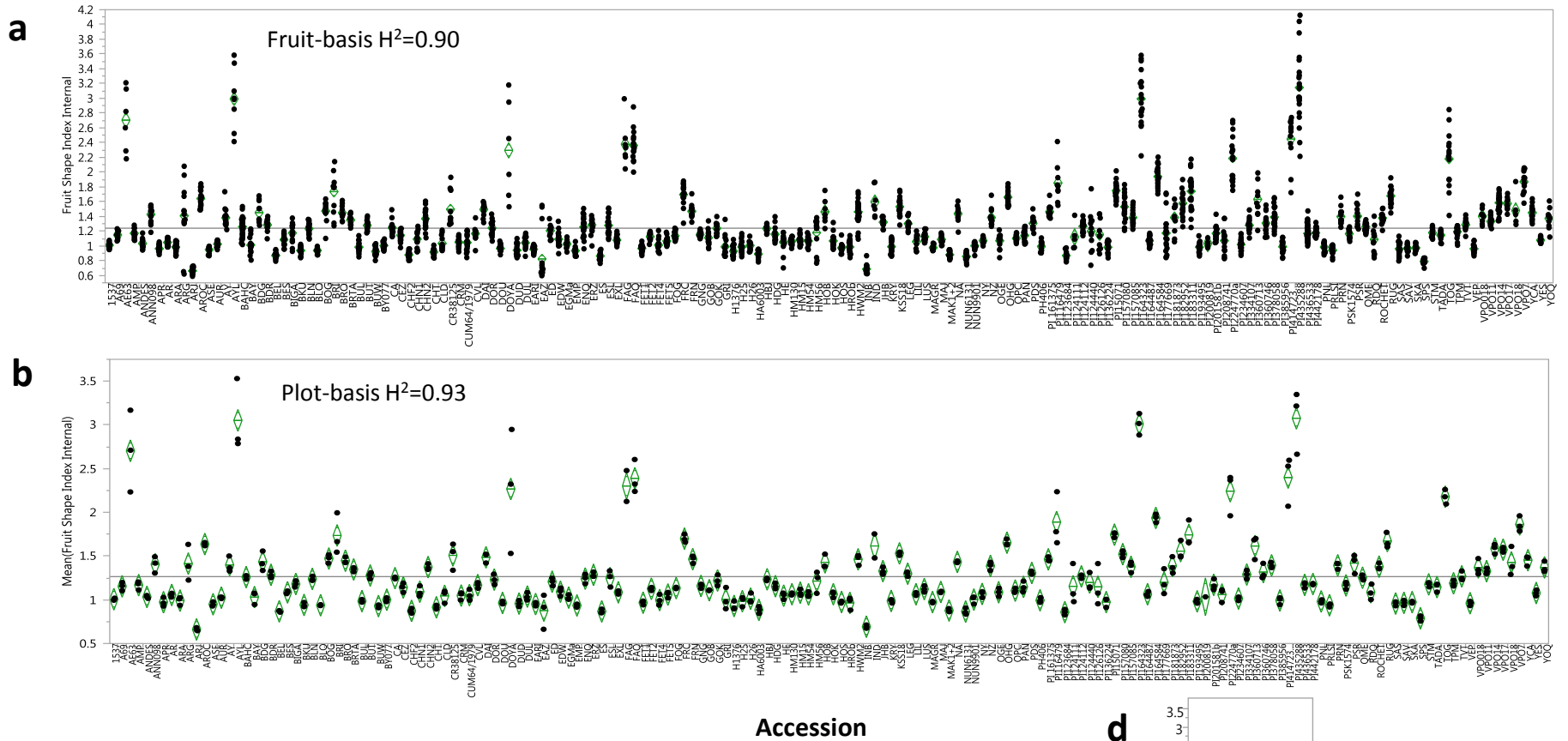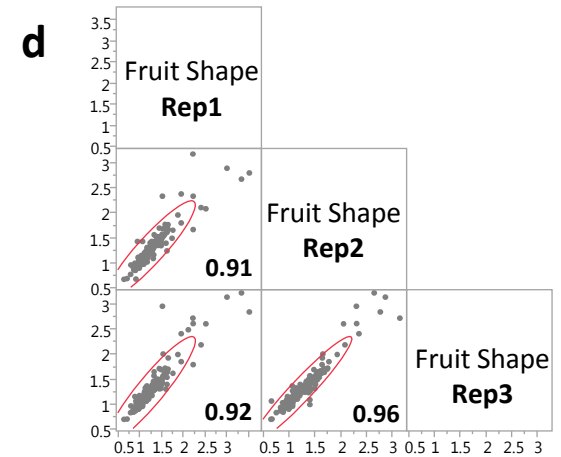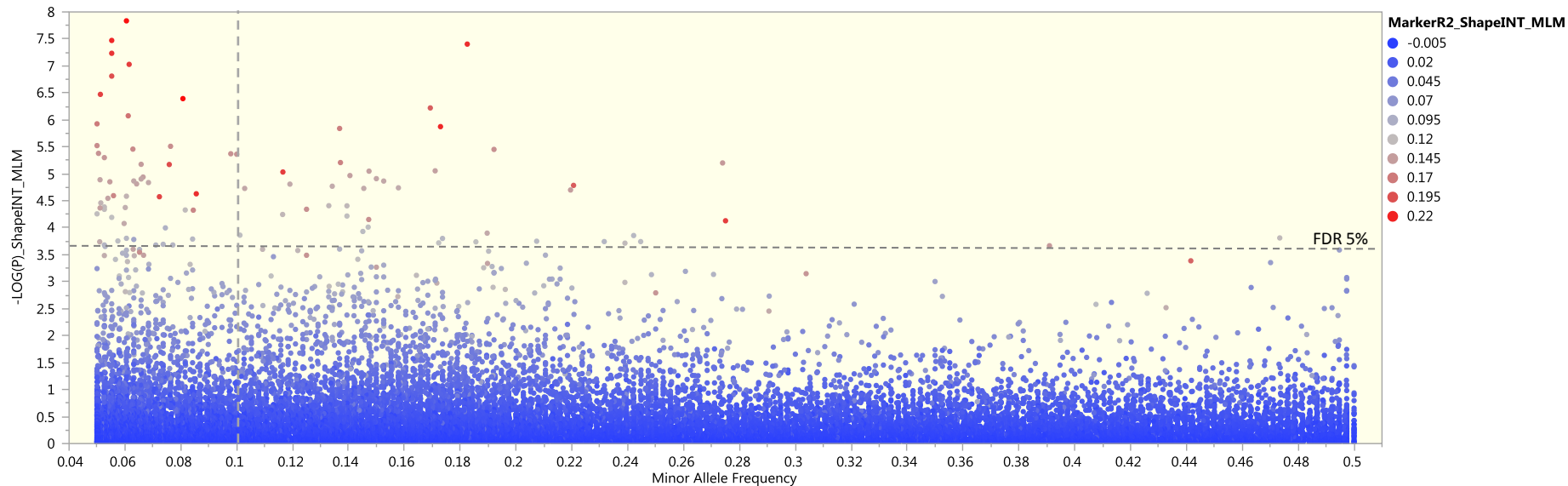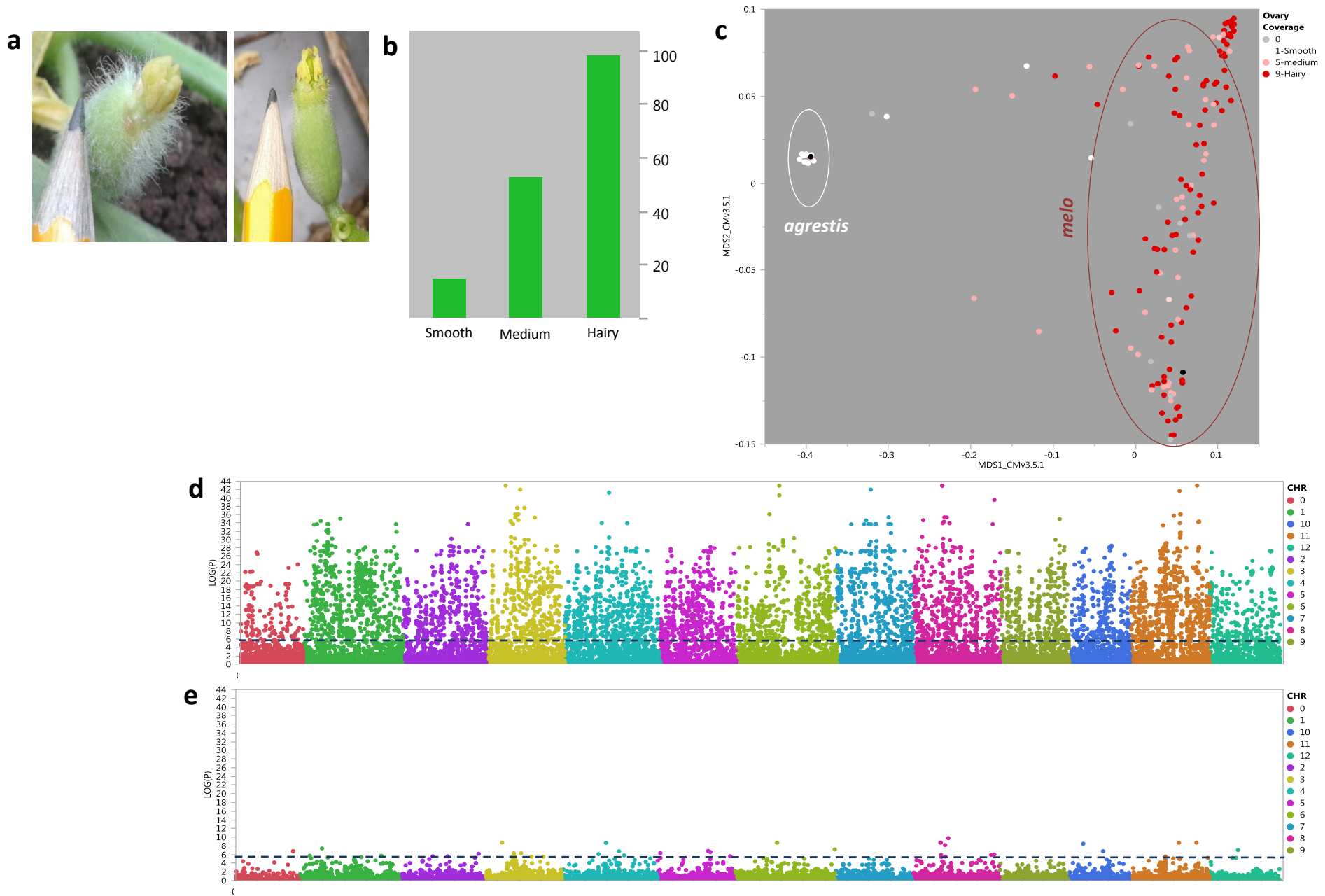
Supplementary Figure S15: **Properties of fruit shape variation .** (a) one-way ANOVA at the fruit basis. (b) one-way ANOVA at the plot basis. (c) variation components table for the fruit basis ANOVA. (d) correlation between replications for plot means.
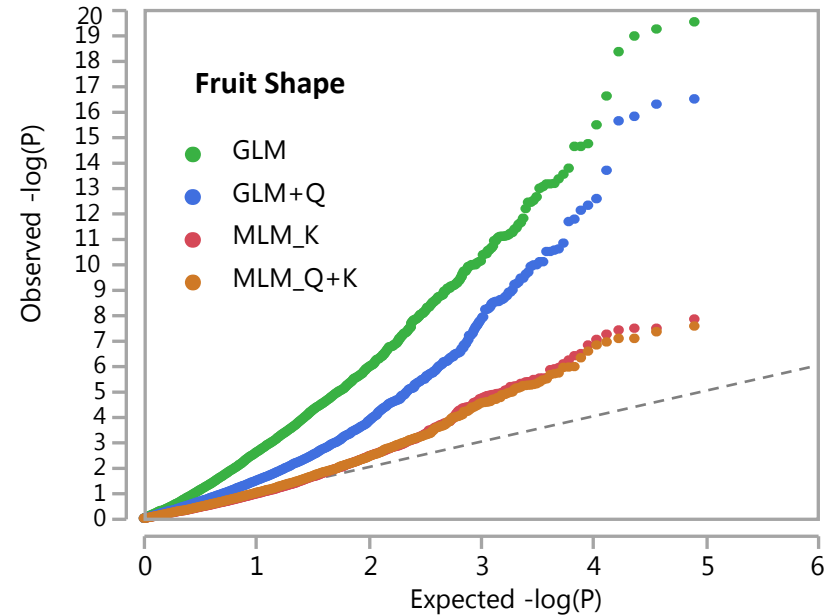
Supplementary Figure S16: Polymorphism effect size for fruit shape index and allele frequencies. SNP allele frequencies plotted against effects on fruit shape index (expressed as -Log P values). Effects are also colored based on SNP R^2.

A

99,263 un-filtered SNPs

MAF

B

29,931 filtered SNPs

MAF
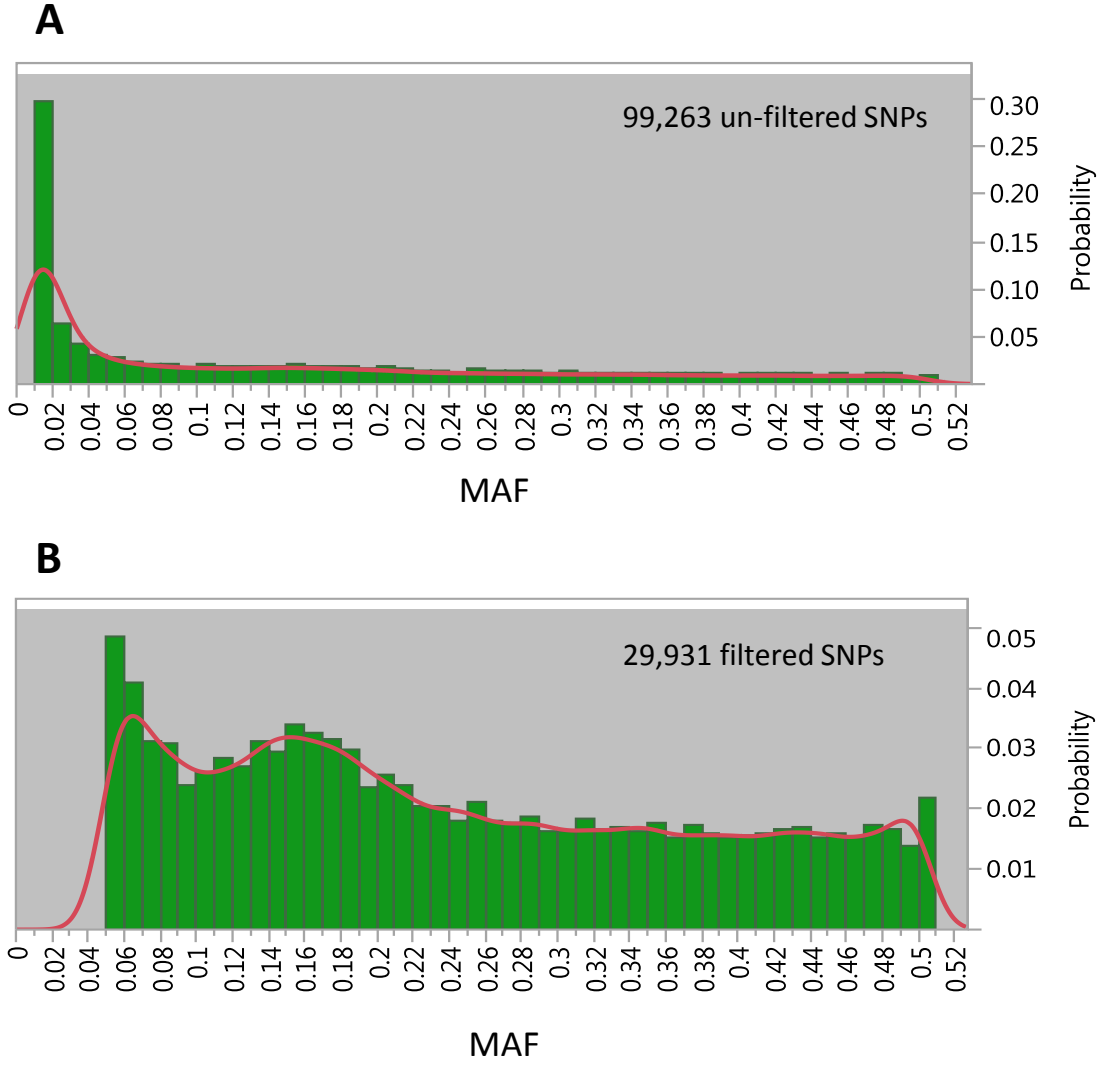
Supplementary Figure S17: **Frequency distribution of minor allele frequencies (MAF).**
(A) Full set – 99,263 SNPs. (B) Filtered set - 23,931 SNPs.

Supplementary Figure S18: **Characterization of ovary hairiness across the diversity panel** – (A) example of spreading hairs (left) and appressed hairs on ovaries at anthesis. (B) Frequency distribution of hairiness classes. (C) Ovary hairiness projected on genetic PCA plot. (D) Manhattan plots of GWA of ovary hairiness. GLM. (E) Manhattan plots of GWA of ovary hairiness. MLM.
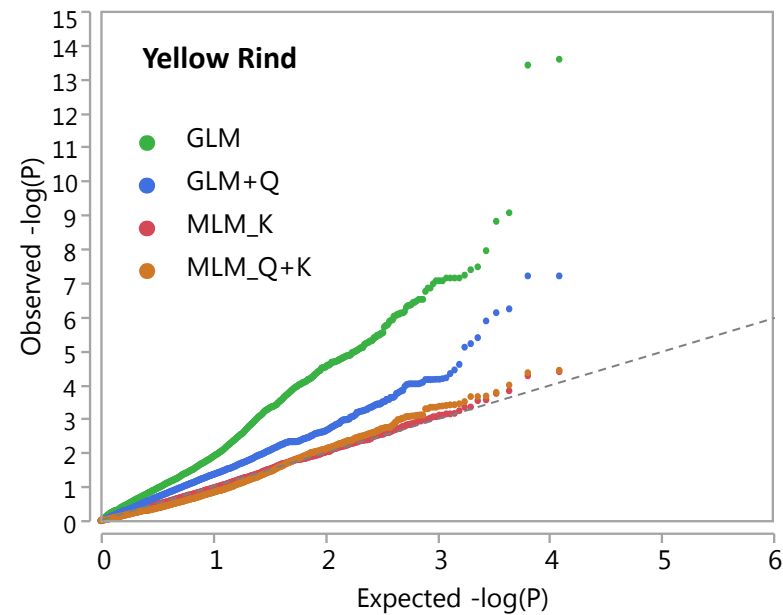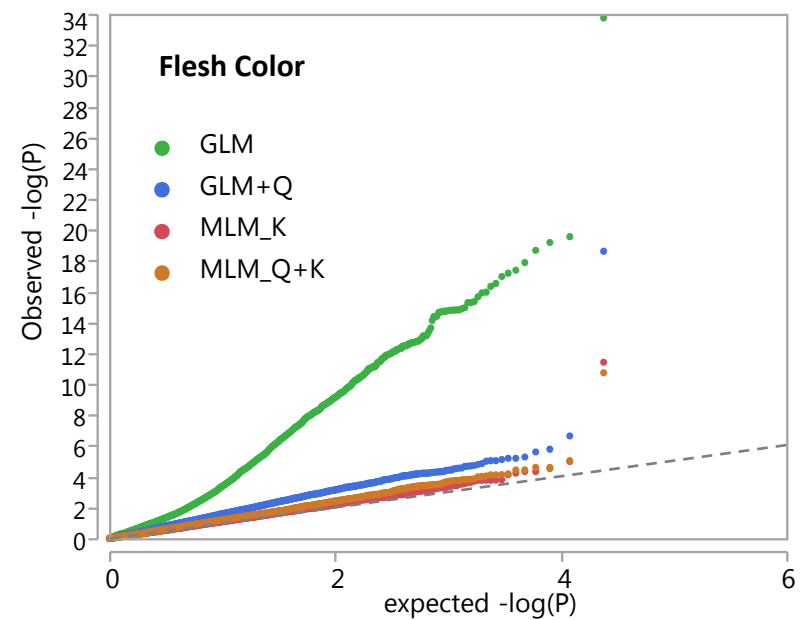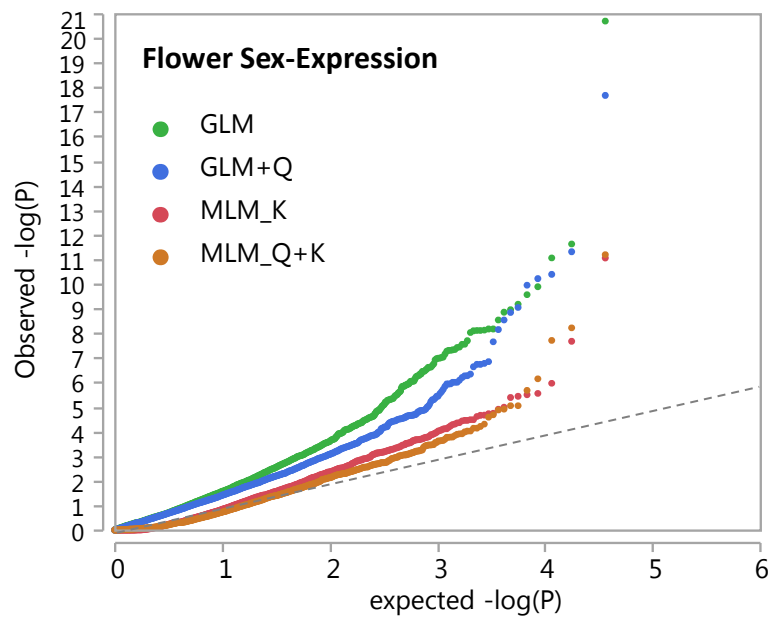
Supplementary Figure S19: **Quantile-quantile (Q-Q) plots comparing distribution of P values at the 4 statistical models used for GWA analyses**. The negative logarithm of the observed (y axis) and the expected (x axis) P value is plotted for each SNP (dot), and the gray dashed line indicates the null hypothesis of no true association.