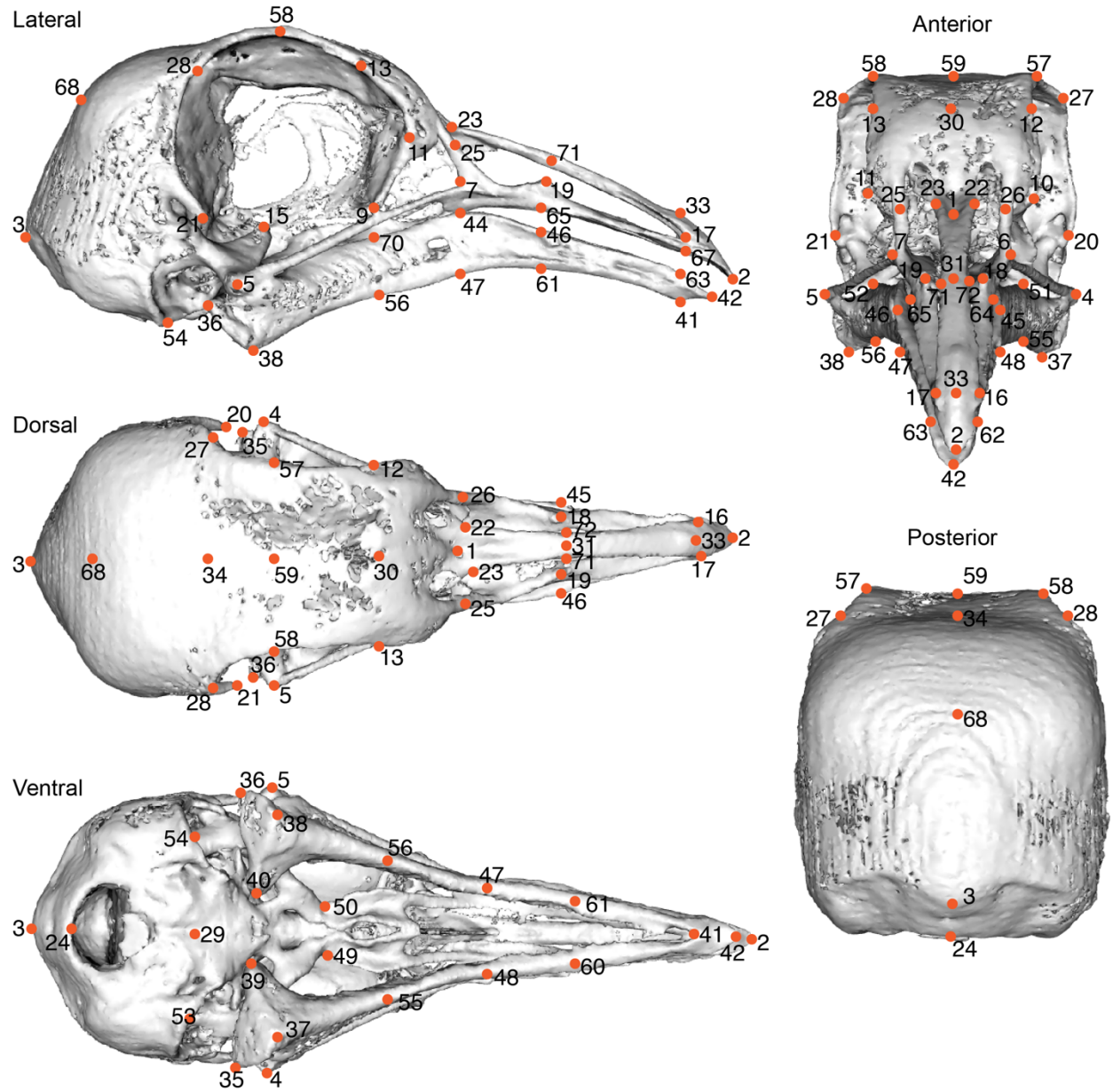
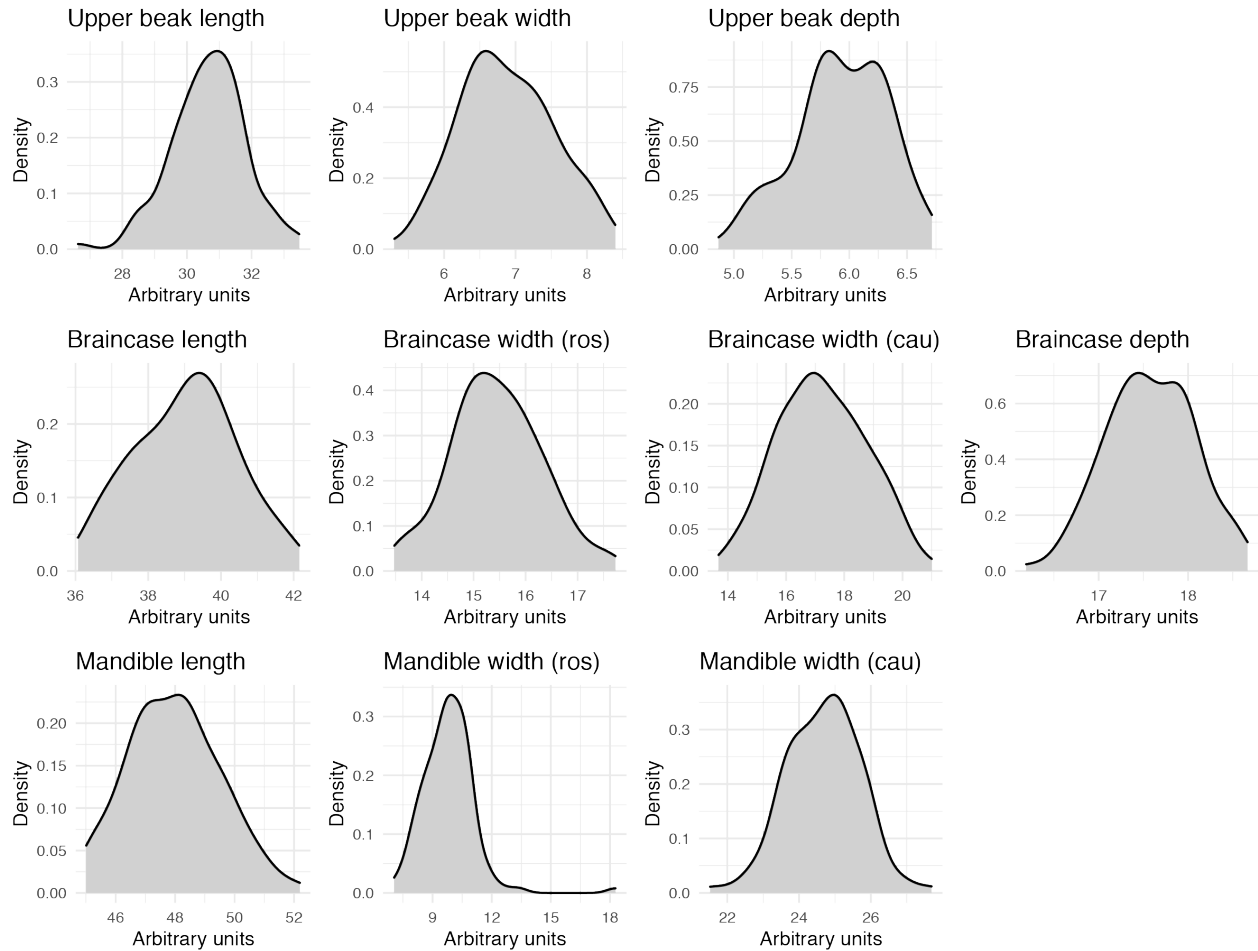


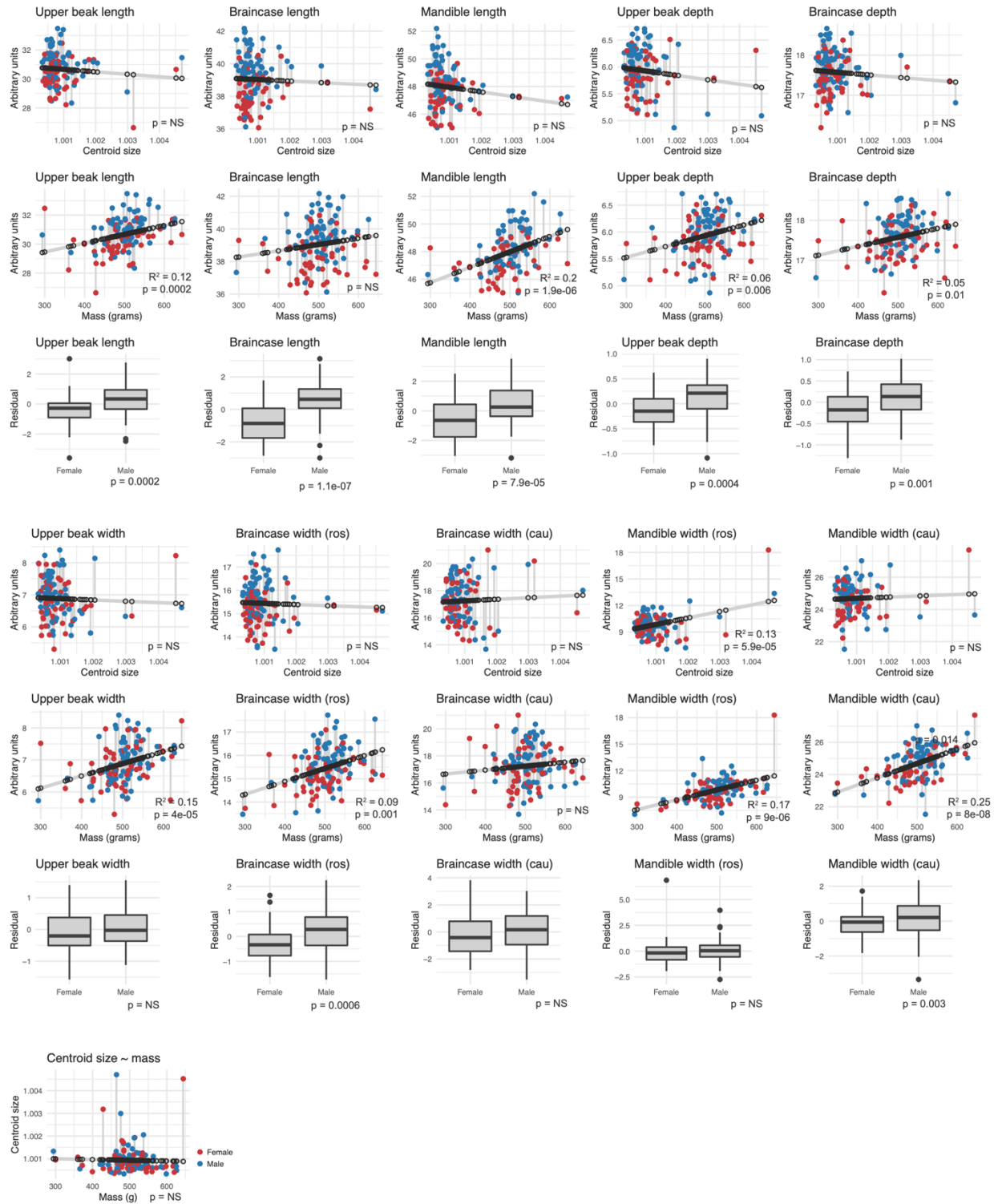
Supplemental Figure 1. Surface models of the Pom x Scan founders. Lateral (left) and dorsal (right) views of the craniofacial skeleton of the male Pom and female Scan founders used to generate the F₂ intercross.



Supplemental Figure 2. Pigeon craniofacial reference landmarks. Landmark positions are indicated by red circles; numbers correspond to descriptions in Supplemental Table 1.

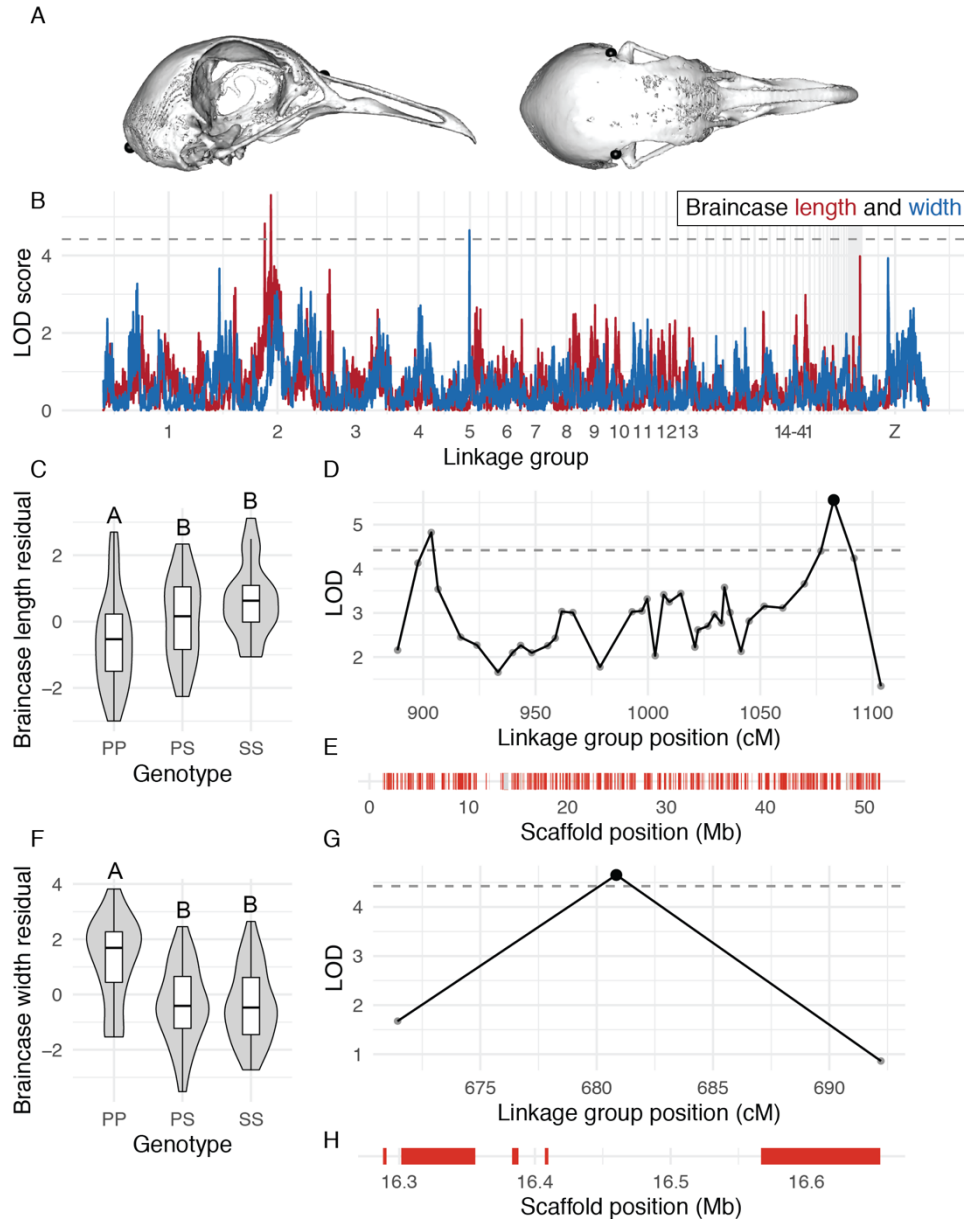


Supplemental Figure 3. Distribution of 11 linear measurements in the Pom x Scan F₂ population. With the exception of rostral mandible width, all linear measurements are normally distributed in the population (Shapiro-Wilk's test, $p > 0.05$). For rostral mandible width, a single F₂ individual is an outlier (MDS079, see Supplemental Figure 10) and causes a deviation from normality (Shapiro-Wilk's $p = 5.2e-09$). Braincase width is linear distance between landmarks 12 and 13 (rostral) or landmarks 27 and 28 (caudal). Mandible width is linear distance between landmarks 60 and 61 (rostral) or landmarks 55 and 56 (caudal).

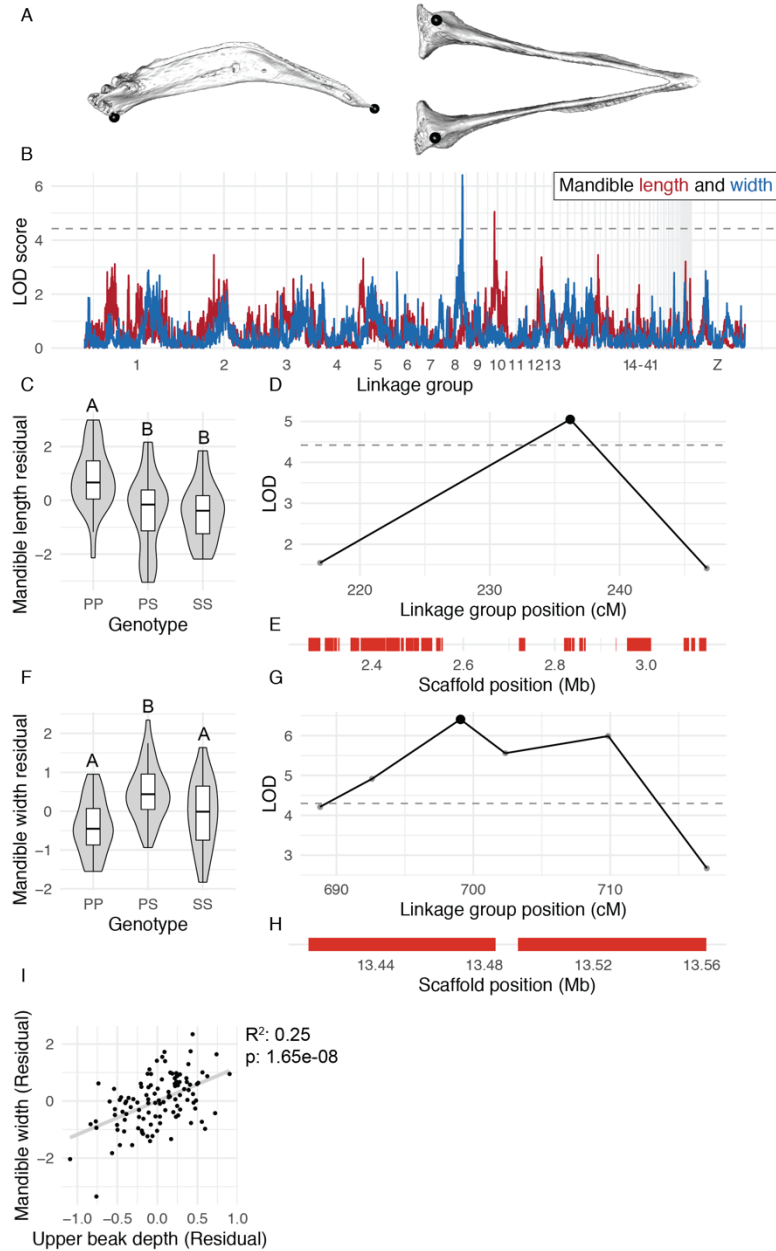


Supplemental Figure 4. Linear regression of 11 craniofacial measurements on cranium centroid size and body mass. For each phenotype, linear measurement ~ cranium centroid size regression is displayed on the top and linear measurement ~ body mass regression is displayed in the middle with associated R^2 and p -value indicated in the bottom right corner of each plot. Each dot represents raw measurement of an F_2 individual, color-coded by sex (male = blue,

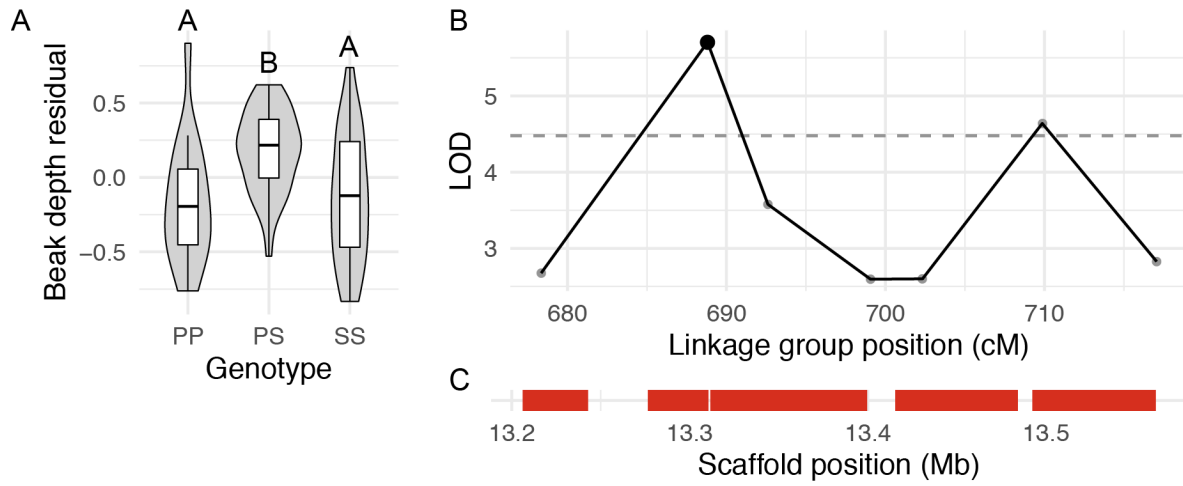
female = red). Each raw measurement is connected to an open circle that indicates its predicted value; grey connecting lines correspond to residual value used for QTL mapping. For each phenotype, the boxplot on the bottom displays residual values of linear measurement ~ body mass regression by sex; outliers are indicated by black dots. Associated p-values are indicated in bottom right corner of each plot (two-sided Wilcoxon test). Cranium centroid size ~ body mass regression is displayed in the bottom left corner of the figure. Braincase width is linear distance between landmarks 12 and 13 (rostral) or landmarks 27 and 28 (caudal). Mandible width is linear distance between landmarks 47 and 48 (rostral) or landmarks 37 and 38 (caudal).



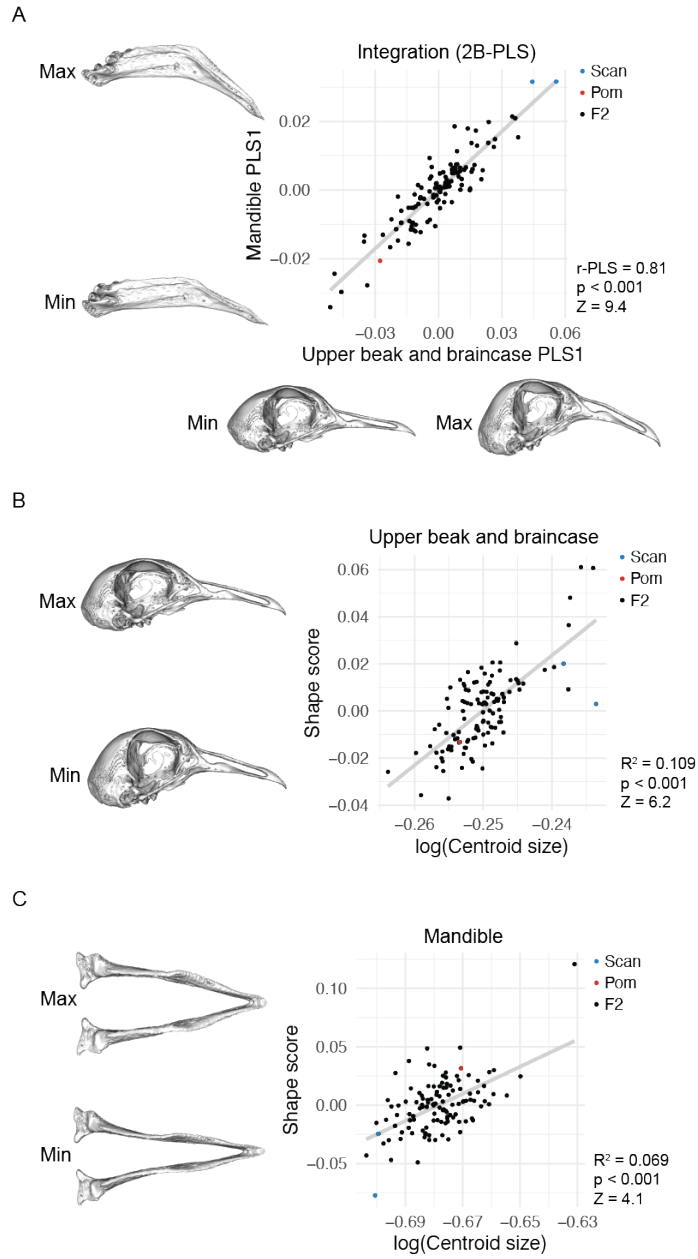
Supplemental Figure 5. QTL associated with braincase length and width. (A) Landmark pairs used to measure braincase length (left) and width (right). (B) QTL scans for braincase length (red) and width (blue). Dashed horizontal lines denotes 5% genome-wide significance threshold. (C) Effect plot for braincase length QTL on LG2. Letters denote significance groups, p-values determined via Tukey test: PP vs. PS = 8.7×10^{-3} , PP vs. SS = 2.2×10^{-4} . (D) LOD support interval for braincase length on LG2. Dots indicate linkage map markers; the black dot highlights the peak marker that was used to estimate QTL effects. (E) Genes located within braincase length QTL LOD support interval, color coded based on if gene is expressed in HH29 facial primordia (red) or not expressed (gray). (F) Effect plot for braincase width QTL on LG5. Letters denote significance groups, p-values: PP vs. PS = 5.8×10^{-5} , PP vs. SS = 7.5×10^{-5} . (G) LOD support interval for braincase width QTL on LG5. (H) Genes located within braincase width QTL on LG5. P = allele from Pom founder, S = allele from Scan founder.



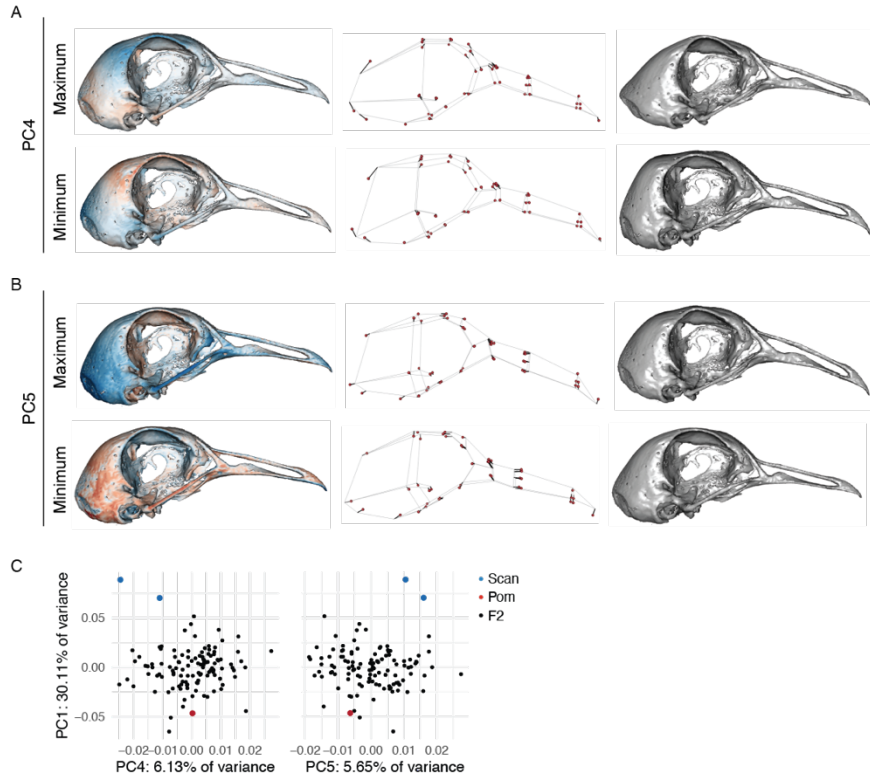
Supplemental Figure 6. QTL associated with mandible length and width. (A) Landmark pairs used to measure mandible length (left) and width (right). (B) QTL scans for mandible length (red) and width (blue). Dashed horizontal lines denotes 5% genome-wide significance threshold. (C) Effect plot for mandible length QTL on LG8. Letters denote significance groups, p-values determined via Tukey test: PP vs. PS = 5.9e-04, PP vs. SS = 1.1e-03. (D) LOD support interval for mandible length QTL. (E) Genes located within mandible length QTL LOD support interval. (F) Effect plot for mandible width QTL on LG10. Letters denote significance groups, p-values determined via Tukey test: PP vs. PS = 2.7e-05, PS vs. SS = 1.7e-02. (G) LOD support interval for mandible width QTL. (H) Genes located within mandible width QTL. (I) Scatterplot of upper beak depth and mandible width residuals for all F₂ individuals. P = allele from Pom founder, S = allele from Scan founder.



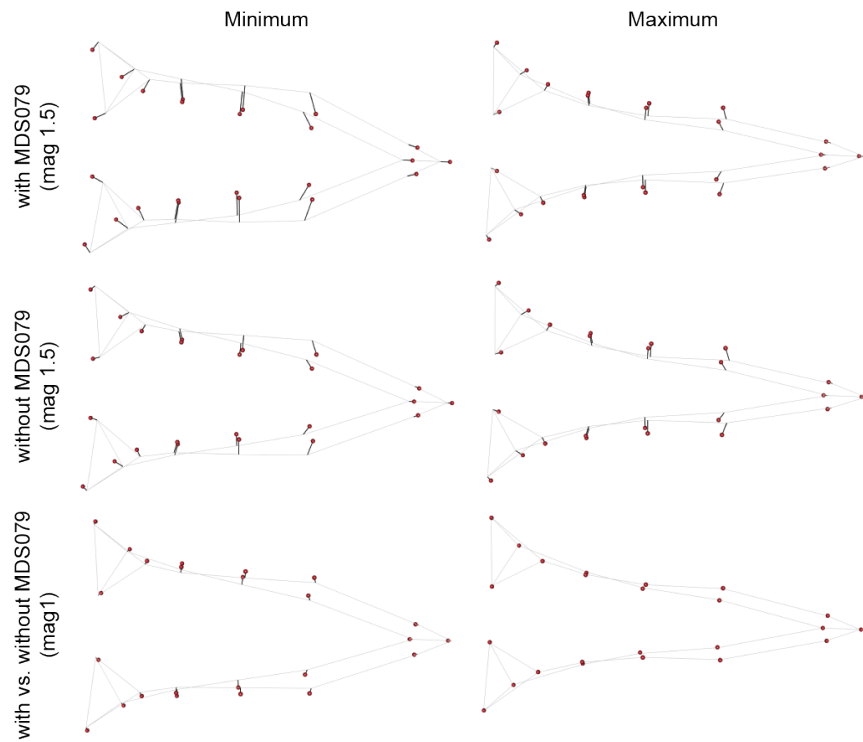
Supplemental Figure 7. QTL on LG8 associated with upper beak width. (A) Effect plot for upper beak depth QTL. Letters denote significance groups, p-values determined via Tukey test: PP vs. PS = 3.9×10^{-4} , PS vs. SS = 1.7×10^{-2} . (B) LOD support interval. (C) Genes in interval. P = allele from Pom founder, S = allele from Scan founder.



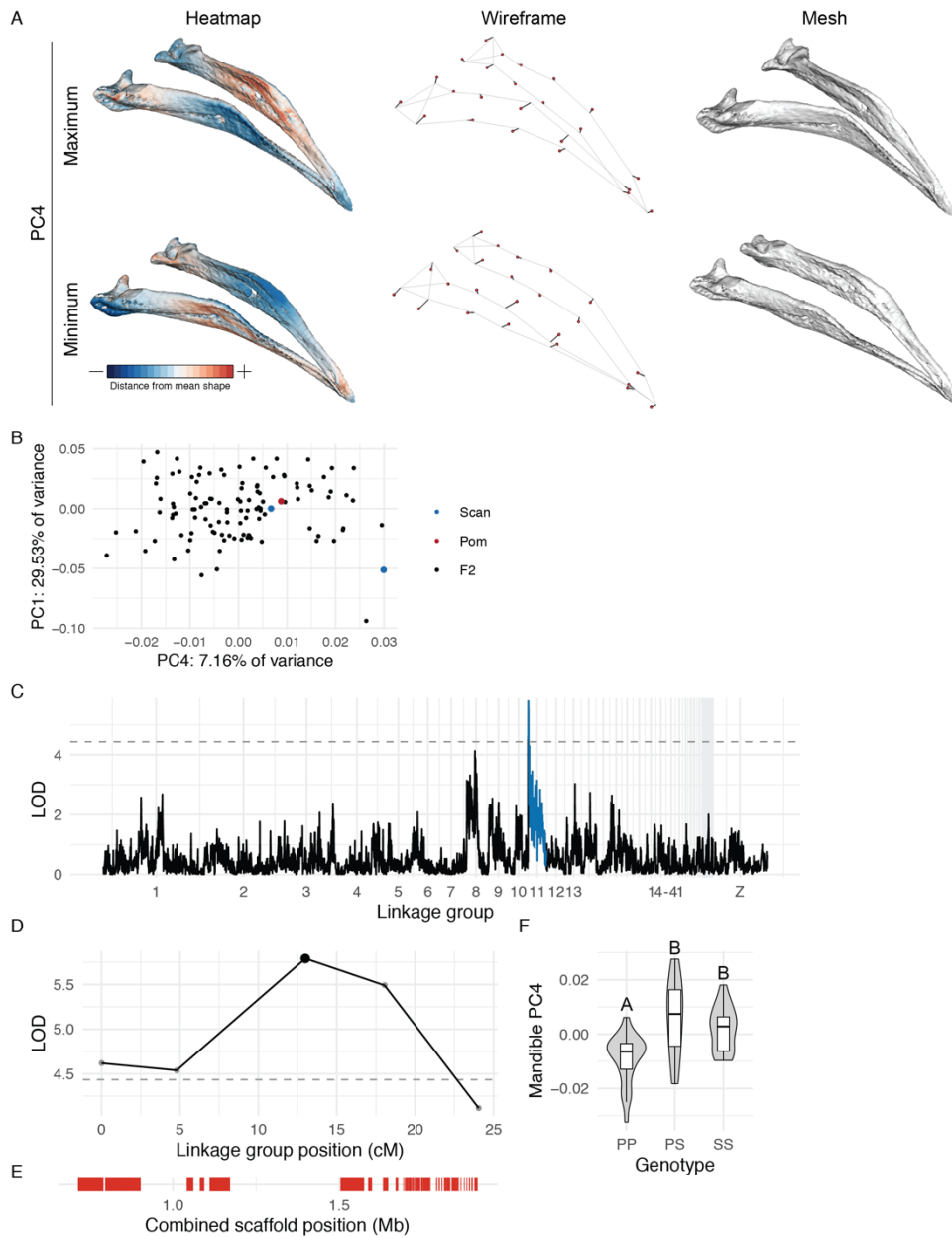
Supplemental Figure 8. UBB and MAN integration and allometry. (A) UBB PLS1 vs. MAN PLS1 shape. (B) UBB shape ~ centroid size linear regression. (C) MAN shape ~ centroid size linear regression. For all panels, minimum and maximum shapes are depicted by warped meshes along corresponding axis. Shape changes were magnified 2x to aid visualization.



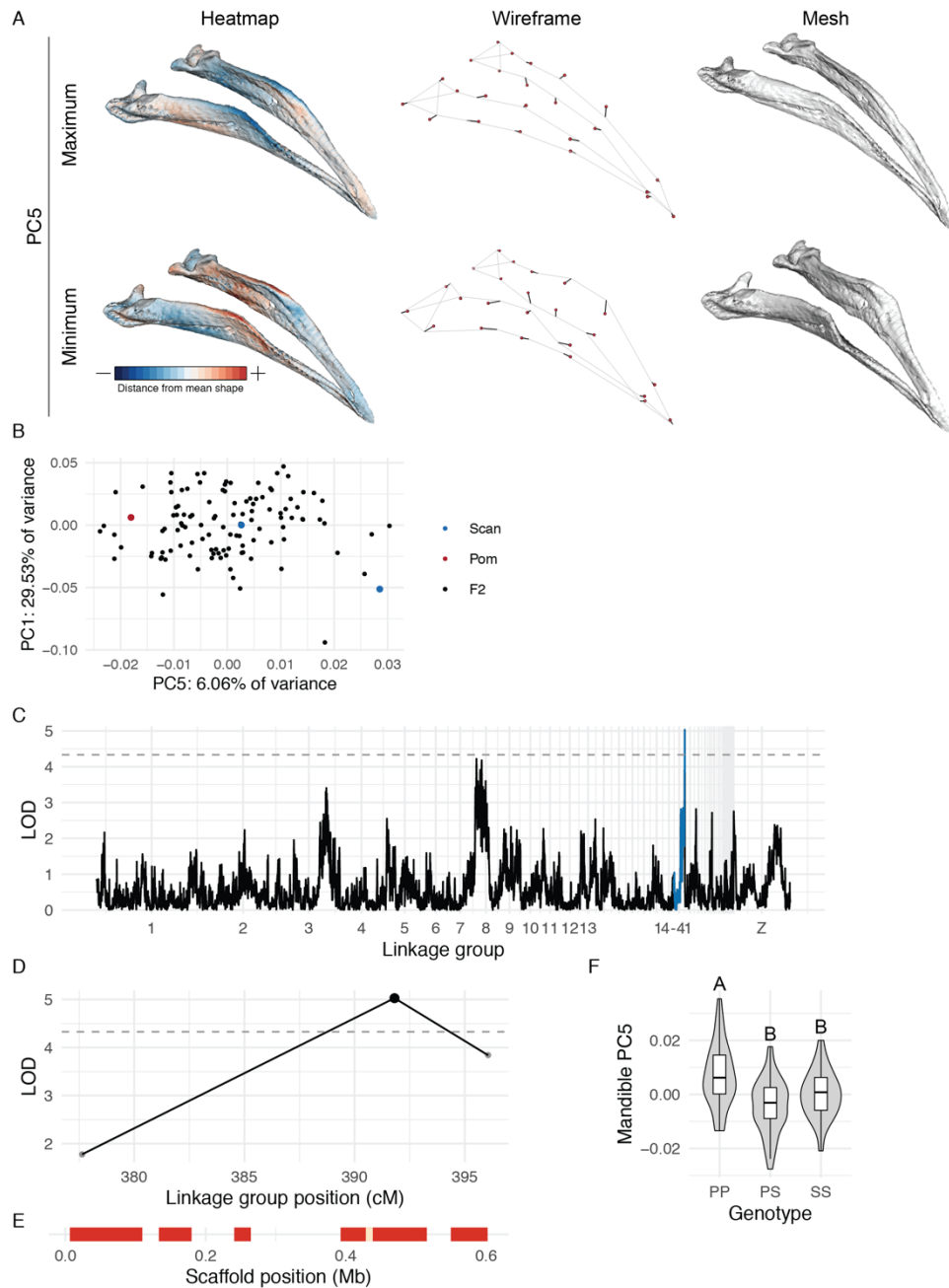
Supplemental Figure 9. UBB PC4 and PC5 shape variation. (A-B) Minimum and maximum UBB PC4 (A) and PC5 (B) shapes, visualized as heatmaps (left), wireframes (center), and warped meshes (right). For wireframes and meshes, UBB PC4 and PC5 shape is magnified 3x to aid visualization. (C) PCA plots of UBB PC1 vs. PC4 (left) and PC5 (right).



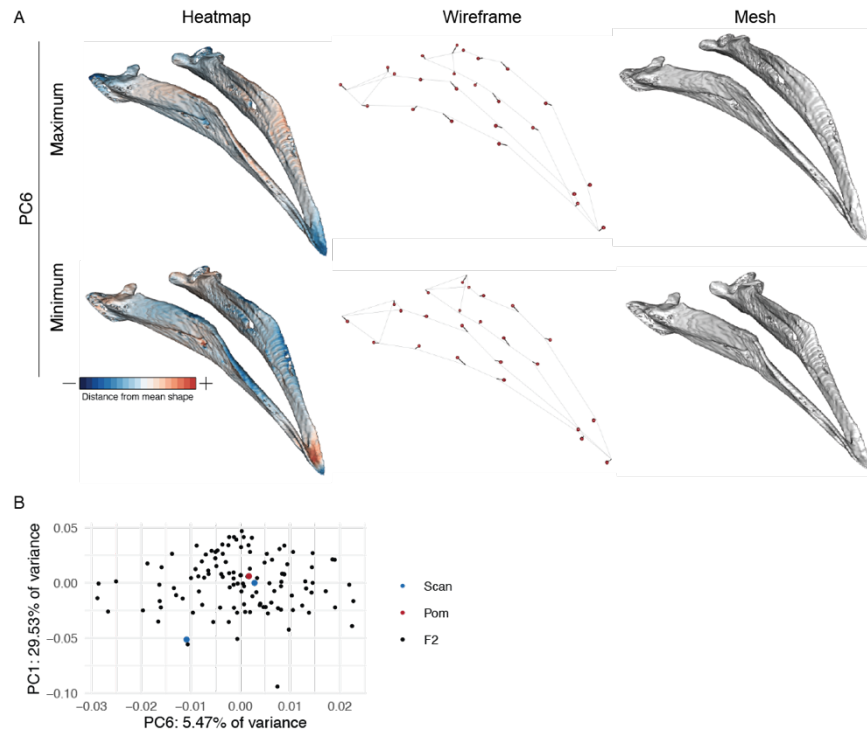
Supplemental Figure 10. MAN PC1 shape with and without MDS079. Dorsal views of MAN wireframes showing minimum (left) and maximum (right) PC1 shapes if MDS079 is included (top panel) or excluded (center panel) from the geometric morphometric analysis. MDS079 had an exceptionally wide mandible and was an outlier from the rest of the F2 population along the MAN PC1 axis (see PCA plot in Figure 4B). Although inclusion of MDS079 changed the magnitude of the PC1 axis, it had virtually no effect on the shape described by MAN PC1, thus it was kept in all downstream analyses.



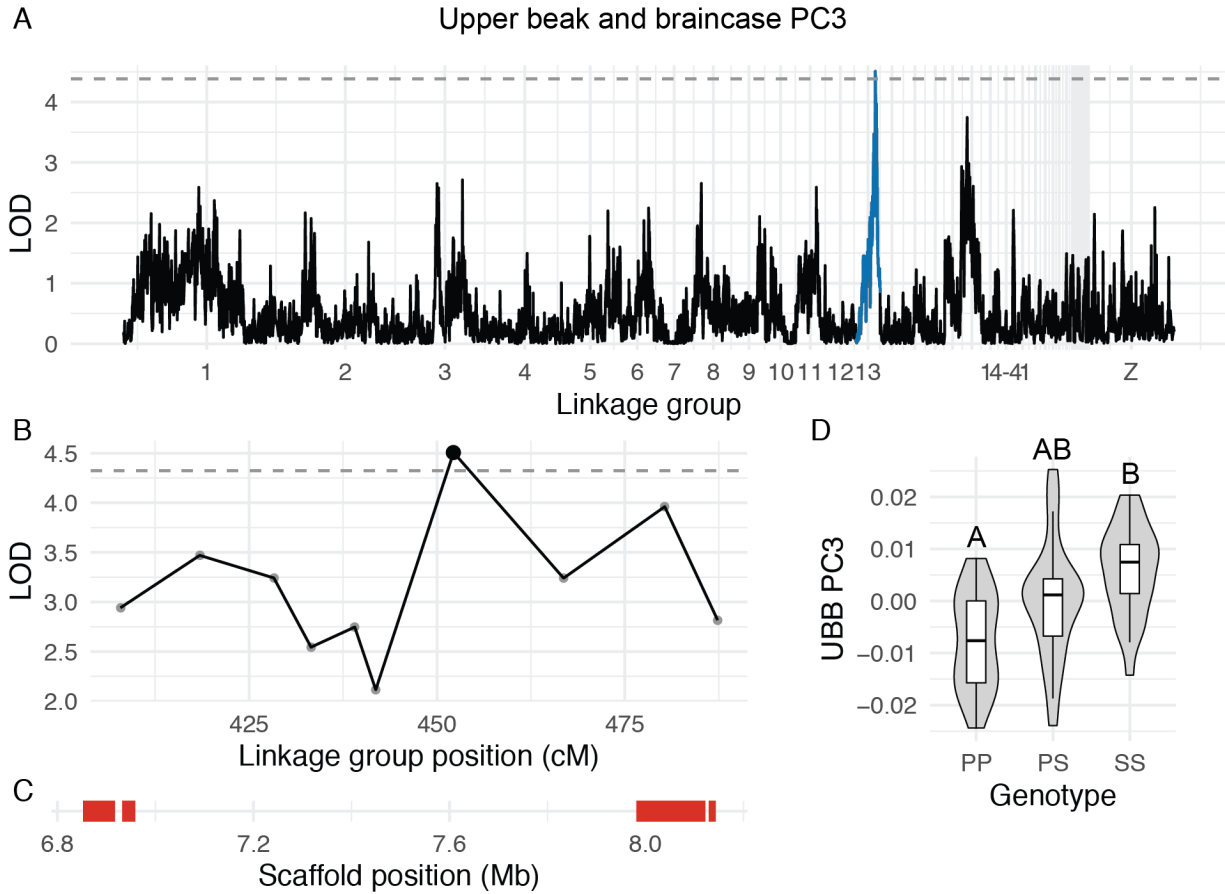
Supplemental Figure 11. MAN PC4 shape variation and associated QTL. (A) Minimum and maximum MAN PC4 shapes, visualized as heatmaps (left), wireframes (center), and warped meshes (right). For wireframes and meshes, shape is magnified 3x to aid visualization. (B) PCA plots of MAN PC1 vs. PC4. (C) Genome-wide QTL scan for MAN PC4. (D) MAN PC4 LOD support interval for QTL on LG11. (E) Genes in LG11 QTL interval. (F) LG11 QTL effect plot. Letters denote significance groups, p-values determined via Tukey test: PP vs. PS = 2.2×10^{-6} , PP vs. SS = 2.2×10^{-3} . P = allele from Pom founder, S = allele from Scan founder.



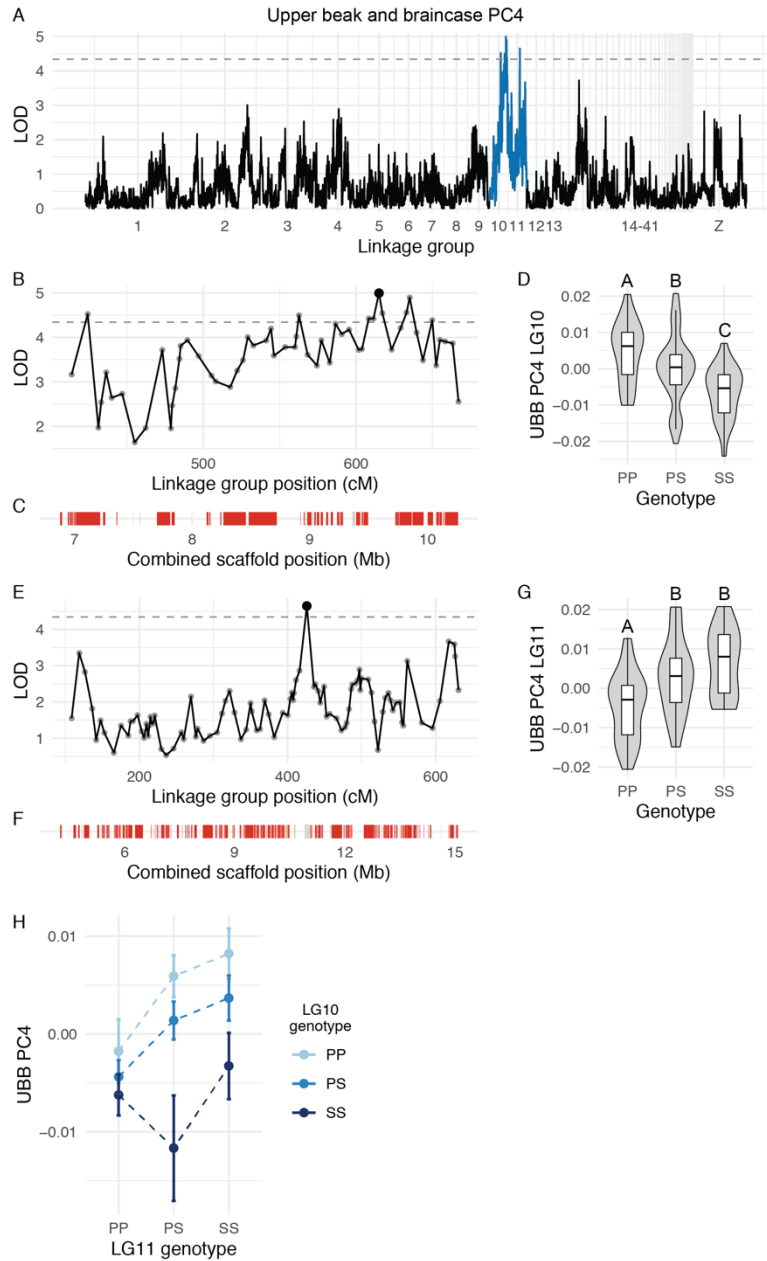
Supplemental Figure 12. MAN PC5 shape variation and associated QTL. (A) Minimum and maximum MAN PC5 shapes, visualized as heatmaps (left), wireframes (center), and warped meshes (right). For wireframes and meshes, shape is magnified 3x to aid visualization. (B) PCA plots of MAN PC1 vs. PC5. (C) Genome-wide QTL scan for MAN PC5. (D) MAN PC5 LOD support interval for QTL on LG20. (E) Genes in LG20 QTL interval. (F) LG20 QTL effect plot. Letters denote significance groups, p-values determined via Tukey test: PP vs. PS = $1.3e-05$, PP vs. SS = $1.9e-02$. P = allele from Pom founder, S = allele from Scan founder.



Supplemental Figure 13. MAN PC6 shape variation. (A) Minimum and maximum MAN PC6 shapes, visualized as heatmaps (left), wireframes (center), and warped meshes (right). For wireframes and meshes, shape is magnified 3x to aid visualization. (B) PCA plots of MAN PC1 vs. PC6.



Supplemental Figure 14. QTL associated with UBB PC3. (A) Genome-wide QTL scan for UBB PC3. (B) UBB PC3 LOD support interval for QTL on LG13. (C) Genes in LG13 QTL interval. (D) LG13 QTL effect plot. Letters denote significance groups, p-values determined via Tukey test: PP vs. SS = $5.3e-04$. P = allele from Pom founder, S = allele from Scan founder.



Supplemental Figure 15. QTL association with UBB PC4. (A) Genome-wide QTL scan for UBB PC4. (B) UBB PC3 LOD support interval for QTL on LG10. (C) Genes in LG10 QTL interval. (D) LG10 QTL effect plot. Letters denote significance groups, p-values determined via Tukey test: PP vs. PS = $1.6e-02$, PP vs. SS = $2.4e-05$, PS vs. SS = $2.5e-02$. (E) LOD support interval for LG11 QTL. (F) Genes in LG11 QTL support interval. (G) LG11 QTL effect plot. Letters denote significance groups, p-values determined via Tukey test: PP vs. PS = $2.4e-03$, PP vs. SS = $7.5e-05$. (H) Interaction between LG10 and LG11 QTL. P = allele from Pom founder, S = allele from Scan founder.