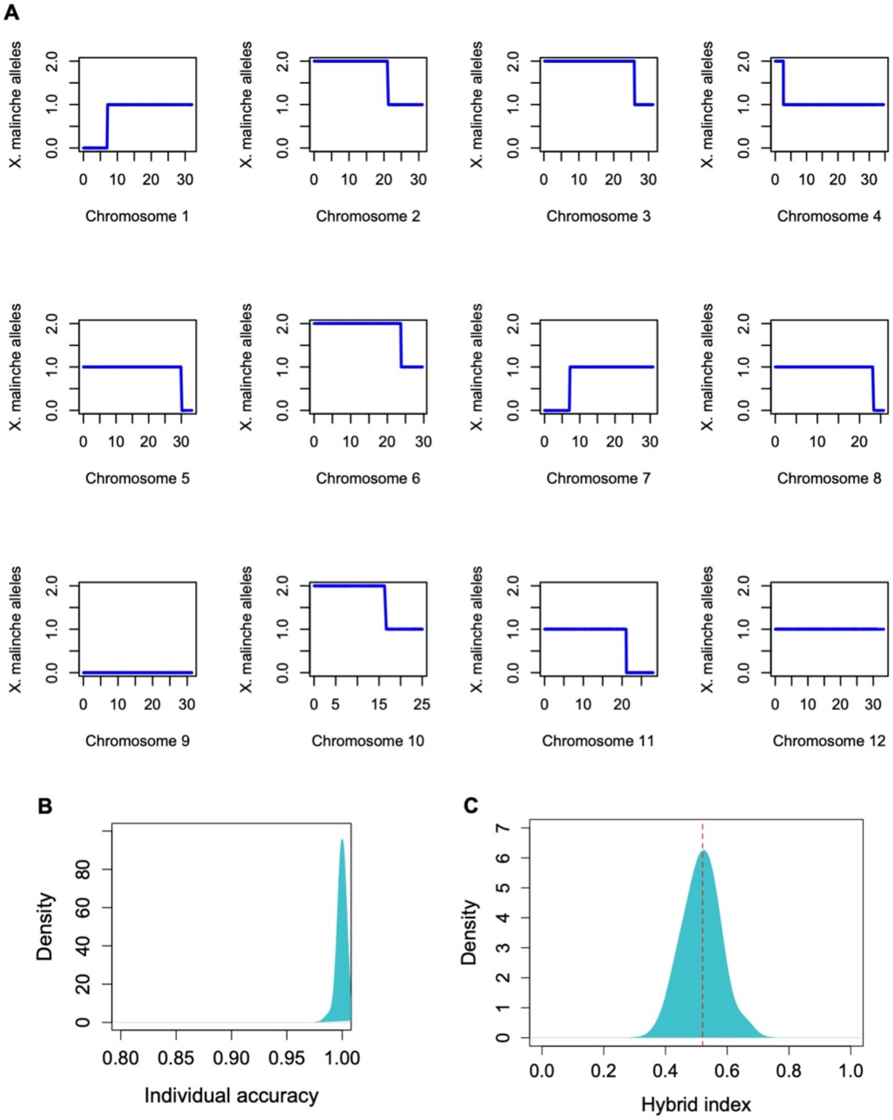


Figure S1. **Results of local ancestry inference of individuals in our study as well as simulations of local ancestry inference, related to Figure 1. A)** Local ancestry along chromosomes 1-12 for an F_2 hybrid individual. Plotted here are the number of *X. malinche* alleles at each ancestry informative site supported by a posterior probability of 0.9 or greater for a given ancestry state. Scale on x-axis corresponds to the chromosome length in megabases. **B)** Expected individual level accuracy from simulations of early generation hybrids. Simulations were conducted using the *mixnmatch* and *ancestryinfer* programs with parameters matching those observed in our study system. **C**) Genome-wide distribution of *X. malinche* ancestry among F_2 hybrids in our mapping population. Although on average individuals derive 50% of their genome from each parental species (dashed red line), there is substantial variance in ancestry generated by the recombination process.

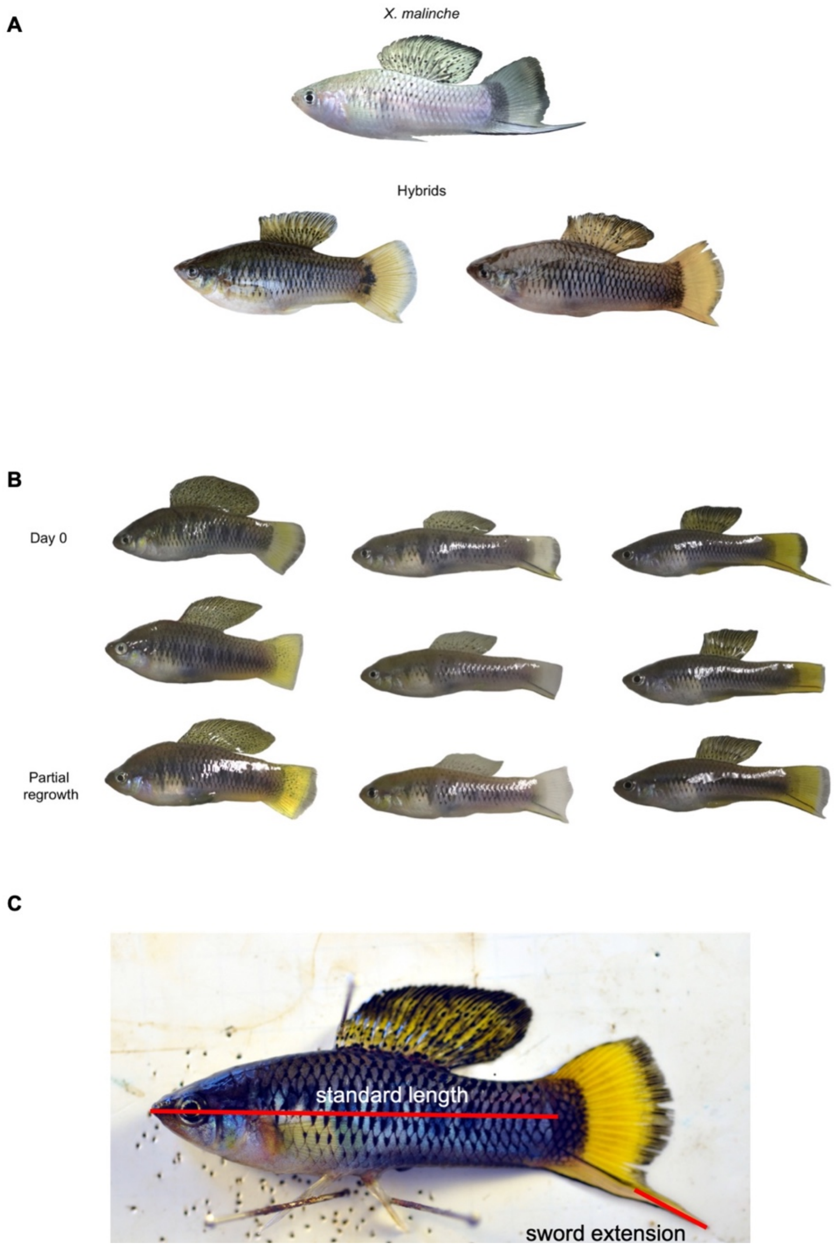


Figure S2. Phenotypic variation seen in hybrids and sword regrowth phenotypes, related to Figure 1. A) Sword length and sword black margin can become decoupled in hybrids even though the traits are always coupled in *X. malinche.* Top - *X. malinche.* Bottom left - male hybrid with a short sword lacking upper and lower pigmented margin. Bottom right - male hybrid with a short sword lacking an upper sword pigmented margin but displaying a lower sword pigmented margin. **B)** Stages of sword regeneration in regrowth experiments. Example of *X. birchmanni* (left), F_1 (middle), and *X. malinche* (right) fish included in the sword regeneration RNAseq experiment. Shown here are phenotypes from fish pre-removal of the edge of the caudal fin (and sword in *X. malinche* and F_1s), post-removal, and after tissue regrowth. **C)** Schematic showing measurement of sword length using ImageJ software. Sword extension was normalized by standard length for QTL analysis of sword length.

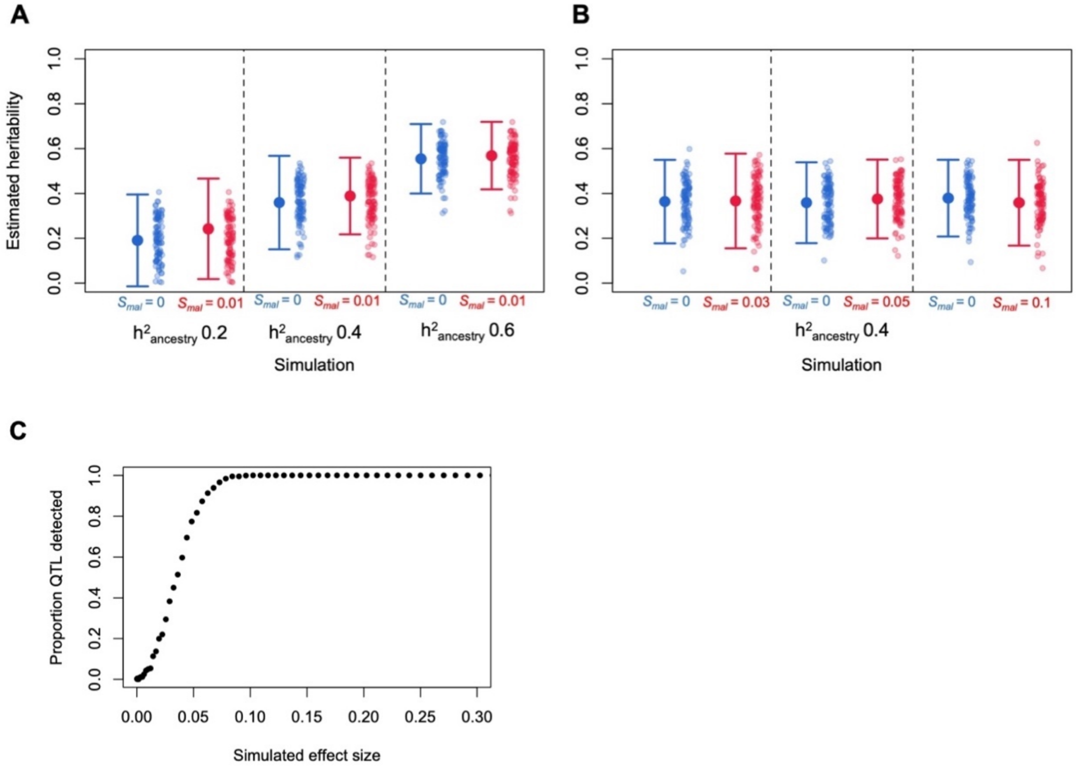


Figure S3. Examination of assumptions made in heritability calculations and predicted power in QTL mapping, related to Figures 1 and 2. Panels A & B show the results of admix'em simulations evaluating accuracy of heritability estimates varying the level of genetic variation for the trait of interest in parental populations. **A**) Broad sense heritability of a phenotype as a function of ancestry was varied from 0.2 to 0.6 across simulations. We also varied whether all phenotypic variation was attributable to ancestry (blue) or whether there was segregating variation for the phenotype within the simulated *X. malinche* population (red). In those simulations we drew allele frequencies for the causal loci in the *X. malinche* population from a random exponential distribution and arbitrarily set the effect size to 1% of the simulated QTL effect size. Large points and whiskers show mean and two standard deviations across 100 simulations, raw data per simulation is shown by individual points. **B**) In another series of simulations, we varied the effect size of the alleles segregating in the simulated *X. malinche* population from 3-10% of the QTL effect size (red). Simulations with no segregating variation in *X. malinche* are shown in blue for comparison. Large points and whiskers show mean and two standard deviations across 100 simulations, raw data per simulation is shown by individual points. See Materials and Methods for a complete description of these simulations. **C**) Predicted power curve for QTL detection in our study as a function of simulated QTL effect size. We varied the proportion of phenotypic variation explained by a single QTL and tested our power to detect it in simulations. Each point represents the proportion of 1,000 simulations in which the QTL was detected at our genome-wide significance threshold.

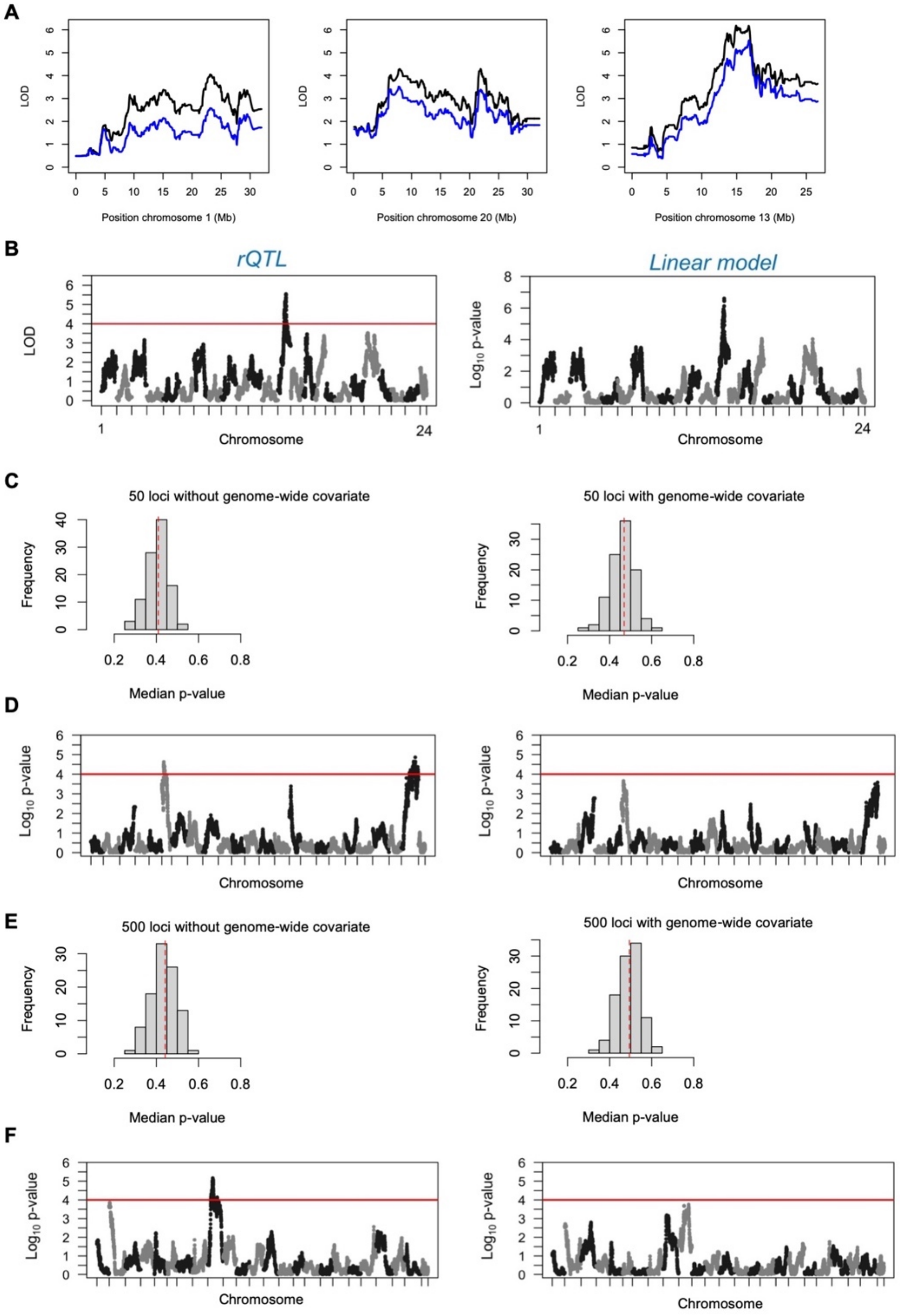
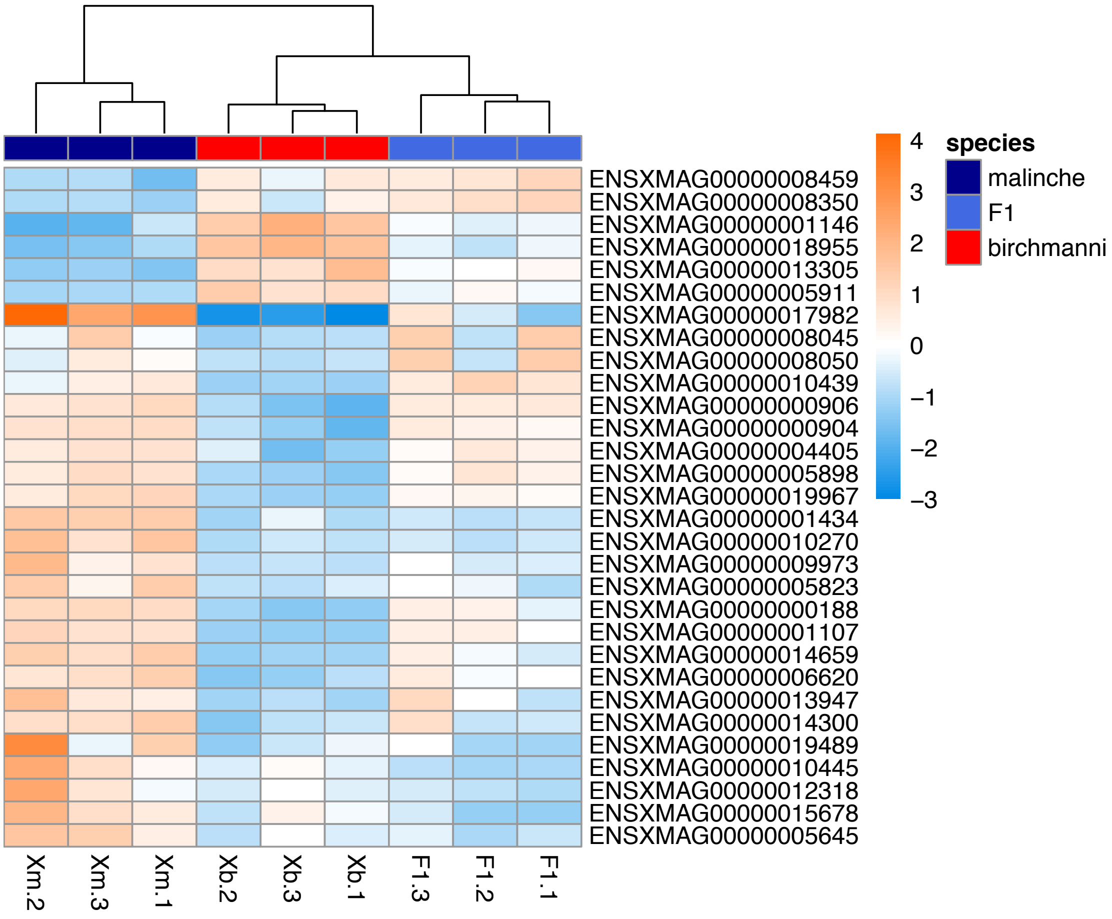


Figure S4. Analysis of impacts of not accounting for genome-wide ancestry in QTL mapping analyses, related to Figures 2 and 3. A) Results of QTL analysis for sword length presented in the main text with and without genome-wide ancestry included as a covariate. Plots show results for three chromosomes with (blue) and without (black) genome-wide ancestry included as a covariate. Without ancestry included as a covariate (black lines), we recover three QTLs that pass the genome-wide significance threshold (LOD=4). However, the peaks on chromosome 1 and chromosome 20 have a relatively flat signal and drop below the genome-wide significance threshold when we account for genome-wide ancestry in R/qtl analysis (blue lines). **B)** Comparison of linear model and R/qtl results. Linear model-based mapping results (right) for sword length mirror results from R/qtl (left). Due to extremely long runtimes it was impractical to use R/qtl in simulations, but these results suggest that a linear model-based approach gives qualitatively similar results. **C)** Distribution of median p-values in simulations of polygenic traits determined by ancestry at 50 underlying loci. Left - Distribution of median p-values at ancestry informative sites in each of 100 simulations when genome-wide ancestry is not included as a covariate in the analysis. Right - Distribution of median p-values at ancestry informative sites in each of 100 simulations when genome-wide ancestry is included as a covariate in the analysis. Red dashed line indicates the median of each distribution. P-values are skewed towards lower values when genome-wide ancestry is not accounted for. **D**) Example results for simulations of 50 loci contributing to variation in a polygenic trait. Left - Manhattan plot showing results without genome-wide ancestry included as a covariate. Right - Manhattan plot showing results for the same simulation with genome-wide ancestry included as a covariate. Red line shows genome-wide significance threshold used in our study. **E**) Distribution of median p-values in simulations of polygenic traits as in C but with 500 underlying loci determining a trait. Given that we should have near zero power to detect loci of these effect sizes in our simulations, the skew observed without correction likely reflects the impacts of ancestry structure on p-value distributions. **F)** Example results as in D but for simulations of 500 loci contributing to variation in a polygenic trait. Left - Manhattan plot showing results without genome-wide ancestry included as a covariate and Right - Manhattan plot for the same simulation with genome-wide ancestry included as a covariate.



FigureS5. Heatmap of 30 differentially expressed genes between *X. birchmanni* and *X. malinche* regenerating tissue with the strongest expression differences between species, related to Figure 4. 30 out of the 3,333 significantly differentially expressed genes are shown. Many of these genes have intermediate expression in F_1 hybrids. Dark blue, light blue, and red rectangles under the dendrogram indicate species identity of each biological replicate, light blue to orange shading in the matrix indicates relative expression level. Color legend to the right indicates expression level of a sample on a $\log_2$ fold change scale, relative to mean expression of that gene across samples.

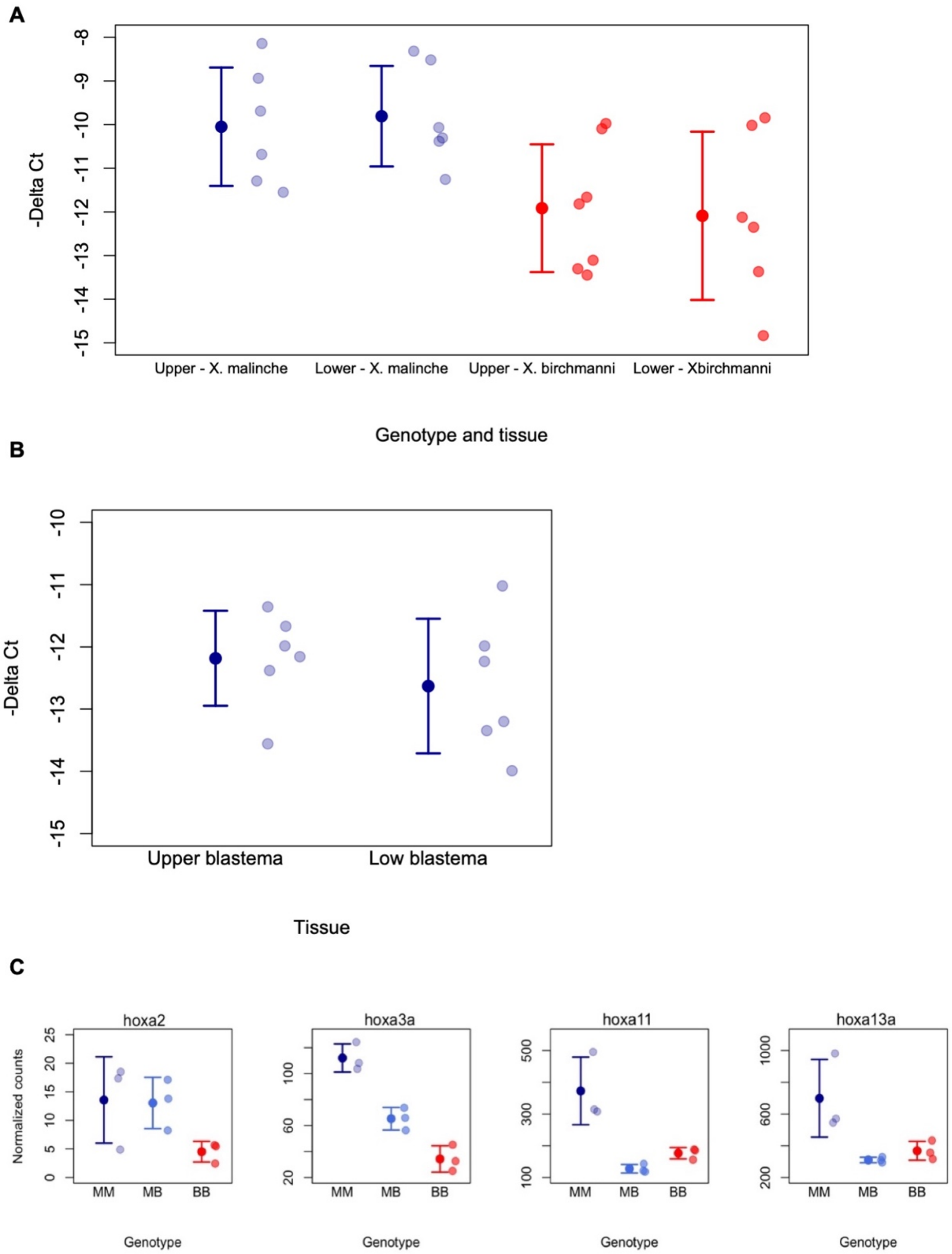


Figure S6. rt-qPCR investigation of candidate genes and RNAseq expression results of additional candidate genes, related to Figures 4 and 5. A) *grn2* expression measured by rt-qPCR in upper and lower regenerating caudal tissue in *X. malinche* and *X. birchmanni.* Semi-transparent points show mean -ΔCt across technical replicates for each individual and solid points and whiskers show the group mean ± one standard deviation. There were no significant differences in *grn2* expression between upper and lower portions of the regenerating blastemas within species. **B)** *sp8* expression measured by rt-qPCR in upper and lower regenerating caudal tissue in *X. malinche.* Semi-transparent points show mean -ΔCt across technical replicates for each individual and solid points and whiskers show the group mean ± one standard deviation. There were no significant differences in *sp8* expression between upper and lower portions of the regenerating blastemas. **C)** Expression levels of differentially expressed *hoxa* genes from RNAseq data in regenerating caudal fin tissue in *X. malinche* (dark blue), F_1 hybrids (light blue) and *X. birchmanni* (red). This cluster of genes is found nearby the QTL region we identify on chromosome 13. Solid points and whiskers show mean and one standard deviation, semi-transparent points show the raw data per individual.

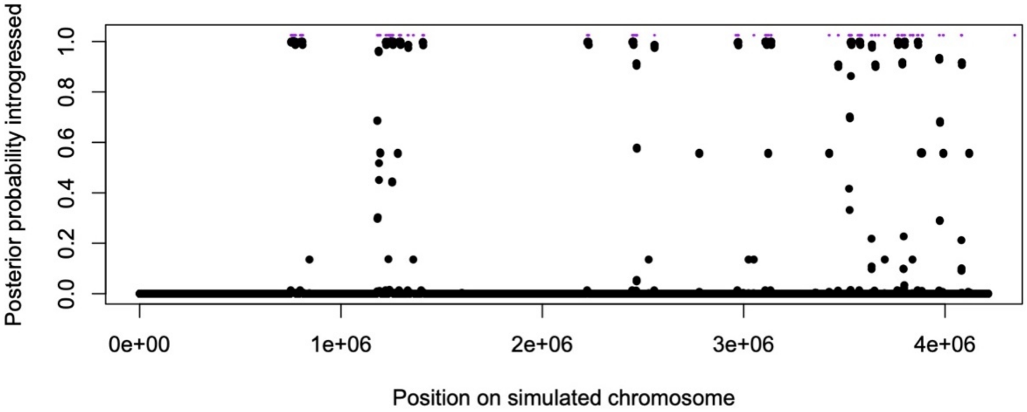


Figure S7. Example of results from simulations of ancient admixture and application of PhyloNet-HMM to infer local ancestry, related to Figure 5. Black dots show the posterior probability inferred by PhyloNet-HMM that the site is hybridization derived. Purple lines above the black dots show the true locations of admixture derived tracts, determined from decoding tree sequences in SLiM.

Closest gene annotation	**start**	**end**	**Width of peak**	**Fold change**	**p-value**	**FDR**	**Type; Distance from gene**
rapgef5a	15645736	15645922	187	3.57	2.87E-05	5.23E-05	Intergenic; -14140
ITGB8	15746149	15746358	210	3.43	1.78E-05	5.23E-05	Intergenic; -10249
BTD	15834337	15834509	173	-4.61	9.41E-20	2.52E-16	Intergenic; -3344
SLC25A13	16441207	16441383	177	2.78	8.04E-05	0.0023	Intergenic; -4898
SLC25A13	16442176	16442361	186	2.85	5.38E-05	0.0017	Intergenic; -5853
tbc1d5	16795427	16795598	172	2.6	3.24E-05	0.00125	Intergenic; -17455
tbc1d5	16801227	16801427	201	2.73	8.27E-05	0.00224	Intergenic; -11655
rftn1a	16909129	16909372	244	-2.8	3.12E-05	0.00125	Intergenic; 2052

Table S3. Species specific differentially accessible peaks identified with ATAC-seq in regenerating *X. birchmanni* and *X. malinche* ventral caudal blastemas, relates to Table 1. Differential peak analysis was performed with the program diffBind and peaks with corrected p-values <0.05 within the QTL region are included in this table. Only five regions under the joint QTL were found to have differential accessibility, and three were found outside of the QTL region on the 3’ edge.

Supplemental Reference

S1. Powell, D.L., García-Olazábal, M., Keegan, M., Reilly, P., Du, K., Díaz-Loyo, A.P., Banerjee, S., Blakkan, D., Reich, D., Andolfatto, P., et al. (2020). Natural hybridization reveals incompatible alleles that cause melanoma in swordtail fish. Science. 368, 731–6.