Figure S6   Fine-mapping recombinant breakpoints with a Hidden Markov Model
(A) Reads were binned into small 10 kilobase (kb) bins. For each bin, the probability of each genotype was determined, and boundaries of the recombination breakpoint were set at the 99% cutoff. (B) Histogram of recombination breakpoint resolution in the FTC x LITC cross. Median resolution was 89 kb. 539 additional recombinants (not shown) had a resolution of greater than 500 kb. (C) Diagram of high-resolution (<350 kb) recombination breakpoints near three marine/freshwater inversions. Each line represents the 99% confidence interval for a recombination breakpoint in a single fish and the grey rectangles represent the previously published marine/freshwater inversion boundaries (Jones et al. 2012). Data is presented for the entire scaffold containing each inversion (scaffold 22 on chromosome 1, scaffold 11 on chromosome 11, and scaffold 16 on chromosome 21). Physical position coordinates from the original assembly are used. No recombination events are observed within these three inversion intervals in each cross. (D) Trained Hidden Markov Model parameters used for fine-mapping recombination breakpoints. SNP genotypes were binned into 10 kb bins. For each fish, each bin was assigned a raw genotype of M (only marine reads), F (only freshwater reads), B (both marine and freshwater reads), or X (no data). These raw genotypes were used as the observed data in a Hidden Markov Model with hidden states M (marine), F (freshwater), or H (heterozygous). The model was trained with the Baum-Welch algorithm on 1 million data points.