Supplementary Material

Supplementary Table 1. Study Arms of the pre-Collaborative Cross experiment. All animals were bred to weaning at the Oak Ridge National Laboratory (ORNL) and shipped to either the National Human Genome Research Institute (NHGRI/NIH) or the University of North Carolina at Chapel Hill (UNC) for phenotyping. Blood, urine, and tissue samples were sent to The Jackson Laboratory (JAX) and National Jewish Medical and Research Center (NJC) for additional molecular phenotyping. All animals were genotyped at UNC.

Supplementary Table 2. Candidate genes for *Bwq14*. *Asph* was the highest priority candidate from three independent analyses. *Cyp7a1, Chd7, Gdf6* were the next highest priority candidates, indicated by both the GO and sequence analyses.

Supplementary Figure 1. Pre-CC subjects by generation. Pre-CC subjects were selected from available lines that had reached G2:F5 or beyond and allocated across the four study arms.

Supplementary Figure 2. Distribution of genome segments. A genome segment is the region between recombination breakpoints in an individual pre-CC genome. a) The average number of recombinations per line was 142.3, which approximated simulated expectations. b) Segment sizes are distributed exponentially with median of 10.46 Mb.

Supplementary Figure 3. Marker-based allele effects for white head-spotting. a) The QTL on Chr 2 is driven completely by a heterozygote effect. There were only two heterozygotes in the population at that locus, and they both happened to have a white spot. This type of false positive occurs as a result of skewed genotype classes, and the heterozygous class is always small in the mostly inbred pre-CC population. b) The Chr 10 QTL allele effects are consistent with recessive white spotting.

Supplementary Figure 4. Histograms of eQTL allele effects. CAST/EiJ and PWK/PhJ allele effects are more often in the extreme ends of the distribution for a given eQTL.

Both also show an excess of negative allele effects relative to positive, which may represent hybridization error due to undiscovered SNPs in these strains.

Supplementary Figure 5. Histogram of liver eQTL density in the pre-CC population. The frequency of corrected eQTL density (fraction between the number of eQTLs and protein coding genes) in non-overlapping 5 Mb intervals.

Notes on the expected number of white head spotted lines

Of 111 nonalbino mice, we expect 16% to be heterozygous based on the observed genome-wide average reported in our results, and we expect one-eighth of the homozygotes to have the recessive white spotting allele (e.g. *P(white spot | nonalbino)* = 0.105). Thus, we would expect twelve of 111 mice would be white spotted average (binomial μ = 11.655), and observing six or fewer is unlikely ($P(x \leq 6) = 0.0465$). If not due to chance, this underrepresentation could result from missed observations, segregation distortion reflecting selection against the white spotting locus, or a heterozygosity rate higher than 16% at the causative locus. The first is the most likely given that we did not record white spotting as a phenotype during the first few cohorts dissected over the two-year long course of the experiment. For these we relied on video recordings from a behavior experiment involving these mice.

Notes on breeding errors

We identified three CC lines that showed evidence of deviation from the CC breeding design. One aspect of the design is that mitochondria (M) are always descended from the leftmost female in the "funnel" diagram, and the Y chromosome is always descended from the rightmost male. In two lines, genotypes indicated that either M or Y had not descended from the expected founder strains and those lines were eliminated from the study before analysis. One line lacked any CAST/EiJ contribution and the other appeared to be a mixture of two advanced lines.

One additional line stood out during our analysis due to its lack of any contribution from the PWK/PhJ strain. We genotyped and analyzed the ancestor of this line from the G2:F1 generation (two generations of outcrossing) and determined that PWK/PhJ contribution was also lacking from that sample as well. This indicates there was no PWK/PhJ contribution to the line and this too was a breeding error.