SUPPLEMENTARY RESULTS

Mapping of Agouti trait using Fisher’s exact: We tested the agouti trait using Fisher’s exact test (8 x 2 contingency table, with eight CC founders, two phenotypic values) per SNP. Figure S2 shows the results of this analysis. We found that Fisher’s exact test was just as effective as the logistic regression model in finding QTL positions. However, its utility was limited for more complex studies since it cannot handle covariates.

Figure S2  Plot of QTL mapping using Fisher’s exact test for the agouti trait.
SUPPLEMENTARY METHODS

Haplotype reconstruction

1. For each Geniad strain:
   a. Extract all 69,245 homozygous non-monomorphic SNPs for the 8 cc founders.
   b. Extract the exact same 69,245 SNPs of the strain tested.
   c. Take out SNPs that are called (‘N’) during genotyping. If greater than 62,320 SNPs remain, then proceed to next step.
   d. Follow rules: A <-> T, G <-> C, to correct and match strand to that of the founder genotypes at all SNPs
   e. At each of 20 chromosomes: SNPs are incrementally parsed through PEDPHASE v3 to determine phased Allele 1 homozygous genotypes.
   f. Check is performed to ensure integrity of phased Allele 1 data against original genotyping data.
   g. Phased allele 2 genotypes are simply homozygous genotypes of the left over allele at each SNP.
   h. HAPPY format '.alleles' files are created for each chromosome with SNPs placed in order of increasing genetic distance. (cM) (refer to HAPPY manual)
   i. HAPPY format '.data' files are created for each chromosome, in which two set of genotypes: phased allele 1 homozygous genotype and phased allele 2 homozygous genotype are kept and labeled 'strain.name_A' and 'strain.name_B'.
   j. HAPPY objects are created for each pair of 20 '.alleles' and '.data' file with options set: phase = 'known', haploid = TRUE
   k. The founder weights for allele 1 genotype and allele 2 genotype are calculated individually using hdesign() method in HAPPY. This will return two sets of weights in the range of 0 -1.
   l. The founder that carries highest weight in each set is determined as say, founder_A and founder_B respectively. In a homozygous block, founder_A will be same as founder_B. Thus concensus is founder_A. In a heterozygous block founder_A and founder_B are not same. Thus concensus is ‘founder_A or founder_B’ (example: “AJ or B6”).
   m. OPTIONAL: For added confidence: go back to step h, place SNPs in reverse order for each chromosome and adjust genetic distance (cM) accordingly, then repeat steps (i) to (l) with SNPs in reverse orientation. Then, sum the hdesign weights of a marker computed in forward and reverse orientation. This will return two summed sets of founder weight in the range 0-2. From this, founder_A and founder_B can be determined as before.
   n. Once founder haplotype for each non-missing SNP for the strain is determined, this is then transferred to the full SNP set containing 69,245 SNPs, where SNPs with missing genotype are assigned (imputed with) haplotype as that
of the SNP with non-missing genotype immediately above the missing SNP. In case, should missing genotype occur
at the start of a chromosome, haplotype imputed is that of the non-missing SNP immediately below.

o. Thus haplotype is reconstructed at all 69,245 markers for the strain

2. For N strains: 2 D Matrix \( M = N \times 69,245 \) Haplotype matrix is obtained by repeating step 1 for all N strains individually.

3. A 3 D matrix ‘geno’ is then created with dimensions \( (N_{\text{strains}} \times 8_{\text{founders}} \times 69,245_{\text{SNPs}}) \) and assigned ‘0’ value throughout. Then for each strain/SNP pair of 2 D matrix \( M \), a homozygous founder haplotype is assigned a value 1, and in the case of heterozygous pair of founder haplotypes, each founder in the pair is assigned a value 0.5; in the ‘geno’ matrix.

4. KINSHIP: \( N \times N \) kinship matrix can be obtained by calculating correlation coefficient between reconstructed haplotypes of pairs of strain. Alternatively IBS matrix can be derived from using raw genotypes and utilized as kinship.

QTL mapping

A. Logistic regression based QTL (qualitative) mapping

1. Using R, following code is used to test each marker:
   a. \( \text{fit1} = \text{glm}(\text{as.factor(phenotype)} \sim \text{geno}[,1], \text{family}=\text{binomial}) \)
   b. \( \text{fit2} = \text{glm}(\text{as.factor(phenotype)} \sim 1, \text{family}=\text{binomial}) \) # Null
   c. \( \text{anova(fit1,fit2, test='Chisq')} \) # anova test

B. Multi-nominal Logistic regression based QTL (qualitative) mapping

1. Using R, the following code is used to test each marker:
   a. \( \text{library(nnet)} \)
   b. \( \text{fit1} = \text{multinom( as.factor(phenotype) \sim \text{geno}[,1])} \)
   c. \( \text{fit2} = \text{multinom( as.factor(phenotype) \sim 1)} \) # Null model
   d. \( \text{anova(fit1,fit2, test='Chisq')} \) # anova test

C Linear regression based QTL (quantitative) mapping

1. Using R, the following code is used to test each marker:
   a. \( \text{fit1} = \text{lm(phenotype} \sim \text{geno}[,1]) \)
   b. \( \text{fit2} = \text{lm(phenotype} \sim 1) \) # Null model
   c. \( \text{anova(fit1,fit2,test='F')} \) # anova F-test

D Linear mixed model based QTL (quantitative) mapping
1. Using R, the following code is used to test each marker:
   
a. `library(DOQTL)`
   
b. use `scanone()` method with appropriate options (ref. manual)