Genetics of mouse *Tas1r3*-independent sucrose intake

Cailu Lin[1], Alexander A. Bachmanov [1,2], Michael G. Tordoff[1], Gary K. Beauchamp[1] and Danielle R. Reed[1*]

[1]Monell Chemical Senses Center, Philadelphia, Pennsylvania, United States of America

[2]Current Location: GlaxoSmithKline, Collegeville, Pennsylvania, United States of America

* Corresponding author

E-mail: reed@monell.org (DRR)

Abstract:

Variation in sucrose intake among inbred mouse strains is due in part to polymorphisms in the gene *Tas1r3,* which encodes a sweet taste receptor subunit and engenders the *Sac* locus on distal mChr4. Here, we eliminated variation in this locus to discover influences of additional genetic variations on sucrose intake. We measured voluntary daily sucrose intake in an F_2 intercross with the *Sac* locus fixed and in several other mapping populations: backcross, reciprocal consomic, and single and double congenic strains. The chromosome mapping results implicated *Scq2*, located on Chr9 between 105.7 and 106.9 Mb (*rs33653996* to *rs3023231*), *Scq3* on Chr14 between 9.7 and 33 Mb (*rs3689508* to *rs3669686*), and epistasis of *Scq2* with *Scq1* on Chr1 (between the centromere and *rs13475771*). Mice with different combinations of *Scq1* and *Scq2* genotypes differed more than threefold in daily sucrose intake. This genotype variation was specific to high concentrations of sucrose and did not generalize to low concentrations of sucrose or to other sweeteners. To understand how these genetic variants increase sucrose intake, we measured resting metabolism, glucose and insulin tolerance, and peripheral taste sensitivity. We found that the combinations of *Scq1* and *Scq2* genotypes influenced thermogenesis and the oxidation of fat and carbohydrate. Results of the glucose, insulin tolerance, and taste tests and gustatory nerve recordings ruled out plasma glucose homoeostasis and, to a lesser extent, peripheral taste sensitivity as major contributors to the differences in voluntary sucrose consumption. Our results provide evidence that non-*Sac* genetic loci in mice strongly influence sucrose intake and change whole-body fuel oxidation.

Keywords: sucrose consumption, metabolism, QTL, epistasis interaction, mouse

Introduction

Mice and rats drink sucrose solutions when they are offered. Indeed, they drink so much that providing access to sucrose in addition to food for a few weeks is a reliable model of dietary obesity [1]. However, average sucrose intake differs markedly among inbred mouse and rat strains [2-4]. Moreover, rat strains have been selectively bred to drink large volumes of solutions containing sugar and other sweeteners [5]. These two lines of evidence point to a genetic explanation for differences in this pattern of sucrose consumption. The evidence from mice was confirmed with the discovery of the *Sac* locus [6], which includes *Tas1r3*, a gene coding for one subunit of a sweet taste receptor protein. Variation in this gene accounts for much (but not all) of the variation in intake of sucrose and other sweeteners among mouse strains [7-11] but not among rat strains [12].

Here our goal was to identify sources of genetic variation in sucrose intake in mice that emerged when the *Sac* locus was identical among all mice. Using intercross and backcross mapping populations, we found regions where mice that were similar in sucrose intake shared inherited segments of the genome at a frequency greater than chance. We then bred mice to incorporate the genomic regions associated with higher sucrose intake into a host inbred strain (congenic mice) and evaluated congenic and control mice for their intake of sweeteners and compounds with other taste qualities. We also assessed their metabolism (with indirect calorimetry), gustatory nerve responses (with electrophysiology), and other measures to help us understand the function of these loci.

Methods

***Animal husbandry*.** All animal study procedures were approved by the Monell Chemical Senses Center Institutional Care and Use Committee. Inbred B6BL/6ByJ (B6; stock no: 001139) and 129P3/J (129; stock no: 000690) mice used for the first stages of breeding were purchased from the Jackson Laboratory; we bred all other mice in the animal vivarium at the Monell Center, located in Philadelphia, Pennsylvania (USA). We fed mice Rodent Diet 8604 in all cases (Harlan Teklad, Madison, WI). The mice drank tap water from glass bottles with stainless steel spouts except while being tested (see below). All mice lived in a 12:12-hr light cycle, with lights off at 7 pm, barring unusual circumstances (e.g., power outages). Pups were weaned at 21-30 days of age, and they were housed in same-sex groups.

***Breeding overview*.** We bred and studied five types of mouse mapping populations (**S1 Table**): F_2 intercross, backcross, consomic, congenic (single and double) strains, all derived from B6 and 129 inbred strains. The specific breeding strategy was different for each population, but all were part of a large-scale program to map taste genes [13, 14].

F_2 intercross. All mice had the B6 form of the *Sac* allele [15]. Females of the B6 inbred strain were mated with two males from an incipient congenic strain 129P3/J.C57BL/6ByJ-*Tas1r3* (N_4F_4) to produce F_1 and then F_2 generations.

Backcross and consomic. We bred backcross and consomic mice (**S2 Table**) using marker-assisted methods [13, 14, 16], and some mice produced in this breeding program we used as founders for the congenic strains described below. (We had no consomic mice available for chromosome 14; thus, we did not investigate the *Scq3* locus further.)

Congenic. We branched all congenic strains from consomic strains by mating heterozygous consomic or partially consomic mice with inbred 129 partners. This strategy allowed us to compare littermates with one copy of the donor region (heterozygous; 129/B6) to those without the donor region (homozygous; 129/129). Each congenic mouse was potentially genetically unique (because the donor region could shorten due to meiotic recombination), so we genotyped all congenic mice to define the donor region breakpoints. We named congenic strains with the prefix 'C' (for 129.B6-*Scq1*, we named two strains as C1 and C2; for 126.B6-*Scq2*, seven strains as C3-C9). All 9 strains descended from three progenitors and the residual genomes of the congenic strains were tested and calculated (**S3 Table**). Specifically, we branched *Scq1* congenic mice by backcrossing one partially consomic male from the N_8 generation (a heterozygous male from 129-Chr1^{B6}) to several inbred 129 females. We branched *Scq2* congenic strains by backcrossing several consomic mice (129.B6-Chr9) from the N_7 generation to 129 females. We bred the double congenic mice by mating *Scq2* heterozygous females (B6/129; N_7) with a homozygous *Scq1* male (B6/B6; N_5F_2). Thus, for the *Scq1* locus in the double congenic strain, all offspring were heterozygous (129/B6). For the *Scq2* locus, half the offspring were heterozygous (129/B6) and half were homozygous (129/129).

Limits of breeding. We had difficulty producing enough *Scq2* heterozygous or homozygous mice for in-depth taste and other phenotyping. As we bred *Scq2* congenic strains with shorter and shorter donor regions, what few mice we could produce had to be used for breeding. These breeding problems, similar to those encountered with the corresponding homozygous consomics [14], made it difficult to create a homozygous *Scq2* congenic strain. In contrast, *Scq1* congenic mice were fertile and reproduced

easily, so we did in-depth taste tests and other phenotyping on *Scq1*, consomic, and F_2 mice only.

Genotyping. We extracted and purified genomic DNA from tail tissue either using a sodium hydroxide method [17] or by proteinase K digestion followed by high-salt precipitation (Gentra/Qiagen, Valencia, CA). We measured the concentration and purity of DNA using a small-volume spectrophotometer (Nanodrop, Wilmington DE, USA). We genotyped these DNA samples using two methods: microsatellites and single-nucleotide polymorphisms. For the microsatellites, we used fluorescently labeled primers to amplify genomic DNA by PCR and scanned the products using an ABI 3100 capillary sequencer (Applied Biosystems, Forest City, CA). For the single-nucleotide polymorphisms, we genotyped using primers and fluorescently labeled probes designed to discriminate between alleles, using an ABI Prism 7000 real-time PCR or Step One system (ABI Assay-by-Design, Applied Biosystems, Foster City, CA). All markers and their physical positions are listed in the **S4 Table.**

Phenotyping. We conducted phenotyping experiments with adult mice for taste testing, body composition, metabolism, glucose and insulin tolerance, and gustatory electrophysiology and each mouse mapping population was assessed separately and described as following.

Taste testing (two-bottle choice tests). We performed two types of taste tests: a basic screen for sucrose intake and a more in-depth taste test using sucrose and other taste solutions [18, 19]. For all tests, mice were 8 or more weeks of age. They were trained to drink deionized water from two graduated drinking tubes for at least 2 days. Daily measurements were made in the middle of the light period by reading fluid volume to the nearest 0.1 mL. We measured body weights before and after each test series.

For the basic screening test, we either offered one tube of water and one tube for 300 mM sucrose for 4 days or, in later studies, both tubes containing 300 mM sucrose. (We learned from earlier tests that there was a ceiling effect because mice drank nearly 100% of their daily fluid intake from the 300 mM sucrose tube. Having two tubes of sucrose ensured that even the most avid drinkers would not drink it all in a 24-hour test.)

To perform the in-depth taste tests, we offered one tube of deionized water and another tube with a taste solution (described below), switching positions of the two drinking tubes every day to control for side bias (some mice prefer to drink from one side regardless of the contents of the tube). Between each test series, mice usually had two tubes of water for at least 2 days.

Mice from the F_2 intercross were tested with the following taste solutions in the order listed: 30 mM glycine, 30 mM D-phenylalanine, 1.6 and 20 mM saccharin, 50 and 300 mM sucrose, 300 mM monosodium glutamate (MSG), 3% and 10% ethanol, and 50 mM $CaCl_2$. All consomic mice, 40 of the *Scq1* homozygous congenic mice, and the inbred control 129 and B6 mice were tested with the following: 30 mM glycine, 300 mM NaCl, 20 mM saccharin, 300 mM sucrose, 300 mM MSG, and 3% and 10% ethanol. We purchased all taste compounds from Sigma Chemical Co. (St. Louis, MO) except the ethanol, which was purchased from Pharmco Products (Brookfield, CT).

All backcross, *Scq1* and *Scq2* heterozygous congenic, and double congenic mice were tested with the basic sucrose two-bottle choice test. The homozygous *Scq1* congenics (C1) and control 129 inbred mice were separated into five groups to efficiently test a large number of taste compounds. We tested as follows: group 1, a concentration series of 10, 30, 100, 300, and 600 mM sucrose (N=19 congenic and N=10 controls); groups 2 and 3, concentration series of 10, 30, 100, 300, 600, and 1200 mM glucose

and fructose, respectively (N=10 mice per genotype group); group 4, 4.6% soybean oil, 6% corn starch, and 10.5% maltodextrin (N=19 congenics and 10 controls).

Body composition. We measured the body composition of *Scq1* mice (C1) at 8 and 36 weeks of age using magnetic resonance (MR) methods (Bruker minispec LF110 horizontal whole-body composition rat and mouse analyzer; Bruker BioSpin Corp., Billerica, MA). After the second MR, we also performed body composition analyses by dual-energy x-ray absorptiometry (DEXA; PIXImus II densitometer; GE software, version 2.00; Lunar Corp., Madison, WI), as well as necropsy using anatomic landmarks to identify and remove the organs [20, 21]. We weighed the following organs to the nearest 0.01 g: spleen, heart, pancreas, brain, kidney, liver, six 'white' adipose depots (pericardial, inguinal, retroperitoneal, subscapular, mesenteric, and gonadal), and a 'brown' adipose depot (subscapular). We also measured body length excluding the tail using electronic calipers (Fowler ProMax, Kelley and Kelly Industrial Supply, Syracuse, NY).

Metabolism, food and water intake, and activity. We examined whether *Scq1* genotype affects metabolism, food intake, water intake, or activity, assessing the *Scq1* homozygous congenic (C1) and a control group of inbred 129 mice. Using the TSE LabMaster (version 5.0.6; TSE Systems, Inc., Chesterfield, MO, USA), we measured mice at 8 weeks of age and the same mice again at 36 weeks of age. We trained the mice for 3-5 days in cages that mimicked the experimental cages to ensure they learned to eat and drink appropriately from the suspended food and water containers. We transferred the mice to the experimental cages and measured oxygen consumption, carbon dioxide production, food and water intake, and physical activity in three dimensions (including rearing as well as walking). We estimated heat production from carbon dioxide and oxygen consumption corrected for lean body mass [22].

Glucose and insulin tolerance tests. We tested homozygous *Scq1* congenic mice (C1) and 129 inbred mice twice (at 8 weeks and again at 36 weeks of age) for changes in plasma glucose in response to exogenous glucose (glucose tolerance test; GTT) or insulin administration (insulin tolerance test; ITT) [23]. All mice were tested during the light phase and had unrestricted access to food prior to testing.

For the GTT, we gave the mice glucose (2 g/kg, 10 ml/kg body weight; Sigma-Aldrich St. Louis, MO) either intraperitoneally (IP) (B6/B6 vs 129/129: N=20 vs 12) or by intragastric gavage (IG) (B6/B6 vs 129/129: N=21 vs 12). For the IP route of administration, the glucose was dissolved in 0.9% saline, whereas for the IG route of administration, the glucose was dissolved in deionized water. For the ITT, we injected the mice (B6/B6 vs 129/129: N = 19 vs 11) with insulin (2 U/kg, IP; insulin, Novo Nordisk A/S, Bagsvaerd, Denmark).

We sampled the blood from the tip of the tail, with duplicate measures at each time point (before and after the administration of glucose or insulin) using a One Touch Ultra glucometer (LifeScan, Inc., USA). For the GTT, we sampled the blood at 0, 5, 15, 30, 45, 60, 90, and 120 min; for the ITT, we sampled at 0, 5, 30, 60, 90, and 120 min. For the GTT, animals were in restraint tubes during the blood draws (they were habituated to them before the tests). During the ITT, mice were unrestrained in their home cages.

Gustatory electrophysiology. We conducted the electrophysiological experiment with heterozygous *Scq1* (C1) and homozygous *Scq2* congenic mice (C7) and host 129 inbred mice (5 for each group) as described elsewhere [24-26]. Mice ranged from 26 to 50 weeks of age. Activity of the whole chorda tympani nerve in response to lingual application of taste solutions was electrophysiologically recorded for the following stimuli

(in mM, except for ethanol in %): 20 quinine hydrochloride, 100 NaCl; 10 HCl; 1000 glucose; 1000 maltodextrin; 1000 fructose; 100, 300, 1000 sucrose; 2, 6, 20 saccharin; 100, 300, 1000 MSG; 3, 10 ethanol. Each taste solution was washed over the tongue for 30 sec; between taste stimuli presentations, the tongue was rinsed with deionized water for at least 1 min. The magnitude of the integrated response at 20 sec after stimulus onset was measured and expressed as a proportion of the average of the previous and following responses to 100 mM NH_4Cl [24] .

Data analyses. We performed several steps to prepare the data for the main statistical analysis for each mouse population. We checked the distribution of the phenotype for normality within each mapping population using the Lilliefors test (**S5 Table**), and non-normal data were transformed as appropriate [27]. Male and female mice were similar in sucrose intake, so we pooled their data. For each mapping population we separately checked the correlations between voluntary sucrose intake and (a) body weight, (b) daily habitual water intake, and (c) age (**S1 Figure**). (Mice that habitually drank the most water also drank the most sucrose, so we used habitual water intake as a covariate in the relevant statistical analyses.) In addition to intake, we also computed preference scores as a ratio of taste solution intake to total fluid intake (taste solution plus water), in percent, although this preference score was an imperfect measure because many mice drink nearly 100% of their fluid intake as 300 mM sucrose when also offered plain water. For all statistical models, we used a type 1 (sequential) sum-of-squares model, and when testing for group differences, we used Tukey's HSD tests. For all data analyses, we computed the statistical tests with R (version 3.3.3) and R-studio (version 1.0.136) and graphed the results using either R or Prism 6 (version 6.05; GraphPad Software, La Jolla, CA).

F_2 intercross. We identified genomic regions shared in common by the F_2 mice that drank the most sucrose using markers from 19 mouse autosomes and X chromosome (R package R/QTL, version 1.41 [28]). We estimated genotype probabilities and genotype errors using the *calc.genoprob* function, with interval mapping by maximum likelihood estimation (EM algorithm) for the main-effect quantitative trail loci (QTLs) using the *scanone* function. We tested the significance of each marker regression against 1000 permutations of the observed data using the n.*perm* function. We defined the confidence intervals as drops of 1 LOD from the peak using the *lodint* function. We estimated the explained phenotypic variance for a single locus as $1 – 10^{-2\ LOD\ /\ n}$, where n is the sample size, using the LOD score from *scanone* [28]. We investigated marker pair interactions using the *scantwo* function, evaluating significance against the 1000 permutation tests and confirming the interaction in a general linear model using the single marker and the interaction of marker pairs as fixed factors.

To assess whether markers identified in the analysis also affected the consumption of other taste solutions tested in the F_2 population, we analyzed solution intakes in a general linear model using genotype of each locus as a fixed factor. We analyzed preference scores of these solutions using one-way ANOVA followed by post hoc tests. In addition, we computed Pearson correlations among all taste solutions using intakes (adjusted for habitual water intakes) and preference scores separately.

Backcrosses. We pooled four backcross generations to form one mapping population (N_3 to N_6) and conducted a general linear model analysis of sucrose intake with genotype as a fixed factor. For each marker, we calculated (a) the genotype means, (b) the p-value test statistic as the negative base 10 logarithm, and (c) the effect size using Cohen's *D* [29]. We report these values, including confidence intervals (defined by 2 units of $–\log_{10}$ p-value drop), for the peak marker from the mapping population.

Statistical thresholds were computed with a Bonferroni correction to an α level of 0.05 for the number of markers [N=67; –log(α/N) ≈ 3.13].

Consomics. Building on the results from the F_2 intercross, which suggested a region of interest on chromosome 9 and an interaction between chromosomes 1 and 9, we studied the appropriate consomic mice for these chromosomes. For chromosome 9 we had reciprocal consomic strains, whereas for chromosome 1 we had only one consomic strain. For these three consomic strains, we analyzed sucrose intakes with a general linear model, treating strain as a fixed factor and using the relevant host inbred strain for comparisons (129-Chr 1^{B6} vs inbred 129, 129-Chr 9^{B6} vs inbred 129, and B6-Chr 9^{129} vs inbred B6). We calculated the genotype effect size with Cohen's *D*, and we analyzed the intakes for taste solutions other than 300 mM sucrose using the same statistical approach. We also used a general linear model to evaluate all consomic and control strains with genotype as a fixed factor and habitual water intake as a covariate, followed by post hoc tests. We analyzed preference scores by strain using t-tests.

Heterozygous congenic analysis. We analyzed the data from congenic mice using the common segment method to find genomic regions shared in common by strains that have the same trait [30]. To that end, for all analyses of congenic mice, we used sucrose intake as the outcome measure and the presence or absence of a donor region as a fixed factor, using a proxy marker for the donor region (*rs3708040* for *Scq1* and *rs13480399* for *Scq2*). If the congenic strain differed from the control strain in voluntary sucrose consumption, we concluded the associated donor region contained the causal genetic variant; if not, then we concluded otherwise.

Homozygous congenic analysis. We conducted some of the taste tests with a concentration series and some with a single concentration (see above). For the concentration series (e.g., fructose), we conducted a two-way mixed-design ANOVA with

strain as the fixed factor (congenics vs host inbred 129) and concentration as the repeated measure. For the single concentrations (e.g., saccharin), we analyzed solution intakes using strain as a fixed factor (congenics vs inbred 129). When informative (no ceiling effect, i.e., toward 100% preference), we analyzed preference scores by t-tests (e.g., for the preferences of soybean oil, corn starch, and maltodextrin).

Body composition. We evaluated the effect of *Scq1* genotype on body composition using a two-way ANOVA (genotype by sex) followed by post hoc tests for all three measures of body composition: MR, DEXA, and necropsy. Within each mapping population, we also used two-way ANOVAs (genotype by sex) to determine whether mice with genotypes that increased voluntary sucrose intake also gained more body weight during the 4-day sucrose intake tests.

Metabolism. We averaged all measures from the TSE LabMaster over the four 24-hr periods. For each individual, parameters were corrected by lean body mass. We used t-tests to compare groups (congenics B6/B6 vs host inbred 129/129).

Glucose and insulin tolerance tests, body composition analysis. We analyzed data from GTT and ITT using one-way repeated ANOVA with genotype as a fixed factor and time of blood collection as the repeated measure.

Electrophysiology. Chorda tympani responses to individual taste solutions were not normally distributed, so we used nonparametric Mann-Whitney U-tests to assess differences between genotype groups [31].

Evaluating genes and variants in the Scq1 and Scq2 region. Using the results of the congenic mapping, we identified specific genomic regions that likely contained the causal variant(s) for each QTL. Within those regions, we identified the known variants and evaluated them in several ways. First, we made a list of all genetic variants using an

online database [32], which contains the results of a large-scale genome sequencing project of many inbred mouse strains, including the 129P2/OlaHsd and C57BL/6J strains [33, 34]—of the inbred mouse strains sequenced, these two strains are the most closely related to the parental inbred strains used in our study [35]. From the list of genetic variants we compiled, we identified sequence variants with the potential to cause functional changes using the Sorting Intolerant From Tolerant algorithm [36]. Twelve genetic variants were excluded from this analysis because we could find no unique identifier for the variant in the dbSNP database (Build 138) [37].

Next, using another online database [38], we identified all genes residing within the genomic coordinates of *Scq1* and *Scq2* defined in the congenic strains and classified these candidate genes using the PANTHER classification system [39]. Finally, using an internet-based tool kit [40], we prioritized candidate genes by comparing their functional annotations against a training set that consisted of 332 known carbohydrate metabolism mouse genes from the KEGG database [41]. We used the Entrez identifiers of the genes within the *Scq1* and *Scq2* regions as input for candidate gene prioritization; therefore only genes with an appropriate identifier were included.

We reasoned that the causal genes from *Scq1* and *Scq2* jointly contribute to the same pathway, so we identified gene pairs that met this criterion (i.e., one gene-pair member was from the *Scq1* region and the other was from the *Scq2* region). We thus assessed the protein-protein associations by searching the protein names in the STRING database [42] with multiple active interaction sources that included three types of data: (1) direct experimental evidence of protein-protein interactions; (2) less direct evidence generated by grouping proteins of known function into metabolic, signaling, or transcriptional pathway; and (3) *in silico* computation techniques for detecting new protein-protein associations. We exported gene pairs for protein-protein interactions with

association scores > 0.9, which were computed by benchmarking them against the training set [42], and identified those pairs that fit our criterion

Results

Initial identification of QTLs and the interaction loci

We mapped QTLs for 300 mM sucrose intake on chromosome 9 (*Scq2*) and chromosome 14 (*Scq3*), with the B6 allele being dominant and increasing sucrose consumption in both cases (**Figure 1A**). The peak *Scq2* LOD score was 28.12 at marker *rs3023231* (106.8 Mb), with a confidence interval from 105 to 108 Mb (between *rs13480395* and *rs4227916*); it explained 37% of mouse-to-mouse variation in sucrose intake. The peak *Scq3* LOD score was 4.13 at marker *D14Mit126* (21.8 Mb), with a confidence interval from 9.7 to 33 Mb (between *rs3689508* and *rs3669686*); this explained about 7% of the trait variance. The results of the two-QTL model genome scan identified a significant interaction of a marker pair (*Scq1* on chromosome 1 and *Scq2**;*** **Figure 1B**) for daily sucrose intake, and this epistatic interaction was confirmed by general linear model analysis (**Figure 1C**).

Variation at these three loci (*Scq1*, *Scq2*, and *Scq3*) had similar effects on the taste preferences of the mice, strongly affecting voluntary consumption of sucrose but not other taste compounds (**S2 Figure**), with the following exceptions: compared with mice with the 129/129 genotypes, mice with one or more *Scq1* B6 alleles drank more of 20 mM saccharin, and mice with one or two *Scq2* B6 alleles drank more 3% and 10% ethanol and 300 mM MSG; compared with mice with the B6/B6 genotype, mice heterozygous for the *Scq2* locus had higher intakes and preference scores for $CaCl_2$. Independent of genotype, mice that drank the most 300 mM sucrose did not tend to

drink more of the other taste solutions, except for a tendency to drink more MSG and ethanol (**S3 Figure**).

Confirmation of QTL and the epistasis effect

Backcross. In the backcross population, we detected *Scq2* at the peak marker *rs302323131*; this is the same location and has an allelic effect of the same direction and magnitude as in the F_2 mapping population, with an LOD score of 15.92 and a Cohen's *D* value of >0.8 (**Figure 2A**). (We were not able to test for the effects of *Scq1* or *Scq3* because these backcross mice were not genotyped for markers on other chromosomes.)

Consomics. The results from the consomics also matched those from the F_2 and backcross populations, confirming the main effects of *Scq1* and *Scq2* and their interaction. For *Scq1*, mice with a B6-introgressed chromosome 1 drank less sucrose solution than did the host control mice (**Figure 2C**), with an effect size > 1.5 (**Figure 2D**). For *Scq2,* mice with a B6-introgressed chromosome 9 drank more sucrose and mice with the 129-introgressed chromosome 9 drank less sucrose compared with host controls, with effect sizes of 1.1 and 2.7, respectively (**Figure 2C, D**). Mice with B6 alleles at both *Scq1* and *Scq2* drank much more sucrose than did mice with 129 alleles at both loci (**Figure 2C, E**). We evaluated all consomic and control mice in a single model and found that mice with different genotypes at *Scq1* and *Scq2* differed in voluntary sucrose intake, with a marked increase in mice with certain combinations of those alleles [**Figure 2E**; *Scq1*: $F(1, 225)=36.8$, $p<0.00001$; *Scq2:* $F(1, 225)=166.9$, $p<0.00001$; *Scq1* × *Scq2:* $F(1, 225)=19.4$, $p<0.0001$]. Consomic mice also differed in their intake of other taste solutions compared with mice from the inbred host strains (**S4 Figure**).

Congenics, Scq1 *region*. The results from congenic mapping support the presence of *Scq1* near the centromere of chromosome 1 (**Figure 3**). We bred two *Scq1* congenic strains (C1 and C2; **Figure 3B**, **S6 Table**). Congenic mice from the C1 strain (with the donor region from the centromere to 25.4 Mb; *rs13475771*) drank more sucrose than did control mice (**Figure 3C**), but control and congenic mice from the C2 congenic strain without this part of the donor region drank the same amount of sucrose (**Figure 3D**). This mapping result was supported by the outcome of the general linear model, with the strongest relationship between sucrose intake and genotypes near the centromere (**Figure 3A**).

Congenics, Scq2 region. The results from congenic mapping support the presence of *Scq2* in the same region on chromosome 9 identified in the F_2 and backcross mapping populations (105.7-106.9 Mb; **Figure 4**). We bred seven *Scq2* congenic strains (C3–C9), and those strains with donor regions within the 1.2-Mb region drank more sucrose compared with control mice, whereas those without this donor region did not (**Figure 4C, S7 Table**). The mapping result was supported by the outcome of the general linear model, with the strongest relationship between sucrose intake and genotype within this 1.2-Mb region (**Figure 4A**). We show the relevant genes in **Figure 4D** as a screen shot from the Ensembl Mouse Genome Brower [43].

Scq1 and Scq2 interaction. We bred a double congenic line that retains *Scq1* and *Scq2* from the B6 strain with a 129 genetic background (**Figure 2F**). At *Scq1* they had a fixed heterozygous (129/B6) genotype; at *Scq2* they segregated: half were 129/B6 and half were 129/129; **Figure 2G**). Mice that had B6 alleles at both loci drank six times more sucrose than those with 129 alleles at both loci (Cohen's *D*=5.0); mice with only one B6 allele drank much less (*Scq1*: Cohen's *D*=0.9; *Scq2*: Cohen's *D*=0.8) (**Figure 2H, I**). The genotypes for *Scq1* and *Scq2,* refer to **Figure 2G** but except for three congenic

strains C2, C3 and C4, we recode these three strains as 129/129 because these trains don't retain the QTLs (**Figure 3B, D** and **Figure 4B, C**). A single model encompassing all the congenic data produced very similar results to those of the F_2 and consomics: mice differed in voluntary sucrose intake depending on the combination of *Scq1* and *Scq2* genotypes [*Scq1*: F(1, 989)=8.65, p<0.005; *Scq2*: F(1, 989)=109.72, p<0.00001; *Scq1* × *Scq2*: F(1, 989) =40.16, p<0.00001 (**Figure 2J**)].

Scq1 homozygosity. Possessing one *Scq1* B6 allele was sufficient to produce detectable increases in voluntary sucrose consumption, but this effect was larger for the mice with two B6 alleles, that is, in a congenic strain we bred to homozygosity [effect size of 2.6 (**Figure 5**)].

Characterization of the *Scq1* QTL effect

For characterization of the *Scq1* QTL effect, we used the homozygous *Scq1* congenics (B6/B6) compared with 129 host inbred mice (129/129). The *Scq1* locus affected the consumption of 300 and 600 mM sucrose, 600 mM glucose, and 600 mM fructose, but it had no effects at other, mostly lower concentrations of these sugars (**Figure 6**). Likewise, the *Scq1* mice did not differ from controls in their consumption of 20 mM saccharin, 30 mM glycine, or 10% ethanol (**S5A, B, F Figure**) or of 4.6% soybean oil, 6% corn starch, or 10.5% maltodextrin (**S6A-C Figure**). However, there were differences between the two groups of mice in their intake of 300 mM NaCl, 300 mM MSG, and 3% ethanol (**S5C-E Figure**).

Body composition analysis. We used three methods to study body composition, MR, DEXA, and necropsy, all of which indicated *Scq1* genotype had no significant effect on body composition (MR, **S8A Figure**; DEXA, **S8B Figure**; necropsy, **S9 Figure**). From the necropsy data, we learned that mice of different genotypes were similar in most

organ weights, except the subscapular adipose depot [42] and heart [F(1,23)=22.5, P<0.0001] (**S9 Figure**). We also learned that drinking sucrose voluntarily for 4-12 days did not result in significant differences in body weight gain among mice of different *Scq* genotypes (**S10 Figure**).

Metabolism. Homozygous *Scq1* mice differed in their metabolic gas exchange compared with inbred host controls, especially in early adulthood. Specifically, mice from the C1 strain (B6/B6) consumed less the oxygen [8 wks: t(1,23) = -4.71, p<0.001; 36 wks: t(1,23) = -3.16 , p<0.01], produced less carbon dioxide [8 wks: t(1,23)=4.03, p<0.001; except for 36 wks: t(1,23)=0.53, p = 0.598], generated less heat [kcal/kg lean, hr; 8 wks: t(1,23) = -5.22, p<0.001; 36 wks: t(1,23) = -3.49, p<0.01], and burned proportionately more calories as fat than carbohydrate compared with inbred 129 mice (129/129; **Figure 7C-F**). However, at both ages tested, these homozygous *Scq1* mice ate the same amount of food (grams per 24 hr), drank the same amount of water (milliliters per 24 hr), and were as active (beam breaks/hr) as inbred 129 mice. The statistical results are as follows: food [8 wks: t(1,23) = -1.11, p=0.282; 36 wks: t(1,23) = -1.10, p=0.284], water [8 wks: t(1, 23)=0.54, p=0.597; 36 wks: t(1,23) = -0.96, p=0.348], and activity [8 wks: t(1,23) = -1.85, p=0.077; 36 wks: t(1,23) = -0.001, p=0.999] (**Figure 7A, B, G**).

Glucose and insulin tolerance tests. Compared with inbred 129 mice, homozygous *Scq1* mice did not differ in plasma glucose or plasma insulin in response to exogenous glucose, administered either by IP injection or IG gavage, at any age tested (**S7A, B, C Figure**), with details as follows: IP, 8 weeks [genotype F(1,30)=1.52, p=0.227; time F(7,210)=73.5, p<0.0001, genotype × time interaction F(7,210)=1.70, p=0.111]; IP, 36 weeks [genotype F(1,24)=0.19, p=0.665; time F(7,168)=107.9, p<0.0001; genotype × time interaction F(7,168)=1.70, p=0.112]; IG, 8 weeks [genotype

F(1,31)=1.95, p=0.173; time F(7, 217)=69.4, p<0.0001; genotype × time interaction [F(7,217)=0.18, p=0.989]; IG, 36 weeks [genotype F(1,24)=0.065, p=0.800; time F(7,168)=62.2, p<0.0001; genotype × time interaction F(7, 168)=0.53, p=0.813].

After injection with insulin, 8-week-old mice of different *Scq1* genotypes did not differ in plasma glucose except for at one time point [genotype F(1,28)=8.68, p=0.006; time F(5, 140)=101.2, p<0.0001; genotype × time interaction F(5,140)=3.70, p=0.003] (**S7C Figure,** left panel). These same mice at 36 weeks showed no effects of genotype at all and no interaction of genotype over time [genotype F(1,23)=1.012, p=0.325; time F(5,115)=80.060, p<0.0001; genotype × time interaction F(5,115)=1.339, p=0.253] (**S7C Figure,** right panel).

Electrophysiology. Mice of different *Scq1* or *Scq2* genotypes did not differ significantly from inbred host controls in the rate or amplitude of taste nerve firing (chorda tympani) for any of the taste stimuli tested (100 and 300 mM sucrose; 20 mM quinine; 100 mM NaCl; 10 mM HCl; 1000 mM glycine; 1000 maltose; 1000 mM fructose; 2, 6, and 20 mM saccharin; 100, 300, and 1000 mM MSG; and 3% and 10% ethanol; **S11 Figure**). The only taste solution they responded differentially to was limited to the *Scq1* group's response to 1000 mM sucrose (p=0.03; **S8 Table**)

Candidate genes

We identified and investigated genes in the *Scq1* and *Scq2* regions defined by the congenic mapping results (*Scq1*, 1:1-25377067; *Scq2*, 9:105650529-106888925). *Scq1* harbors 193 genes, and *Scq2* harbors 55 genes (**S9 Table**), of which 49 are protein-coding genes [21] and 199 are other gene types, such as long noncoding RNA (lncRNA) genes. Among 41 of the protein-coding genes, there were 168 relevant missenses and stop codon gain/loss variants (**S9 Table**). We grouped the biological

processes for these genes into categories (**S12A Figure**), and in doing so, we found that 31 of these known genes are involved in carbohydrate metabolism (**S9 Table**). We identified pairs with the highest association confidence scores and that contained one gene from the *Scq1* and another from the *Scq2* region. For example *Rb1cc1* from *Scq1* and *Pik3r4* from *Scq2* have an interaction confidence score > 0.9 (**S12B Figure**). The prioritization of these two genes ranks 1st and 11th based on the similarity to other genes involved in carbohydrate metabolism, and one member of the pair has a missense variation (*Pik3r4*; **S9 Table**). Additional gene pairs also had protein-protein association scores > 0.9: *Oprk1-Grm2* (opiate and glutamine receptors), *Grm2-Npbwr1* (glutamine and neuropeptide receptors), and *Rpl29-Rpl*7 (ribosomal proteins) (**S10 Table**).

Discussion

We identified three small regions of the mouse genome containing variations that affect the voluntary intake of concentrated sucrose solution, which emerged when influential variation in the *Tas1r3* gene was removed. We used B6 and 129 mice for these experiments in part because we were familiar with the marked differences in sucrose consumption between the two strains [44], and their utility was confirmed by the initial intercross results showed a striking influence of a locus on chromosome 9 that interacts with a separate region on chromosome 1.

Mice with the B6 version of the *Scq1* and *Scq2* alleles had a distinct phenotype: they consumed large amounts of high concentrations of simple sugars—especially sucrose but also glucose and fructose—but not polymeric carbohydrates (corn starch and maltodextrin), non-caloric sweeteners (saccharin), or most other taste solutions, with the exception of ethanol. This seems to be a 'sugar' phenotype, which includes fermented and non-fermented sugars, that connects to whole-body fuel use. Under basal conditions, mice with genotypes of *Scq1* that decreased voluntary sucrose consumption also oxidized more carbohydrate and less fat than did mice with opposing genotypes. These genetic effects on voluntary sucrose intake and carbohydrate oxidation do not seem to arise from the earliest steps of intestinal absorption because there were few or no differences by genotype in the outcomes of glucose and insulin tolerance tests. Peripheral taste sensitivity might be a contributor to the differences in voluntary sucrose consumption, but three points make this explanation unlikely: (a) genotype affects voluntary intake at high but not low sucrose concentrations (the ability to detect the taste is more important at lower concentrations), (b) genotype effects were absent or small for

peripheral nerve responses, and (c) these small effects were in the opposite direction from those of another genetic model of voluntary sucrose intake [24].

There are at least two ways to approach the evaluation of genes within the critical intervals for their effect on voluntary sucrose consumption. The first method is to examine all genes within the regions to determine if there are known connections between those genes and voluntary sucrose consumption (narrowly) and metabolism (more broadly). No genes were linked to voluntary sucrose consumption in the databases surveyed, but many were part of carbohydrate and fat metabolism pathways, including *Pik3r4* [45], *Alas1* [46], *Acy1* [47], *Cops5* [48], *Atp6v1h* [49, 50], *Tlr9* [51], *Lypla1* [52, 53], *Pkhd1* [54], and *Gdap*1 [55]. We further reasoned that the interaction of the *Scq1* and *Scq2* regions indicated that causal genes might be part of the same pathway. That approach led us to identify particular gene pairs that jointly contribute to a shared function (e.g., *Pik3r4* and *Rb1cc1*); however, their involvement in voluntary sucrose intake awaits direct tests.

The second approach is to find genes or gene pathways that have a similar constellation of traits, with the idea that the genes in the region could be unknown members of the known pathway. Using this second method, we learned that many of the traits of the *Scq1* and *Scq2* mice studied here are similar to mice with knockout or overexpression of the *FGF21* gene. These traits include the increased intake of sugar but not of non-caloric sweeteners or corn starch, and the lack of large effects on peripheral taste nerves and body weight or fatness [56]. A next step would be to evaluate the genes in the region experimentally for their roles in the *FGF21* signaling pathway. We note that two proteins that regulate this signaling pathway (*Prdm14* [57] and *Sulf1* [58]) are located in the *Scq1* region.

Genetic controls of sucrose intake

Through our use of consomic mouse strains, we indirectly mapped other genes for taste-related traits. Consomic strains contain an entire introgressed chromosome; thus, there is a greater potential to capture more causal genes and their variants compared with the congenic strains, which have smaller introgressed regions. As expected, our consomic mapping results discovered new regions of interest distinct from the Scq loci and also confirmed many previously reported regions that contain variation that affects the intake several other taste compounds by mice [59-62].

This study was conducted in mice, but what we learned has connections to human health. The contribution of sucrose, intake especially in beverages, to the ills of the modern diet remains controversial, but the benefits of understanding the biological causes of excessive sugar consumption are likely to be helpful. An earlier example of this type of benefit comes from the discovery of a subunit of the mouse sweet receptor, T1R3 [7-11, 63]. This discovery was useful in understanding taste biology, but it is now apparent that this sugar-sensing receptor has many nongustatory functions, including insulin regulation [64] and adipogenesis [65]. Likewise, we hope this study provides insights toward understanding the biological drive to consume sugar and that this knowledge may be harnessed to prevent the negative effects of excessive sugar consumption by understanding and allaying the human 'sweet tooth'.

References

1. Ramirez, I., *When does sucrose increase appetite and adiposity?* Appetite, 1987. **9**(1): p. 1-19.
2. Lewis, S.R., et al., *Inbred mouse strain survey of sucrose intake.* Physiol Behav, 2005. **85**(5): p. 546-56.
3. Lush, I.E., *The genetics of tasting in mice. VI. Saccharin, acesulfame, dulcin and sucrose.* Genetical Research, 1989. **53**: p. 95-99.
4. Tordoff, M.G., L.K. Alarcon, and M.P. Lawler, *Preferences of 14 rat strains for 17 taste compounds.* Physiol Behav, 2008. **95**(3): p. 308-32.
5. Dess, N.K., *Responses to basic taste qualities in rats selectively bred for high versus low saccharin intake.* Physiol Behav, 2000. **69**(3): p. 247-57.
6. Fuller, J.L., *Single-locus control of saccharin preference in mice.* J Hered, 1974. **65**(1): p. 33-6.
7. Kitagawa, M., et al., *Molecular genetic identification of a candidate receptor gene for sweet taste.* Biochem Biophys Res Commun, 2001. **283**(1): p. 236-42.
8. Max, M., et al., *Tas1r3, encoding a new candidate taste receptor, is allelic to the sweet responsiveness locus Sac.* Nat Genet, 2001. **28**(1): p. 58-63.
9. Montmayeur, J.P., et al., *A candidate taste receptor gene near a sweet taste locus.* Nat Neurosci, 2001. **4**(5): p. 492-8.
10. Sainz, E., et al., *Identification of a novel member of the T1R family of putative taste receptors.* J Neurochem, 2001. **77**(3): p. 896-903.
11. Bachmanov, A.A., et al., *Positional cloning of the mouse saccharin preference (Sac) locus.* Chemical Senses, 2001. **26**(7): p. 925-933.
12. Lu, K., et al., *No relationship between sequence variation in protein coding regions of the Tas1r3 gene and saccharin preference in rats.* Chem Senses, 2005. **30**(3): p. 231-40.
13. Lin, C., et al., *Adiposity QTL Adip20 decomposes into at least four loci when dissected using congenic strains.* PLoS One, 2017. **12**(12): p. e0188972.
14. Lin, C., et al., *Body composition qtls identified in intercross populations are reproducible in consomic mouse strains.* PLoS One, 2015. **10**(11): p. e0141494.
15. Reed, D.R., et al., *Polymorphisms in the taste receptor gene (Tas1r3) region are associated with saccharin preference in 30 mouse strains.* J Neurosci, 2004. **24**(4): p. 938-46.
16. Lin, C., et al., *Burly1 is a mouse QTL for lean body mass that maps to a 0.8-Mb region of chromosome 2.* Mamm Genome, 2018.
17. Truett, G.E., et al., *Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (HotSHOT).* Biotechniques, 2000. **29**(1): p. 52, 54.
18. Tordoff, M.G. and A.A. Bachmanov, *Influence of test duration on the sensitivity of the two-bottle choice test.* Chem Senses, 2002. **27**(9): p. 759-68.
19. Tordoff MG, B.A., *Monell mouse taste phenotyping project. Monell Chemical Senses Center.* www.monell.org/MMTPP 2001.
20. Cinti, S., *The adipose organ.* Milano: Editrice Kurtis, 1999.
21. Hayakawa, T., H. Yamasita, and T. Iwaki, *A Color Atlas of Sectional Anatomy of the Mouse.* Braintree Scientific Inc, 2001.
22. Mouse Genomes Project - Query SNPs, i.o.S.I.c.A.f.h.w.s.a.u.s.M.S.r.-.

23. Murovets, V.O., A.A. Bachmanov, and V.A. Zolotarev, *Impaired Glucose Metabolism in Mice Lacking the Tas1r3 Taste Receptor Gene.* PLoS One, 2015. **10**(6): p. e0130997.
24. Inoue, M., et al., *Allelic variation of the Tas1r3 taste receptor gene selectively affects behavioral and neural taste responses to sweeteners in the F_2 hybrids between C57BL/6ByJ and 129P3/J mice.* J Neurosci, 2004. **24**(9): p. 2296-303.
25. Inoue, M., et al., *Soa genotype selectively affects mouse gustatory neural responses to sucrose octaacetate.* Physiological Genomics, 2001. **5**: p. 181-186.
26. Inoue, M., et al., *Whole-nerve chorda tympani responses to sweeteners in C57BL/6ByJ and 129P3/J mice.* Chemical Senses, 2001. **26**: p. 915-923.
27. Delignette-Muller, M., Pouillot, R., Denis, J., and Dutang, C. , *fitdistrplus: Help to Fit of a Parametric Distribution to Non-Censored or Censored Data. R package. 2014.*
28. Broman, K.W., et al., *R/qtl: QTL mapping in experimental crosses.* Bioinformatics, 2003. **19**(7): p. 889-90.
29. Cohen, J., *Statistical power analysis for the behavioral sciences. 2nd ed. 1988, Hillsdale, N.J.: L. Erlbaum Associates. xxi, 567 p.*
30. Shao, H., et al., *Analyzing complex traits with congenic strains.* Mammalian genome : official journal of the International Mammalian Genome Society, 2010. **21**(5-6): p. 276-86.
31. Inoue, M., et al., *Allelic variation of the Tas1r3 taste receptor gene selectively affects taste responses to sweeteners: evidence from 129.B6-Tas1r3 congenic mice.* Physiol Genomics, 2007. **32**(1): p. 82-94.
32. Anonymous, *Mouse Genomes Project - Query SNPs, indels or SVs*. 2011: Wellcome Trust Sanger Institute.
33. Keane, T.M., et al., *Mouse genomic variation and its effect on phenotypes and gene regulation.* Nature, 2011. **477**(7364): p. 289-94.
34. Yalcin, B., et al., *Sequence-based characterization of structural variation in the mouse genome.* Nature, 2011. **477**(7364): p. 326-9.
35. Yang, H., et al., *Subspecific origin and haplotype diversity in the laboratory mouse.* Nat Genet, 2011. **43**(7): p. 648-55.
36. Anonymous. *SIFT web server: predicting effects of amino acid substitutions on proteins*. 2012 [cited 2016 10/7/2016]; August 15th 2015:[W542-7 (Link)]. Available from: http://siftdna.org/www/SIFT_dbSNP.html.
37. Anonymous. *dbSNP, Single Nucleotide Polymorphisms; http://www.ncbi.nlm.nih.gov/SNP/index.html*. Build 138 2016 [cited 2016 10/7/2016]; Available from: http://www.ncbi.nlm.nih.gov/SNP/index.html.
38. Anonymous, *Mus musculus (laboratory mouse) genome view*. 2015, National Center for Biotechnology Information
39. Mi, H., A. Muruganujan, and P.D. Thomas, *PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees.* Nucleic Acids Res, 2013. **41**(Database issue): p. D377-86.
40. Chen, J., et al., *ToppGene Suite for gene list enrichment analysis and candidate gene prioritization.* Nucleic Acids Res, 2009. **37**(Web Server issue): p. W305-11.
41. Kanehisa, M., et al., *KEGG: new perspectives on genomes, pathways, diseases and drugs.* Nucleic Acids Res, 2017. **45**(D1): p. D353-D361.
42. von Mering, C., et al., *STRING: known and predicted protein-protein associations, integrated and transferred across organisms.* Nucleic Acids Res, 2005. **33**(Database issue): p. D433-7.
43. Hubbard, T.J., et al., *Ensembl 2007.* Nucleic Acids Res, 2007. **35**(Database issue): p. D610-7.

44. Bachmanov, A.A., M.G. Tordoff, and G.K. Beauchamp, *Sweetener preference of C57BL/6ByJ and 129P3/J mice.* Chemical Senses, 2001. **26**: p. 905-913.
45. Nemazanyy, I., et al., *Defects of Vps15 in skeletal muscles lead to autophagic vacuolar myopathy and lysosomal disease.* EMBO Mol Med, 2013. **5**(6): p. 870-90.
46. Simcox, J.A., et al., *Dietary iron controls circadian hepatic glucose metabolism through heme synthesis.* Diabetes, 2015. **64**(4): p. 1108-19.
47. Alexandre-Gouabau, M.C., et al., *Postnatal growth velocity modulates alterations of proteins involved in metabolism and neuronal plasticity in neonatal hypothalamus in rats born with intrauterine growth restriction.* J Nutr Biochem, 2012. **23**(2): p. 140-52.
48. Tian, L., et al., *Essential roles of Jab1 in cell survival, spontaneous DNA damage and DNA repair.* Oncogene, 2010. **29**(46): p. 6125-37.
49. Olsson, A.H., et al., *Decreased expression of genes involved in oxidative phosphorylation in human pancreatic islets from patients with type 2 diabetes.* Eur J Endocrinol, 2011. **165**(4): p. 589-95.
50. Baumeier, C., et al., *Caloric restriction and intermittent fasting alter hepatic lipid droplet proteome and diacylglycerol species and prevent diabetes in NZO mice.* Biochim Biophys Acta, 2015. **1851**(5): p. 566-76.
51. Lewis, C.J. and B.A. Cobb, *Carbohydrate oxidation acidifies endosomes, regulating antigen processing and TLR9 signaling.* J Immunol, 2010. **184**(7): p. 3789-800.
52. Hacia, J.G., et al., *Determination of ancestral alleles for human single-nucleotide polymorphisms using high-density oligonucleotide arrays.* Nat Genet, 1999. **22**(2): p. 164-7.
53. Mootha, V.K., et al., *Integrated analysis of protein composition, tissue diversity, and gene regulation in mouse mitochondria.* Cell, 2003. **115**(5): p. 629-40.
54. Kaimori, J.Y., et al., *Polyductin undergoes notch-like processing and regulated release from primary cilia.* Hum Mol Genet, 2007. **16**(8): p. 942-56.
55. Lopez Del Amo, V., et al., *A Drosophila model of GDAP1 function reveals the involvement of insulin signalling in the mitochondria-dependent neuromuscular degeneration.* Biochim Biophys Acta, 2017. **1863**(3): p. 801-809.
56. von Holstein-Rathlou, S., et al., *FGF21 mediates endocrine control of simple sugar intake and sweet taste preference by the liver.* Cell Metab, 2016. **23**(2): p. 335-43.
57. Grabole, N., et al., *Prdm14 promotes germline fate and naive pluripotency by repressing FGF signalling and DNA methylation.* EMBO Rep, 2013. **14**(7): p. 629-37.
58. Buresh, R.A., et al., *Sulfatase 1 is an inhibitor of ductal morphogenesis with sexually dimorphic expression in the urogenital sinus.* Endocrinology, 2010. **151**(7): p. 3420-31.
59. Phillips, T.J., et al., *Localization of genes affecting alcohol drinking in mice.* Alcohol Clin Exp Res, 1994. **18**(4): p. 931-41.
60. Tordoff, M.G., D.R. Reed, and H. Shao, *Calcium taste preferences: genetic analysis and genome screen of C57BL/6J x PWK/PhJ hybrid mice.* Genes Brain Behav, 2008. **7**(6): p. 618-28.
61. Vadasz, C., et al., *Mapping of quantitative trait loci for ethanol preference in quasi-congenic strains.* Alcohol, 2000. **20**(2): p. 161-71.
62. Tarantino, L.M., et al., *Confirmation of quantitative trait loci for alcohol preference in mice.* Alcohol Clin Exp Res, 1998. **22**(5): p. 1099-105.

63. Nelson, G., et al., *Mammalian sweet taste receptors.* Cell, 2001. **106**(3): p. 381-90.
64. Geraedts, M.C., et al., *Transformation of post-ingestive glucose responses after deletion of sweet taste receptor subunits or gastric bypass surgery.* Am J Physiol Endocrinol Metab, 2012.
65. Simon, B.R., et al., *Artificial sweeteners stimulate adipogenesis and suppress lipolysis independent of sweet taste receptors.* Journal of Biological Chemistry, 2013.

Figure 1. Identification of sucrose consumption loci and their interaction. (**A**) A genome-wide scan for intake of 300 mM sucrose by C57BL/6ByJ x 129P3/J.C57BL/6ByJ-*Tas1r3* F_2 mice (N=279). The pink horizontal dashed line represents the significant linkage threshold determined by 1000 permutation tests (threshold=3.64). The black curves trace the logarithm of the odds (LOD) scores (y-axis). The genotype effect of the two significant loci (*Scq2* and *Scq3*) were confirmed in a general linear model using habitual water intake as a covariate and a post hoc comparison of least-square means among the genotypes of these loci tested. The letters 'a' and 'b' above horizontal lines depicting means ± SEMs show the significant differences in daily sucrose consumption by genotype group. (**B**) Heat map of the two-dimensional genome scan for daily sucrose consumption in the F_2 generation using a two-QTL model. The maximum LOD score for the full model (two QTLs plus an interaction) is indicated in the lower right. The maximum LOD score for the interaction model is indicated in the upper left. A yellow circle shows significant interactions between *Scq2* and *Scq1.* The color-coded scales display values of significance for the interaction model (left scale; LOD threshold = 6.5) and the full model (right scale; LOD threshold = 9.1). (**C**) Least square means ± SEs are grouped by genotype to show the interaction of the marker pairs (*Scq1* and *Scq2*: *D1Mit316* and *rs3023231*). The letters 'a', 'b', and 'c' show significant group differences as determined by post hoc tests.

Figure 2. Confirmation of *Scq1*, *Scq2*, and their interaction in backcross, consomic, and congenic mice. (**A**) Backcross mice. *Scq2* was detected in the pooled backcross mice (N_3, N_4, N_5, and N_6; N=438). The x-axis is the location of the markers in Mb (see **S4 Table**) on mouse chromosome 9 (mChr9); the y-axis is the –log p-value of the statistical test by each marker genotype. The small blue bar under the peak indicates

the confidence interval (2 units of –log10 p-value drop). The horizontal red line shows statistical threshold, calculated using Bonferroni correction [–log(0.05/68) ≈ 3.13]. The inset shows the least-square mean (LSM) and standard error of daily sucrose intake of mice grouped by peak marker genotype (*rs3023231*)**.** (**B-E**) Consomic mice. (**B**) A schematic of the comparisons of mouse consomic strains and their host inbred strains, as well as the corresponding genotypes of *Scq1* and *Scq2* in each mouse strain. (**C**) The LSMs and their standard error of the consomic strain 129.B6-Chr1 and the reciprocal consomic strains 129.B6-Chr9 and B6.129-Chr9, compared with their host inbred strain, 129 or B6, separately. (**D**) Effect size in Cohen's *D* for each group computed by comparing homozygous consomic genotypes vs host inbred genotypes. The large difference in effect size between the two reciprocal consomic strains 129.B6-Chr9 and B6.129-Chr9 supports the epistatic interaction of the homozygous alleles for *Scq2* and other chromosomes. (**E**) The interaction effect of homozygous alleles between *Scq1* and *Scq2* was confirmed in the consomic and host inbred strains. (For mice genotypes corresponding to *Scq1* and *Scq2,* see panel B.) (**F-J**) Congenic mice. (**F**) A schematic of the 129.B6-double congenic strain, created by mating *Scq1* homozygous congenic males with *Scq2* heterozygous females. These double congenic mice thus have heterozygous genotypes at the *Scq1* (B6/129 for all mice) locus and segregate genotypes (B6/129 or 129/129) at the *Scq2* locus*.* (**G**) Comparisons of littermates from the mouse congenic strains and corresponding genotypes of *Scq1* and *Scq2* in each group of mice. (**H**) LSM and their standard error of the congenic strains: 129.B6-*Scq1*, 129.B6-*Scq2*, and 129.B6-double congenics, with one copy of B6 donor fragment (B6/129), compared with their littermates without the B6-donor fragment (129/129). (**I**) Effect size in Cohen's *D* computed for each group of comparisons by comparing heterozygous congenic genotype (B6/129) vs their littermates' genotype (129/129). A large difference in effect size between the double congenic strain and 129.B6-*Scq1* or

129.B6-*Scq2* confirmed the epistatic interaction of the heterozygous alleles between *Scq2* and *Scq1*. (**J**) The interaction effect of heterozygous alleles between *Scq1* and *Scq2* was confirmed in congenic strains. (For mice corresponding genotypes of *Scq1* and *Scq2,* see panel G; mice from three congenic strains C2, C3 and C4 that don't retains QTLs, genotypes of both *Scq1* and *Scq2* were recoded as 129/129).

Figure 3. The *Scq1* locus region isolated by comparing the two informative congenic strains. (**A**) For average daily sucrose intake, we computed trait marker associations using a general linear model among all 129.B6-*Scq1* heterozygous congenic mice grouped by genotype at each marker. The x-axis gives marker positions in Mb on chromosome 1 (mChr1); the y-axis, $-\log_{10}$-transformed p-values. (**B**) Donor region of the congenic strain: black bar, strain retained the *Scq1* locus; gray bar, strain lost the locus. Blue indicates the region contributed by the host strain. We determined whether each congenic strain retained the *Scq1* locus by comparing the daily sucrose intakes of littermates with and without the donor fragment. The *Scq1*-positive strain (C1) retains a region (red line; 25.4 Mb from centromere to *rs13475771* at 25.4 Mb) that the *Scq1*-negative strain (C2) does not share. The allele effect direction matches the results from the consomic mice, with the B6 strain allele decreasing the trait. (**C and D**) Average daily sucrose intake of the *Scq1*-positive (C1) and *Scq1*-negative (C2) strains compared using a general linear model**.** Strain comparisons with sample size (n) of each genotype within each strain.

Figure 4. The *Scq2* locus region isolated by comparing seven congenic strains. (**A**) For average daily sucrose intake, we computed trait-marker associations using a general linear model among all 129.B6-*Scq2* congenic mice grouped by genotype at each marker. The x-axis gives marker positions in Mb on chromosome 9 (mChr9); the y-axis, $-\log_{10}$-transformed p-values. (**B**) Donor region of each congenic strain: black bar, strains

retained the *Scq2* locus; gray bar, strains lost the locus; blue indicates the region contributed by the host strain. We determined whether congenic strains (shown at left) retained the *Scq2* locus (i.e., were "positive") by comparing the daily sucrose intakes of littermates with and without the donor fragment. (**C**) Strain comparisons with sample size (n) of each genotype within each strain. C3 and C4, *Scq2* negative; C5–C9, *Scq2* positive. (**D**) *Scq2*-positive strains (C5-C9) share a common region (red dashed line; 1.2-Mb donor region between 105.7 and 106.9 Mb, from *rs33653996* to *rs3023231*) that the *Scq2*-negative strains (C3 and C4) do not share. The allele effect direction matches that from the F_2, backcross, and consomic mice, with the B6 strain allele increasing the trait. We show noncoding RNA genes, protein-coding genes, pseudogenes, and processed transcripts within the 1.2 Mb *Scq2* region, which we obtained from the Ensemble Mouse Genome Brower @GRCm38.p5.

Figure 5. The *Scq1* homozygous allele effect on voluntary sucrose intake. The *Scq1* locus affects 300 mM sucrose consumption in homozygous 129.B6-*Scq1* congenic mice (B6/B6; C1) vs host inbred 129 mice (129/129), with an effect size.

Figure 6. The effects of S*cq1* are largest for the most concentrated sugar solutions. Average daily intake (left panels) and preference (right panels) for sucrose (**A**), glucose (**B**), and fructose (**C**) by homozygous 129.B6-*Scq1* congenic mice (C1; for sucrose, N=19; for glucose, N=10; for fructose, N=10) and host inbred 129 mice (all N=10) in two-bottle 48-h preference tests. Error bars show SEM. *p<0.05, significant difference between the congenic and host inbred mice (post hoc tests). The mice corresponding genotypes for *Scq1* and *Scq2* for congenic mice are denoted by blue circles, and the control 129 inbred mice by black circles.

Figure 7. The *Scq1* genotype affects fuel oxidation but not food intake. Metabolic assessments in *Scq1* congenic mice (black bars; N=17; C1) and control 129 inbred mice

(gray bars; N=8) at 8 and 36 weeks of age: food intake (**A**), water intake (**B**), heat production (**C**), respiratory exchange ratio (RER; **D**), oxygen consumption (VO_2; **E**), carbon dioxide production (VCO_2; **F**), and activity over the entire 24-hr light/dark cycle (**G**). Data are mean ± SEM. The mice corresponding genotypes for *Scq1* and *Scq2* for congenic mice are in black bars and the control 129 inbred mice are in grey bars.

Supplemental Figures

S1 Figure. Correlations between sucrose intake (ml/d) and body weight (g), daily habitual water intake (ml/d), and age (weeks) of mice within each genetic mapping population. We report sample size in **S1 Table**. BW_M=body weight in males, BW_F=body weight in females, Wat=daily habitual water intake.

S2 Figure. F_2 mice (N=279) showed few or no genotype effects of the *Scq1, Scq2*, and *Scq3* loci on intake or preference for most taste solutions (N=279), including a low concentration (50 mM) of sucrose. Group means that do not differ either (1) have no letter subscript or (2) share a letter subscript in common. Genotype annotation: A=B6/B6, H=129/B6, B=129/129.

S3 Figure. Heat map of Pearson correlations for taste solution intakes and preference scores of the F_2 mice (N=279).

S4 Figure. Chromosomes 1 and 9 harbor QTLs for the consumption of other taste solutions. A significant difference in intake and/or preference of a solution between a consomic strain and host inbred 129 or B6 strain indicates presence of a QTL on this chromosome for the taste solution ($p<0.05$). For intakes, we used a general linear model, with strain as fixed factor and habitual water intake as a covariate, followed by post hoc test. For preference scores, we used t-tests.

S5 Figure. Effects of the *Scq1* locus on intakes and preference for 20 mM saccharin (**A**), 30 mM glycine (**B**), 300 mM NaCl (**C**), 300 mM MSG (**D**), 3% ethanol (**E**), and 10% ethanol (**F**) by homozygous 129.B6-Scq1 congenics (C1: B6/B6) and host inbred 129 mice (129/129) in two-bottle preference tests.

S6 Figure. No effect of the *Scq1* locus on average daily intake and preference for 4.6% soybean oil (**A**), 6% corn starch (**B**), or 10.5% maltodextrin (**C**) by homozygous 129.B6-*Scq1* congenics (C1: B6/B6; N=19) and host inbred 129 mice (129/129; N=10) in two-bottle preference tests.

S7 Figure. No *Scq1* genotype response to tolerance of glucose administered intraperitoneally (IP GTT; **A**) or intragastrically (IG GTT; **B**) or insulin (**C**) in homozygous *Scq1* congenic mice (black squares; N=21) and the host inbred control mice (gray circles; N=12) at age 8 weeks and 36 weeks. Data are mean ± SEM of genotype, and the significance of the genotype effect was evaluated by post hoc tests.

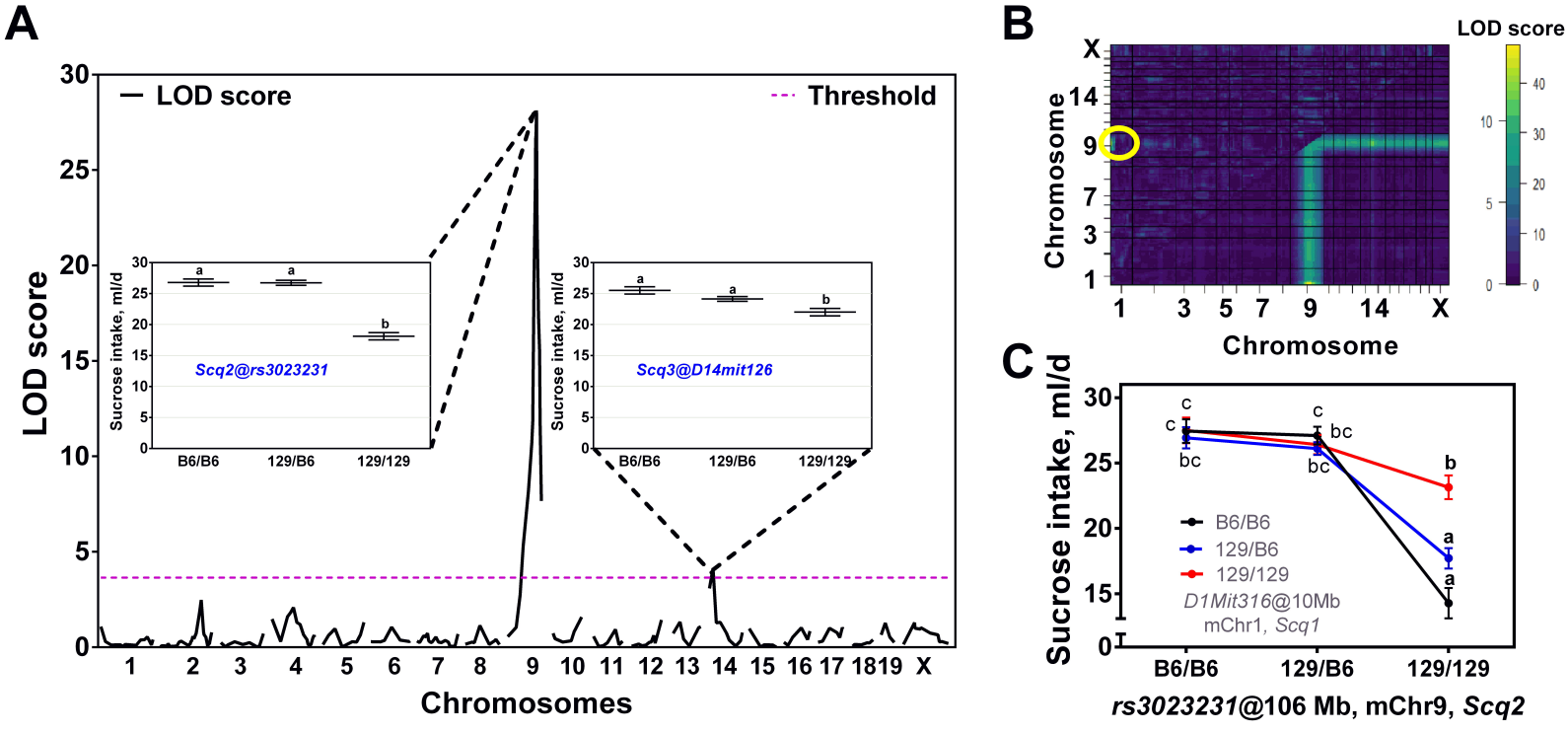
S8 Figure. No effect of the *Scq1* genotype on mouse body composition at 8 or 36 weeks of age. (**A**) Mouse body composition as measured by MR (8 weeks: N=32, males=15, females=17; 36 weeks: N=22, males=9, females=13). (**B**) Mouse body composition as measured by DEXA (N=26, males=12, females=14). BMC=bone mineral content (g); BMD=bone mineral density (g/cm^2). Homozygous *Scq1* congenics (C1; B6/B6) vs host inbred 129 controls (129/129); Groups that do not share a common subscript letter differ by post hoc testing.

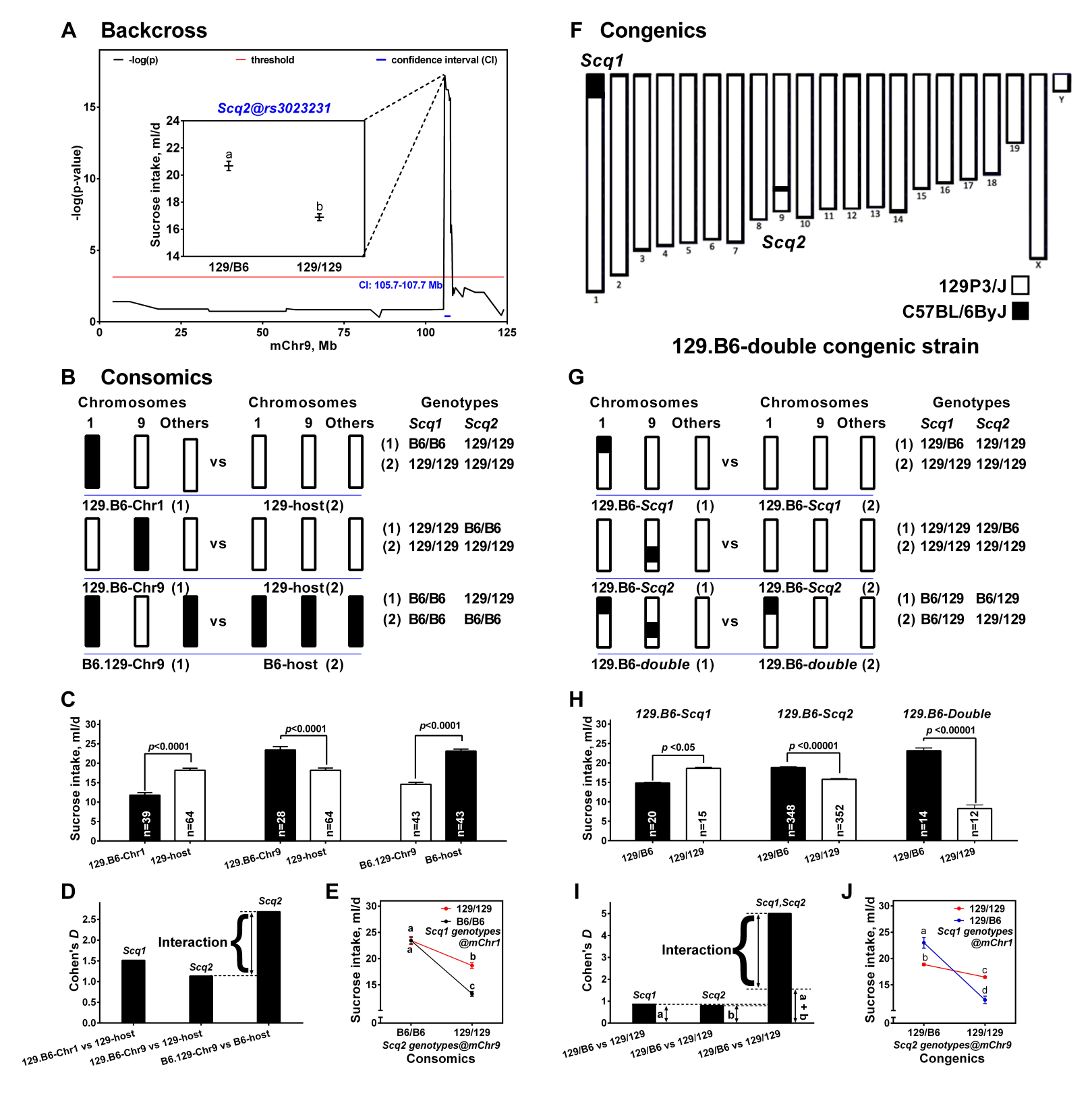
S9 Figure. No genotype effect of *Scq1* on organ and fat weight of mice dissected at 180 days of age. Homozygous *Scq1* congenics (C1; B6/B6; males=7, females=12) vs host inbred 129 controls (129/129; males=5, females=2). Groups that do not share a common subscript letter differ by post hoc testing.

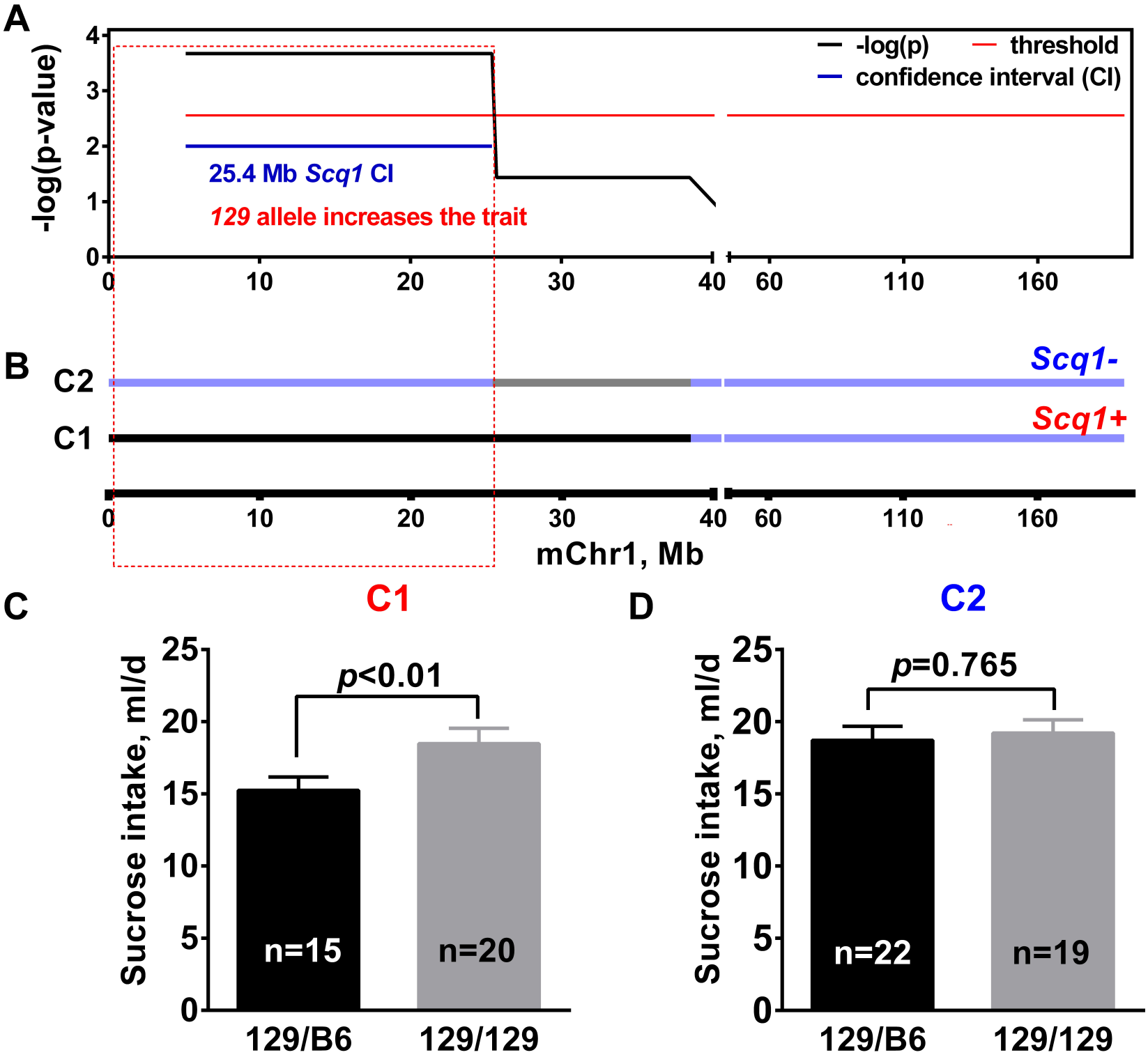
S10 Figure. No genotype effect on body weight changes when mice were fed chow and drank sucrose (**A-H**), glucose (**I**), or fructose (**J**) solution voluntarily for 4-12 days in F_2 mice (**A-C**; N=279), consomic mice (**E-G**; 129.B6-Chr1 and 129 controls: N=105; 129.B6-Chr9 and 129 controls: N=87; B6.129-Chr9 and B6 controls: N=86), or congenic mice (**H-J**; sucrose: N=105; glucose: N=20; fructose: N=20). See **S9 Figure** caption for other details. A=B6/B6 genotype; H=129/B6 genotype; B=129/129 genotype.

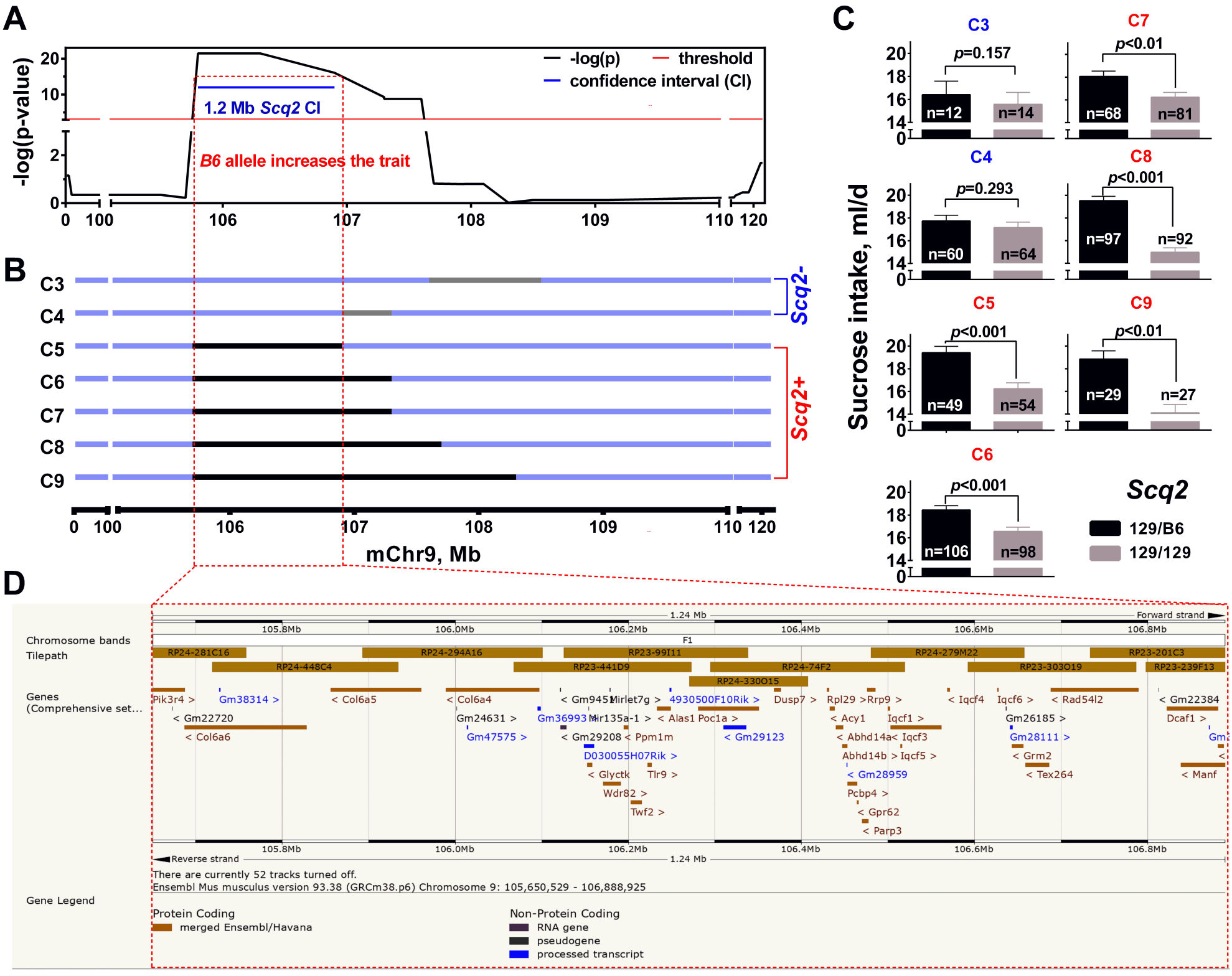
S11 Figure. Both *Scq1* and *Scq2* showed no genotype effect on the responses of the chorda tympani gustatory nerve to lingual application of taste stimuli. We compared the heterozygous 129.B6-*Scq1* congenic mice (C1; B6/129; N=5) or homozygous 129.B6-*Scq2* congenic mice (C7: B6/B6; N=5) and host inbred 129 mice (129/129; N=5) using Mann-Whitney *U*-tests for individual solutions. We observed no significant genotype differences (p>0.05) in responses of the chorda tympani gustatory nerve for the taste stimuli, except for 1000 mM sucrose for *Scq1*, with * p=0.03 (**S8 Table**). Values are median ± median absolute deviations.

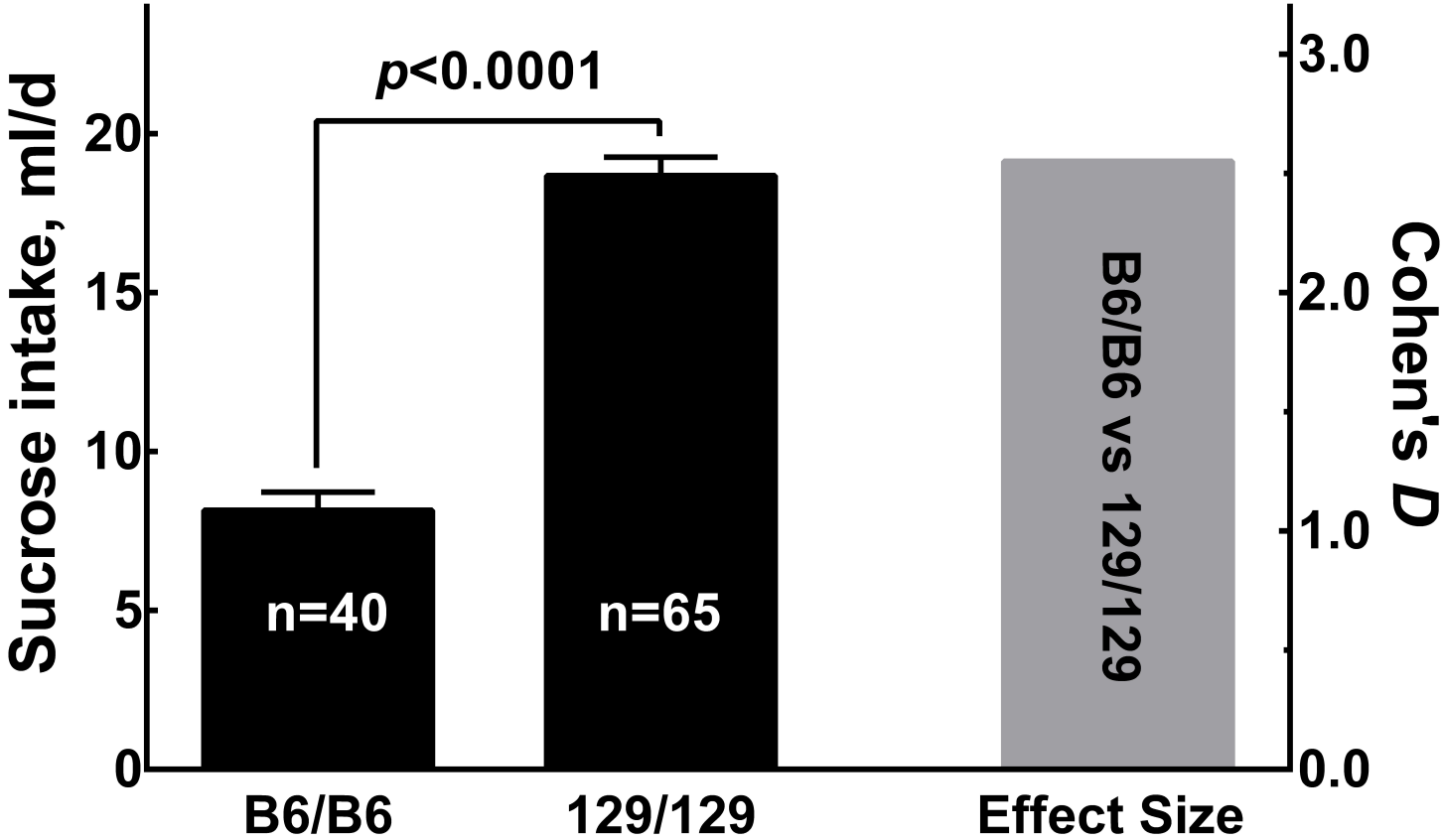
S12 Figure. (**A**) Gene ontology analysis of the biological process for genes within the *Scq1* and *Scq2* regions. The Ensembl mouse gene identifier IDs (N=195; no Ensembl gene identifier IDs were available for the 45 unknown genes) were batch-uploaded into the Panther database for *Mus musculus*, and in total, 92 genes and 125 hits were processed. (**B**) Protein-protein associations for genes within *Scq1* and *Scq2* (**S10 Table**) were identified by searching protein names using information from high-throughput experimental data, mining of literature databases, and predictions based on genomic context analysis.

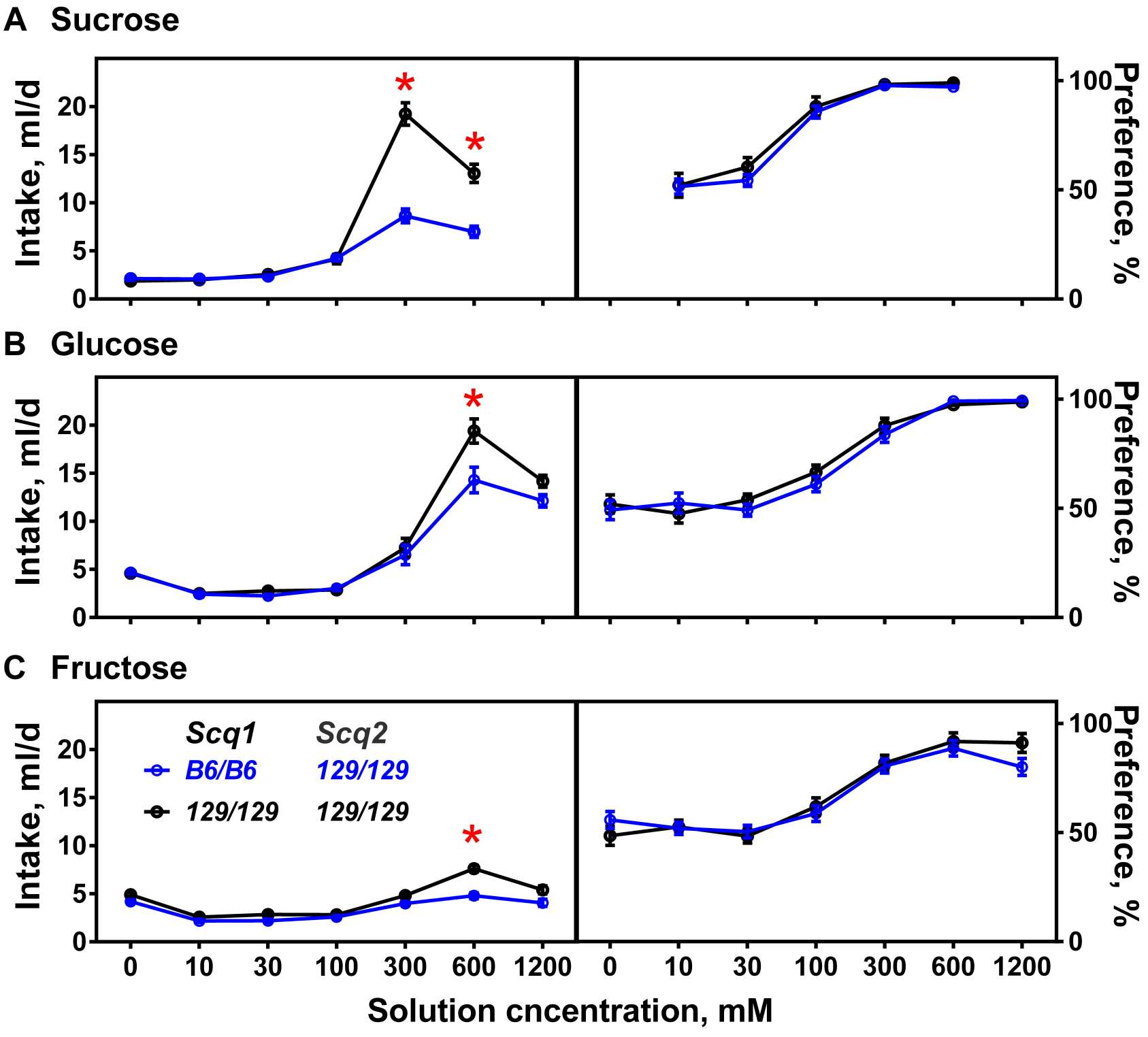

A
LOD score
Threshold
LOD score
Chromosomes
Sucrose intake, ml/d
Scq2@rs3023231
B6/B6
129/B6
129/129
Scq3@D14mit126
B
LOD score
Chromosome
Chromosome
C
Sucrose intake, ml/d
B6/B6
129/B6
129/129
D1Mit316@10Mb
mChr1, Scq1
rs3023231@106 Mb, mChr9, Scq2

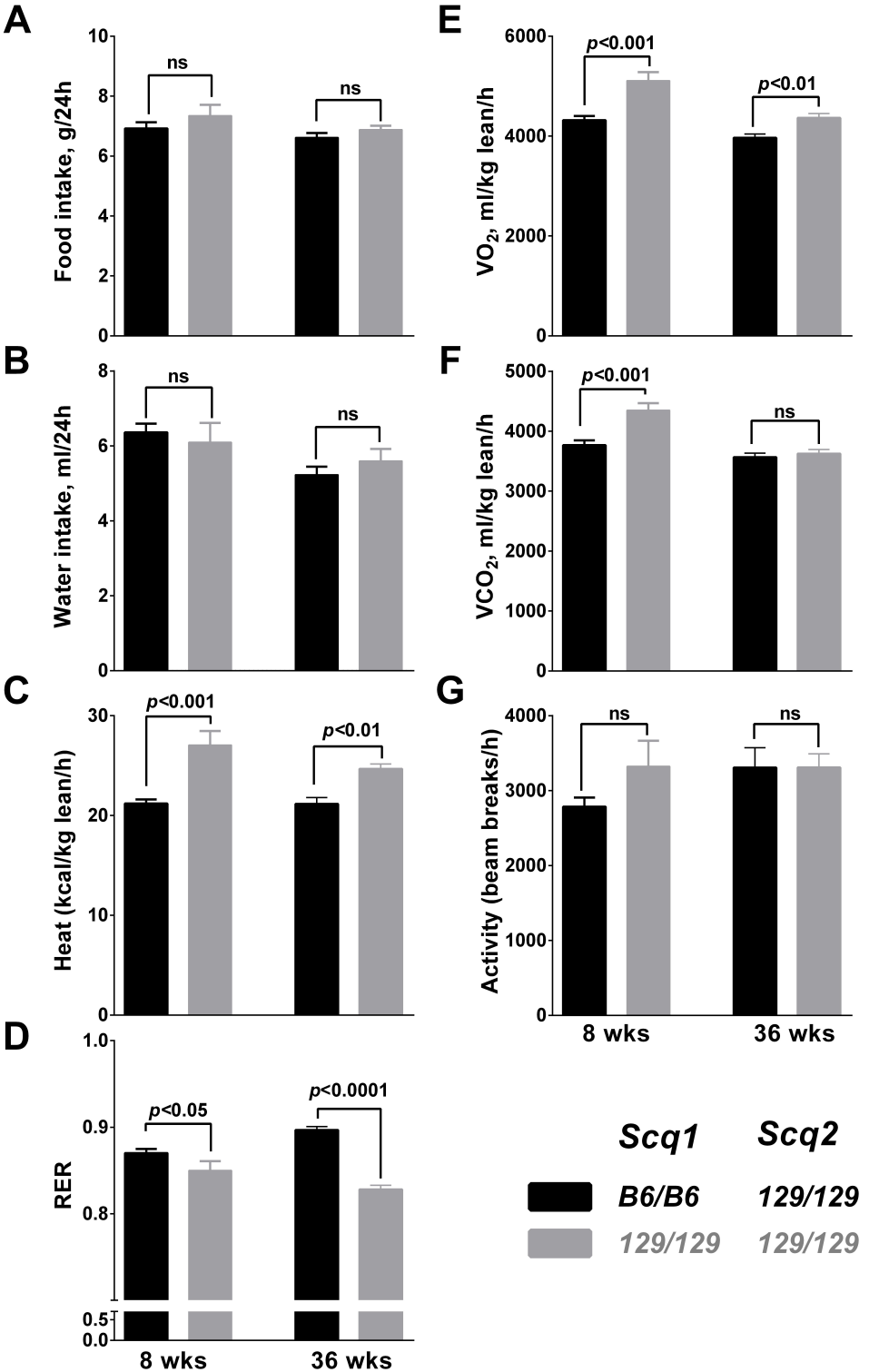

A Backcross
-log(p)
threshold
confidence interval (CI)
Scq2@rs3023231
Sucrose intake, ml/d
129/B6
129/129
CI: 105.7-107.7 Mb
-log(p-value)
mChr9, Mb
B Consomics
Chromosomes
Genotypes
Scq1
Scq2
129.B6-Chr1 (1)
129-host(2)
129.B6-Chr9 (1)
B6.129-Chr9 (1)
B6-host (2)
C
p<0.0001
129.B6-Chr1
129-host
129.B6-Chr9
B6.129-Chr9
B6-host
n=39
n=64
n=28
n=43
D
Cohen's D
Interaction
129.B6-Chr1 vs 129-host
129.B6-Chr9 vs 129-host
B6.129-Chr9 vs B6-host
E
129/129
B6/B6
Scq1 genotypes @mChr1
Scq2 genotypes@mChr9
Consomics
F Congenics
Scq1
Scq2
129P3/J
C57BL/6ByJ
129.B6-double congenic strain
G
129.B6-Scq1
129.B6-Scq2
129.B6-double
H
p <0.05
p <0.00001
n=20
n=15
n=348
n=352
n=14
n=12
I
Scq1,Scq2
129/B6 vs 129/129
J
Congenics


A
-log(p-value)
-log(p)
threshold
confidence interval (CI)
25.4 Mb Scq1 CI
129 allele increases the trait
B
C2
C1
Scq1-
Scq1+
mChr1, Mb
C
C1
Sucrose intake, ml/d
p<0.01
n=15
n=20
129/B6
129/129
D
C2
Sucrose intake, ml/d
p=0.765
n=22
n=19
129/B6
129/129


A
-log(p-value)
-log(p)
threshold
confidence interval (CI)
1.2 Mb Scq2 CI
B6 allele increases the trait
B
C3
C4
C5
C6
C7
C8
C9
Scq2-
Scq2+
mChr9, Mb
C
Sucrose intake, ml/d
C3
p=0.157
n=12
n=14
C7
p<0.01
n=68
n=81
C4
p=0.293
n=60
n=64
C8
p<0.001
n=97
n=92
C5
p<0.001
n=49
n=54
C9
p<0.01
n=29
n=27
C6
p<0.001
n=106
n=98
Scq2
129/B6
129/129
D
Chromosome bands
Tilepath
Genes (Comprehensive set...
Forward strand
Reverse strand
There are currently 52 tracks turned off.
Ensembl Mus musculus version 93.38 (GRCm38.p6) Chromosome 9: 105,650,529 - 106,888,925
Gene Legend
Protein Coding
merged Ensembl/Havana
Non-Protein Coding
RNA gene
pseudogene
processed transcript


p<0.0001
Sucrose intake, ml/d
20
15
10
5
0
n=40
n=65
B6/B6 vs 129/129
3.0
2.0
1.0
0.0
Cohen's D
B6/B6
129/129
Effect Size


A Sucrose
B Glucose
C Fructose
Intake, ml/d
Preference, %
Scq1
Scq2
B6/B6
129/129
129/129
129/129
0
5
10
15
20
0
50
100
0
10
30
100
300
600
1200
Solution cncentration, mM


A
Food intake, g/24h
ns
ns
B
Water intake, ml/24h
ns
ns
C
Heat (kcal/kg lean/h)
p<0.001
p<0.01
D
RER
p<0.05
p<0.0001
8 wks
36 wks
E
VO2, ml/kg lean/h
p<0.001
p<0.01
F
VCO2, ml/kg lean/h
p<0.001
ns
G
Activity (beam breaks/h)
ns
ns
8 wks
36 wks
Scq1
Scq2
B6/B6
129/129
129/129
129/129