Report

Neuron

Genetic Background Limits Generalizability of Genotype-Phenotype Relationships

Highlights

- *Cacna1c* and *Tcf7l2* null allele effects were strongly influenced by genetic background
- Some interactions with genetic background supported opposite conclusions
- Genetic background modulated mutant phenotypes more strongly than sex
- These results do not negate the contributions of mutant mice to biomedical research

Authors

Laura J. Sittig, Peter Carbonetto, Kyle A. Engel, Kathleen S. Krauss, Camila M. Barrios-Camacho, Abraham A. Palmer

Correspondence

aap@ucsd.edu

In Brief

Sittig et al. demonstrate low generalizability of mouse null allele phenotypes across a panel of F_1 genetic backgrounds, which suggests that the use of single strains is a barrier to robust characterization of genotype-phenotype relationships.

Sittig et al., 2016, Neuron *91*, 1253–1259
September 21, 2016
http://dx.doi.org/10.1016/j.neuron.2016.08.013

CellPress

Neuron
Report

Genetic Background Limits Generalizability of Genotype-Phenotype Relationships

Laura J. Sittig,[1,2] Peter Carbonetto,[1] Kyle A. Engel,[1] Kathleen S. Krauss,[1] Camila M. Barrios-Camacho,[1] and Abraham A. Palmer[1,2,3,*]
[1]Department of Human Genetics, University of Chicago, 920 East 58th Street, Chicago, IL 60637, USA
[2]Department of Psychiatry, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA
[3]Lead Contact
*Correspondence: aap@ucsd.edu
http://dx.doi.org/10.1016/j.neuron.2016.08.013

SUMMARY

Genome-wide association studies (GWASs) have identified numerous loci that influence risk for psychiatric diseases. Genetically engineered mice are often used to characterize genes implicated by GWASs. These studies are based on the assumption that observed genotype-phenotype relationships will generalize to humans, implying that the results would at least generalize to other inbred mouse strains. Given current concerns about reproducibility, we sought to directly test this assumption. We produced F_1 crosses between male C57BL/6J mice heterozygous for null alleles of *Cacna1c* and *Tcf7l2* and wild-type females from 30 inbred laboratory strains. We found extremely strong interactions with genetic background that sometimes supported diametrically opposing conclusions. These results do not negate the invaluable contributions of mouse genetics to biomedical science, but they do show that genotype-phenotype relationships cannot be reliably inferred by studying a single genetic background, and thus constitute a major challenge to the status quo.

INTRODUCTION

The mouse is the premiere mammalian model organism. Since the development of transgenic and knockout mice, thousands of publications have used mice to define relationships between genotypes and phenotypes (Brandon et al., 1995; Capecchi, 1989). The advent of nuclease-mediated genome editing strategies including CRISPR/Cas9 is expected to accelerate the use of mice for this purpose (Singh et al., 2015). Mutant alleles are typically studied in a single inbred strain background, reflecting the widely held reductionist worldview that seeks to examine a single genetic difference while holding all other genetic and environmental variables constant (Little and Colegrave, 2016). Moreover, most studies focus on relatively young animals, and until recently, most mutant alleles were studied in only one sex, typically males. As human genome-wide association studies (GWASs) continue to implicate new loci in a wide range of common diseases, mutant mice are being used to examine which genes within these loci may influence disease liability and to examine the underlying molecular mechanisms. At the same time, the International Knockout Mouse Consortium (IKMC) is generating and phenotyping a knockout mouse line for every one of the >20,000 protein coding genes in the mouse genome, mostly on a single genetic background (Bradley et al., 2012; Collins et al., 2007). Thus, a huge literature continues to accumulate that describes the phenotypes caused by mutant alleles in single strains.

Recent discussions have focused on failures to replicate results in biomedical science. Replicability is a paramount consideration when conducting phenotypic screening among mutant mice in single laboratories and is also a major issue for high-throughput efforts such as the IKMC. Standard operating procedures for phenotyping and robust methodologies for data analysis have been developed in order to maximize power and reproducibility (Hrabě de Angelis et al., 2015; Karp et al., 2015; Kilkenny et al., 2010; Mallon et al., 2008). In addition, the NIH has implemented a major policy change that requires inclusion of both sexes as a biological variable in preclinical research studies (Clayton and Collins, 2014). Yet there is no standard in the field for establishing generalizability across inbred strains, despite an awareness among mouse geneticists that genetic background is another potential source of variability (Doetschman, 2009; Phillips et al., 1999; Sanford et al., 2001; Threadgill et al., 1995). Examples of epistatic interactions among naturally occurring alleles were reported a century ago (Castle, 1919; Dexter, 1914).

We examined the phenotypic effects of two null alleles using F_1 crosses between C57BL/6J and a panel of 30 different inbred strains. The two genes that we chose to study have been strongly implicated in multiple disorders by human GWASs. The first gene, *CACNA1C*, has been associated with bipolar disorder (Ferreira et al., 2008; Green et al., 2013; Sklar et al., 2008) and schizophrenia (Hamshere et al., 2013; Nyegaard et al., 2010; Ripke et al., 2011), as well as with cross-disorder risk for these and several other major psychiatric disorders (Smoller et al., 2013). Mice with a null allele for *Cacna1c* have been reported to display numerous behavioral phenotypes (Dao et al., 2010), which lends support to the human studies implicating *CACNA1C* in psychiatric diseases. The second gene we studied, *TCF7L2*, is among the most strongly associated and best replicated genetic

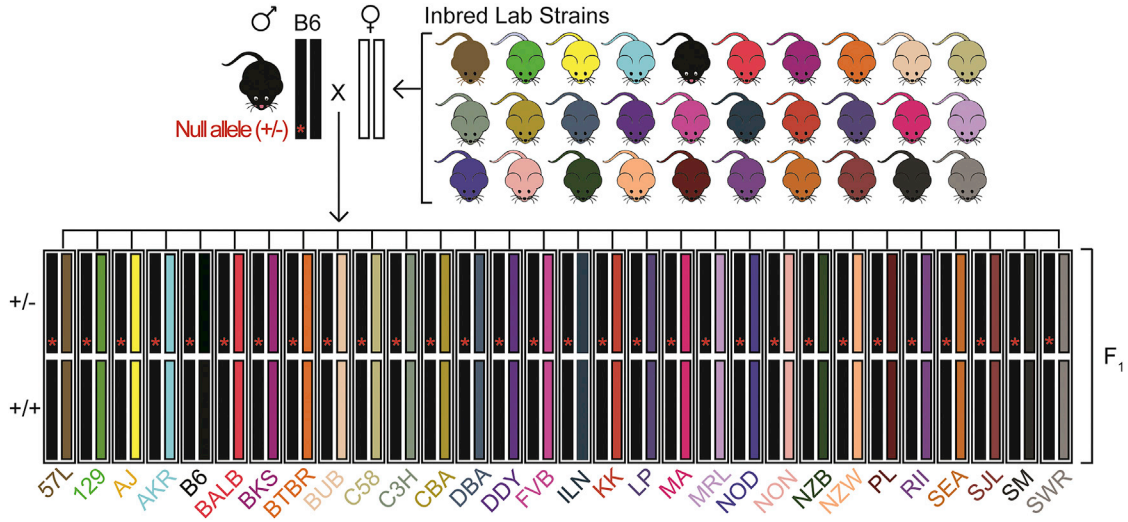

Figure 1. Breeding Design for Placing Null Alleles on Different Genetic Backgrounds

Males heterozygous for the null allele mutation (red asterisk) on a C57BL/6J background were bred with females from 30 inbred laboratory mouse strains (represented by schematic chromosomes of arbitrarily selected colors). This generated 30 distinct F_1s that were genetically identical except at the targeted allele, where they were either +/− or +/+.

risk factors for type 2 diabetes (Fuchsberger et al., 2016; Grant et al., 2006; Morris et al., 2012). In addition to its metabolic role, *TCF7L2* has also been associated with schizophrenia and bipolar disorder (Alkelai et al., 2012; Hansen et al., 2011; Winham et al., 2014). Mice with a null allele for *Tcf7l2* exhibit both physiological and behavioral phenotypes including improved glucose tolerance, altered fear learning, and anxiety (Savic et al., 2011a, 2011b). Here we evaluated the generalizability of *Cacna1c*$^{+/-}$ and *Tcf7l2*$^{+/-}$ phenotypes using many different genetic backgrounds. Our results illustrate how the reproducibility and robustness suffer when only a single strain is considered.

RESULTS

We generated a structured panel of heterozygous null and wild-type mice from 30 different F_1 backgrounds by breeding +/− C57BL/6J males to +/+ females from 30 inbred strains (Figure 1). This produced a panel of +/+ and +/− littermates that were isogenic at all other loci, analogous to the approaches taken by Dorman et al. (2016), Lifsted et al. (1998), and Park et al. (2014). We used this breeding design to generate two cohorts of mice, one for the *Cacna1c* null allele and one for the *Tcf7l2* null allele. Each cohort of mice was produced independently. Mice in the *Cacna1c* cohort (n = 723) were tested for exploratory activity, anxiety, methamphetamine sensitivity, depression-like behavior, and acoustic startle response. Mice in the *Tcf7l2* cohort (n = 630) were tested for exploratory activity, anxiety, fear conditioning, and sensorimotor gating, as well as several metabolic traits: body weight, fasted blood glucose levels, and baseline blood glucose levels. Altogether, we obtained data for 15 phenotypes, 12 of which were behavioral (Tables 1 and S3, available online).

We estimated the variance in each phenotype explained by the factors of interest in our experimental design: genotype (+/+ or +/−), F_1 genetic background ("strain"), sex, and the two- and three-way interactions among these factors. The interaction between genotype and strain indicates the degree to which the null allele's effect depended on strain and is therefore a measure of generalizability of the null allele's phenotypic effects. Strikingly, the majority of phenotypes showed significant or trending genotype × strain interactions, indicating that they depended on the strain in which the null allele was expressed (Table 1; Figures 2A–2D; Table S3). The interactions affected phenotypes with highly penetrant, weak, and non-existent main effects of the null allele (Figure S1).

Phenotypic data from different F_1s frequently supported dramatically different conclusions. For example, *Cacna1c* genotype affected methamphetamine sensitivity in approximately half of the F_1s, while the remaining F_1 +/− mice were not significantly different from +/+ mice (Figure 2A). This spectrum of vulnerability was also observed for the *Tcf7l2* null allele (e.g., Figures 2B and 2C). However, the impact of genetic background was not always merely a matter of degree; there were four instances in which directionally opposite effects of the same allele occurred in different genetic backgrounds (Tables 1 and S3). For example, a genotype × strain interaction precluded a main effect of the *Tcf7l2* null allele on acoustic startle response (Figure 2D). Post hoc tests demonstrated that for three of the F_1s (ILN, BKS, and DBA) *Tcf7l2* haploinsufficiency decreased acoustic startle response, whereas the opposite was true for two other F_1s (AJ and NZB).

Finally, we considered the role of sex as a biological variable. Three of the traits showed modest but significant interactions between genotype and sex (startle, body weight, and prepulse inhibition). For body weight, there was a significant three-way interaction between genotype, strain, and sex (Table S3).

Table 1. Genotype × Strain Interactions

Phenotype	Genetic Background × Null Allele p Value	Phenotypic Variance Explained by Background × Null Allele
Cacna1c Cohort		
Percent time in the light box	0.20[d]	4.3%
Exploratory activity	0.01[a]	3.4%
Duration in center	9.9×10^{-8}[c]	7.4%
Methamphetamine sensitivity	9.5×10^{-3}[b]	4.2%
Immobility	7.4×10^{-3}[b]	4.4%
Acoustic startle response	0.08[d]	1.6%
Tcf7l2 Cohort		
Body weight	0.42	0.6%
Exploratory activity	6.8×10^{-6}[c]	7.0%
Duration in center	6.8×10^{-7}[c]	5.9%
Contextual fear learning	0.23	3.5%
Cued fear learning	0.08[d]	4.2%
Acoustic startle response	8.9×10^{-8}[c,d]	3.9%
Prepulse inhibition	8.7×10^{-4}[c]	7.6%
Fasted blood glucose levels	0.24	3.4%
Baseline blood glucose levels	0.05	4.3%

See Table S3 for the ANOVA of each phenotype. See Figure S1 for a comparison of main and interactive effects.
[a]$p < 0.05$.
[b]$p < 0.01$.
[c]$p < 0.001$.
[d]Opposite effects of the null allele occurred in different F_1 backgrounds.

DISCUSSION

Single strains are often used to establish the presence, absence, and severity of phenotypes in genetically engineered mice. The conclusions from such experiments inform our understanding of genetic and physiological systems, yet these conclusions are predicated on the assumption that they generalize to humans, which presupposes that they would at least generalize to other inbred mouse strains. Here we evaluated the phenotypic effects of two null alleles on a panel of F_1 mice derived from commonly used inbred mouse strains. Our breeding design produced heterozygous null littermates that were genetically identical to their respective wild-type controls at all loci besides the targeted allele. The majority of null phenotypes observed in *Cacna1c*$^{+/-}$ and *Tcf7l2*$^{+/-}$ mice were not generalizable: phenotypic responses often varied from strongly affected to unaffected in different genetic backgrounds, and in several cases there were directionally opposite effects of the same allele. Overall, the prevalence and strength of interactions with genetic background were greater than the interactions with sex. This study illustrates that the choice of genetic background can have a dramatic effect on the null allele phenotype, challenging the reductionist idea that mutant alleles have a specific phenotype that can be readily determined using a single strain.

The null alleles we evaluated had been phenotyped previously using single genetic backgrounds, allowing us to compare our results to those previous studies. The importance of procedural and environmental differences complicates direct comparisons of our phenotypic data to previous results (Crabbe et al., 1999; Wahlsten et al., 2003; Chesler et al., 2002; Sorge et al., 2014). However, in the case of *Cacna1c*, the one assay that was performed and analyzed with very similar procedures replicated a sex-specific decrease of the startle response in C57BL/6J mice (Dao et al., 2010) (Table S3). For *Tcf7l2*, the null allele was previously evaluated in outbred CD-1 mice in our laboratory (Savic et al., 2011a). We identified five phenotypes from that study that could be compared to our current results, three of which exhibited strong interactions with genetic background that precluded cross-study comparisons of single phenotypic effects. The other two were different from what was reported in CD-1 mice, and for one of them, contextual fear learning, the direction of the effect was directionally opposite to that previously reported. These results illustrate how genetic background can dramatically alter experimental conclusions.

Our statistical power to detect effects of the null alleles was limited by sample size in this study design. Certain F_1s that appeared sensitive or insensitive to one of the mutant alleles may represent type 1 or 2 errors; accordingly, the number of significant post hoc tests was reduced when we performed permutations of the phenotypic data (data not shown). However, correction for multiple testing would not have been appropriate because our goal was to compare the experimental conclusions that would have been reached if only a single genetic background was considered. Our study design could have been extended to include reciprocal crosses or even a full diallele cross; however, our design was sufficient to demonstrate the importance of genetic context. Our use of F_1 mice most likely reduced interactions when compared to what would be seen across fully inbred strains. Had we introduced the same mutation using nuclease-mediated genome editing onto multiple pure inbred lines rather than F_1s, we might have seen stronger interactions with genetic background, but these could have been due to off-target mutations that would have differed from strain to strain.

This study was not designed to address the prevalence of epistatic interactions involving *CACNA1C* and *TCF7L2* in humans; indeed, we examined null alleles, which are presumably more severe than the corresponding human risk alleles. Furthermore, it is possible that the unique population history of inbred mouse strains contributes to the strength of the observed interactions; for example, by amplifying the frequency of alleles that are rare in wild mouse populations. It is not clear whether interactions with genetic background would have been as prevalent had we used wild rather than laboratory mice. It has been argued that evaluation of mutant phenotypes in single strains of F_1 hybrid mice, such as those we used here, is a preferable strategy to using single inbred strains because F_1s capture greater genetic diversity and are less likely to show anomalous phenotypic effects (Silva et al., 1997). However, our data suggest that using F_1

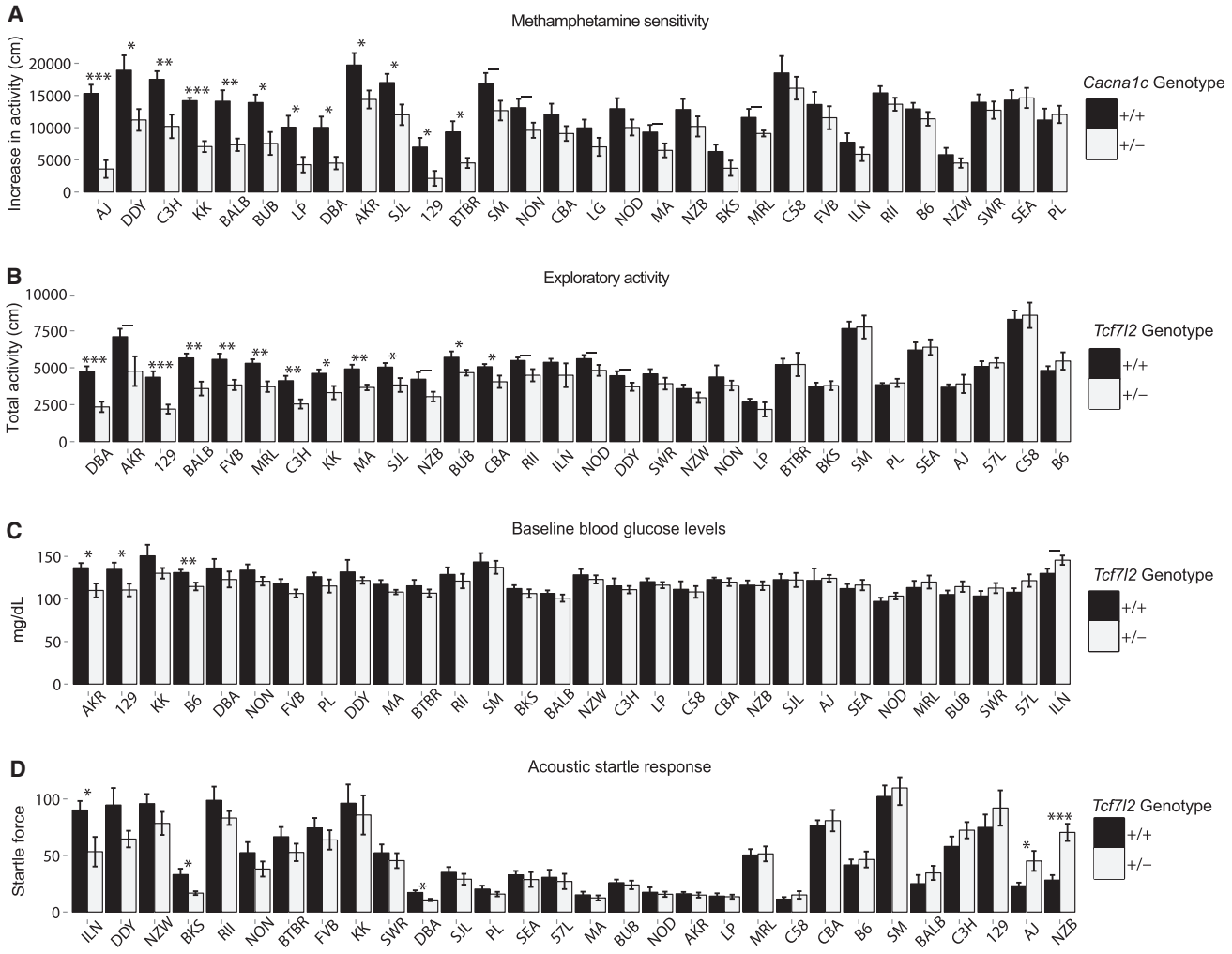

Figure 2. Null Allele Phenotypes on 30 Different Genetic Backgrounds

(A) Locomotor response to methamphetamine in *Cacna1c*$^{+/-}$ and *Cacna1c*$^{+/+}$ mice. On days 1 and 2, a saline injection was given before a 30 min exposure to an open field. On day 3, methamphetamine (2 mg/kg i.p. [intraperitoneally]) was given before the open field. The difference in locomotor activity on day 3 compared to activity on day 2 is shown.

(B) Exploratory activity during a 30 min open field test in *Tcf7l2*$^{+/-}$ and *Tcf7l2*$^{+/+}$ mice.

(C) Baseline blood glucose levels in *Tcf7l2*$^{+/-}$ and *Tcf7l2*$^{+/+}$ mice.

(D) Startle response to a 120-dB acoustic stimulus in *Tcf7l2*$^{+/-}$ and *Tcf7l2*$^{+/+}$ mice.

Bars show the mean phenotype ± SEM. F_1s are labeled according to the maternal strain and are ordered along the x axis according to the difference between +/− and +/+ mice. See also Table S3. Null allele effects within strains were evaluated using two-tailed t tests. $^{-}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

animals does not circumvent the confounding effects of genetic background.

What can be done to remedy this problem? The strong effects of genetic background represent both a blessing and a curse. While they complicate the evaluation of mutant alleles, they also create new experimental opportunities. It is now feasible to create mutations in several strains using nuclease-mediated genome editing or, when mutant phenotypes of interest are dominant, by using an F_1 breeding scheme similar to the one presented here. Differences in susceptibility across mouse strains allow the identification of gene-gene interactions (Nadeau, 2001), an approach that has been used successfully in the past (Dietrich et al., 1993; Hamilton and Yu, 2012; Heydemann et al., 2009; Hide et al., 2002; Pinto et al., 2013; Rozmahel et al., 1996). There are major opportunities to leverage the genetic diversity among inbred mouse strains to reveal functional biological networks that underlie disease processes, especially as the engineering of mutant alleles becomes increasingly efficient. Doing so will be key to unraveling the genetic basis for disease-relevant traits as well as developing new therapeutic avenues to intervene in the associated pathophysiology. We must broaden our focus beyond single strains to realize this potential.

EXPERIMENTAL PROCEDURES

Null Alleles

Cacna1c$^{+/-}$ mice were originally developed by Deltagen and were obtained from the Jackson Laboratory (strain 005783). The line was backcrossed to C57BL/6J for at least five generations prior to arrival at our facility. We backcrossed the line for three additional generations. We examined residual heterozygosity in the backcrossed *Cacna1c*$^{+/-}$ animals using the Mega Mouse Universal Genotyping Array (MegaMUGA) (http://csbio.unc.edu/CCstatus/index.py). The founders used to breed the F_1 panel were 99.8% identical to C57BL/6J based on 73,178 informative SNPs.

Tcf7l2$^{+/-}$ mice on a C57BL/6J background were generated using a zinc-finger nuclease construct (Savic et al., 2011b). We obtained a mutant line with a 10 bp frameshift-inducing deletion. The founder male was backcrossed to C57BL/6J for one generation before generating F_1 crosses.

It is possible that off-target mutations in the *Tcf7l2* line or residual heterozygosity in the *Cacna1c* line may have had phenotypic effects that were falsely attributed as phenotypes of the null allele. However, our use of littermate controls ensures that these could not have been the source of the interactions between the null allele and genetic background. Additional details about the mouse lines and genotyping the null alleles are found in the Supplemental Experimental Procedures.

Generation of F_1 Offspring and Phenotyping

All animal procedures were approved by the University of Chicago Institutional Animal Care and Use Committee. Mice were housed in a single pathogen-free barrier facility. Lights were on a 12 hr on/12 hr off cycle with lights on at 0600 hr. Mice were housed in standard polycarbonate cages with corn cob bedding and ad libitum access to water and laboratory chow. Water (filtered by reverse osmosis) was available in each cage. Breeders received Envigo (formerly Harlan) 2919 19% protein chow. F_1 offspring received Envigo 2918 18% protein chow after weaning. The length of time required to produce and phenotype all F_1 mice was 8 months for the *Tcf7l2* cohort and 6 months for the *Cacna1c* cohort. The housing room, caging systems, diet, water source, and husbandry practices were held constant throughout each cohort.

The 30 laboratory mouse strains we selected are priority strains in large community genotyping and phenotyping efforts (Table S1). Females from each strain were obtained from the Jackson Laboratory at 6–8 weeks of age and acclimated for 1 week before being placed in harems with a heterozygous (+/−) C57BL/6J male. Once females were visibly pregnant, they were singly housed. F_1 offspring were weaned at 21–24 days of age. They were housed in same-sex cages containing at least two and no more than five littermates of the same genetic background. In the event that there was only one male or female littermate, the animal was excluded from the study. Randomization of genotype and sex occurred as a result of our breeding scheme as wild-type versus heterozygous and male versus female littermates were produced at equal ratios. Litters were combined into testing groups containing up to 48 mice. Females from all 30 strains were bred on a rotating basis such that the production of litters, and the composition of each test group of F_1 mice, was randomized with respect to strain.

All mice within the same cohort (e.g., *Cacna1c* or *Tcf7l2*) had the same behavioral testing schedule (Table S2). All testing was conducted between 0900 hr and 1600 hr. The experimenter conducting each assay was held constant. Experiments began after 7–11 weeks of age and progressed from relatively least stressful to more stressful with at least 4 days between tests. Before each test, mice were acclimated to the test room for at least 30 min in their home cages. They were then placed in a clean holding cage to await the start of testing. Each animal was assigned an arbitrary five-digit identification number that obscured their strain and genotype from the experimenter. Animals were placed in the appropriate testing apparatus by the experimenter and were then directly returned to their home cages after completion of the test. Testing equipment was cleaned with 10% isopropanol between animals. All behavior was monitored and scored by automated software systems, removing the possibility of human bias. Factors specific to the testing environment, including the particular test box in which the animal was placed and the time of day (morning or afternoon), were recorded as potential covariates. Details of each phenotypic test are given in the Supplemental Experimental Procedures.

Statistics

The R statistical environment was used to fit a linear regression model for each phenotype. When appropriate, we transformed some phenotypes onto the logarithmic (base 10) scale, and transformed others using the logit function, $\text{logit10}(x) = \log10(x/(1 - x))$ to ensure that the distribution of residuals met the assumption of normality. We modeled each phenotype as a linear combination of covariates including sex, body weight, coat color, and other experimental factors relevant to each phenotypic analysis. Covariates explaining less than 2% of the variance were not used in the linear regression. Outlying data points with a residual more than two SDs away from the mean of each strain were removed. We confirmed that the residuals of our linear models had empirical quantiles that closely matched the expected quantiles under the normal distribution. ANOVA was conducted to evaluate main effects and interactions using the anova.lm function in R (see Supplemental Information). Because we were interested in the effect of the null alleles in each individual F_1, we performed t tests comparing +/+ to +/− within each F_1. We did not apply a multiple testing correction (e.g., Bonferroni) because we wanted to illustrate the different conclusions that would have been reached had only one genetic background been considered, which is the common practice when evaluating the phenotypic consequences of a mutant allele. The complete phenotype data and code used to complete the analyses are available at https://github.com/pcarbo/neuron.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, one figure, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.neuron.2016.08.013.

A video abstract is available at http://dx.doi.org/10.1016/j.neuron.2016.08.013#mmc4.

AUTHOR CONTRIBUTIONS

A.A.P. and L.J.S. conceived and designed the experiments; L.J.S., K.A.E., K.S.K., and C.M.B.-C. performed the experiments; P.C. and L.J.S. analyzed the data; and L.J.S., A.A.P., and P.C. wrote and edited the manuscript.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Bruce Hamilton for providing comments on a draft of the manuscript. The studies were supported by NIH grants MH102931 and MH102728.

Received: May 24, 2016
Revised: July 11, 2016
Accepted: August 8, 2016
Published: September 8, 2016

REFERENCES

Alkelai, A., Greenbaum, L., Lupoli, S., Kohn, Y., Sarner-Kanyas, K., Ben-Asher, E., Lancet, D., Macciardi, F., and Lerer, B. (2012). Association of the type 2 diabetes mellitus susceptibility gene, TCF7L2, with schizophrenia in an Arab-Israeli family sample. PLoS ONE *7*, e29228.

Bradley, A., Anastassiadis, K., Ayadi, A., Battey, J.F., Bell, C., Birling, M.-C., Bottomley, J., Brown, S.D., Bürger, A., Bult, C.J., et al. (2012). The mammalian gene function resource: the International Knockout Mouse Consortium. Mamm. Genome *23*, 580–586.

Brandon, E.P., Idzerda, R.L., and McKnight, G.S. (1995). Knockouts. Targeting the mouse genome: a compendium of knockouts (Part I). Curr. Biol. *5*, 625–634.

Capecchi, M.R. (1989). Altering the genome by homologous recombination. Science *244*, 1288–1292.

Castle, W.E. (1919). Piebald rats and the theory of genes. Proc. Natl. Acad. Sci. USA *5*, 126–130.

Chesler, E.J., Wilson, S.G., Lariviere, W.R., Rodriguez-Zas, S.L., and Mogil, J.S. (2002). Influences of laboratory environment on behavior. Nat. Neurosci. *5*, 1101–1102.

Clayton, J.A., and Collins, F.S. (2014). Policy: NIH to balance sex in cell and animal studies. Nature *509*, 282–283.

Collins, F.S., Rossant, J., and Wurst, W.; International Mouse Knockout Consortium (2007). A mouse for all reasons. Cell *128*, 9–13.

Crabbe, J.C., Wahlsten, D., and Dudek, B.C. (1999). Genetics of mouse behavior: interactions with laboratory environment. Science *284*, 1670–1672.

Dao, D.T., Mahon, P.B., Cai, X., Kovacsics, C.E., Blackwell, R.A., Arad, M., Shi, J., Zandi, P.P., O'Donnell, P., Knowles, J.A., et al.; Bipolar Genome Study (BiGS) Consortium (2010). Mood disorder susceptibility gene CACNA1C modifies mood-related behaviors in mice and interacts with sex to influence behavior in mice and diagnosis in humans. Biol. Psychiatry *68*, 801–810.

Dexter, J. (1914). The analysis of a case of continuous variation in Drosophila by a study of its linkage relations. Am. Nat. *48*, 712–758.

Dietrich, W.F., Lander, E.S., Smith, J.S., Moser, A.R., Gould, K.A., Luongo, C., Borenstein, N., and Dove, W. (1993). Genetic identification of Mom-1, a major modifier locus affecting Min-induced intestinal neoplasia in the mouse. Cell *75*, 631–639.

Doetschman, T. (2009). Influence of genetic background on genetically engineered mouse phenotypes. Methods Mol. Biol. *530*, 423–433.

Dorman, A., Baer, D., Tomlinson, I., Mott, R., and Iraqi, F.A. (2016). Genetic analysis of intestinal polyp development in Collaborative Cross mice carrying the Apc (Min/+) mutation. BMC Genet. *17*, 46.

Ferreira, M.A.R., O'Donovan, M.C., Meng, Y.A., Jones, I.R., Ruderfer, D.M., Jones, L., Fan, J., Kirov, G., Perlis, R.H., Green, E.K., et al.; Wellcome Trust Case Control Consortium (2008). Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. Nat. Genet. *40*, 1056–1058.

Fuchsberger, C., Flannick, J., Teslovich, T.M., Mahajan, A., Agarwala, V., Gaulton, K.J., Ma, C., Fontanillas, P., Moutsianas, L., McCarthy, D.J., et al. (2016). The genetic architecture of type 2 diabetes. Nature *536*, 41–47.

Grant, S.F.A., Thorleifsson, G., Reynisdottir, I., Benediktsson, R., Manolescu, A., Sainz, J., Helgason, A., Stefansson, H., Emilsson, V., Helgadottir, A., et al. (2006). Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. Nat. Genet. *38*, 320–323.

Green, E.K., Hamshere, M., Forty, L., Gordon-Smith, K., Fraser, C., Russell, E., Grozeva, D., Kirov, G., Holmans, P., Moran, J.L., et al.; WTCCC (2013). Replication of bipolar disorder susceptibility alleles and identification of two novel genome-wide significant associations in a new bipolar disorder case-control sample. Mol. Psychiatry *18*, 1302–1307.

Hamilton, B.A., and Yu, B.D. (2012). Modifier genes and the plasticity of genetic networks in mice. PLoS Genet. *8*, e1002644.

Hamshere, M.L., Walters, J.T.R., Smith, R., Richards, A.L., Green, E., Grozeva, D., Jones, I., Forty, L., Jones, L., Gordon-Smith, K., et al.; Schizophrenia Psychiatric Genome-wide Association Study Consortium; Wellcome Trust Case Control Consortium+; Wellcome Trust Case Control Consortium 2 (2013). Genome-wide significant associations in schizophrenia to ITIH3/4, CACNA1C and SDCCAG8, and extensive replication of associations reported by the Schizophrenia PGC. Mol. Psychiatry *18*, 708–712.

Hansen, T., Ingason, A., Djurovic, S., Melle, I., Fenger, M., Gustafsson, O., Jakobsen, K.D., Rasmussen, H.B., Tosato, S., Rietschel, M., et al. (2011). At-risk variant in TCF7L2 for type II diabetes increases risk of schizophrenia. Biol. Psychiatry *70*, 59–63.

Heydemann, A., Ceco, E., Lim, J.E., Hadhazy, M., Ryder, P., Moran, J.L., Beier, D.R., Palmer, A.A., and McNally, E.M. (2009). Latent TGF-beta-binding protein 4 modifies muscular dystrophy in mice. J. Clin. Invest. *119*, 3703–3712.

Hide, T., Hatakeyama, J., Kimura-Yoshida, C., Tian, E., Takeda, N., Ushio, Y., Shiroishi, T., Aizawa, S., and Matsuo, I. (2002). Genetic modifiers of otocephalic phenotypes in Otx2 heterozygous mutant mice. Development *129*, 4347–4357.

Hrabě de Angelis, M., Nicholson, G., Selloum, M., White, J.K., Morgan, H., Ramirez-Solis, R., Sorg, T., Wells, S., Fuchs, H., Fray, M., et al.; EUMODIC Consortium (2015). Analysis of mammalian gene function through broad-based phenotypic screens across a consortium of mouse clinics. Nat. Genet. *47*, 969–978.

Karp, N.A., Meehan, T.F., Morgan, H., Mason, J.C., Blake, A., Kurbatova, N., Smedley, D., Jacobsen, J., Mott, R.F., Iyer, V., et al. (2015). Applying the ARRIVE guidelines to an in vivo database. PLoS Biol. *13*, e1002151.

Kilkenny, C., Browne, W.J., Cuthill, I.C., Emerson, M., and Altman, D.G. (2010). Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. PLoS Biol. *8*, e1000412.

Lifsted, T., Le Voyer, T., Williams, M., Muller, W., Klein-Szanto, A., Buetow, K.H., and Hunter, K.W. (1998). Identification of inbred mouse strains harboring genetic modifiers of mammary tumor age of onset and metastatic progression. Int. J. Cancer *77*, 640–644.

Little, T.J., and Colegrave, N. (2016). Caging and uncaging genetics. PLoS Biol. *14*, e1002525.

Mallon, A.-M., Blake, A., and Hancock, J.M. (2008). EuroPhenome and EMPReSS: online mouse phenotyping resource. Nucleic Acids Res. *36*, D715–D718.

Morris, A.P., Voight, B.F., Teslovich, T.M., Ferreira, T., Segrè, A.V., Steinthorsdottir, V., Strawbridge, R.J., Khan, H., Grallert, H., Mahajan, A., et al.; Wellcome Trust Case Control Consortium; Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC) Investigators; Genetic Investigation of ANthropometric Traits (GIANT) Consortium; Asian Genetic Epidemiology Network–Type 2 Diabetes (AGEN-T2D) Consortium; South Asian Type 2 Diabetes (SAT2D) Consortium; DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium (2012). Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat. Genet. *44*, 981–990.

Nadeau, J.H. (2001). Modifier genes in mice and humans. Nat. Rev. Genet. *2*, 165–174.

Nyegaard, M., Demontis, D., Foldager, L., Hedemand, A., Flint, T.J., Sørensen, K.M., Andersen, P.S., Nordentoft, M., Werge, T., Pedersen, C.B., et al. (2010). CACNA1C (rs1006737) is associated with schizophrenia. Mol. Psychiatry *15*, 119–121.

Park, S.-Y., Ludwig, M.Z., Tamarina, N.A., He, B.Z., Carl, S.H., Dickerson, D.A., Barse, L., Arun, B., Williams, C.L., Miles, C.M., et al. (2014). Genetic complexity in a Drosophila model of diabetes-associated misfolded human proinsulin. Genetics *196*, 539–555.

Phillips, T.J., Hen, R., and Crabbe, J.C. (1999). Complications associated with genetic background effects in research using knockout mice. Psychopharmacology (Berl.) *147*, 5–7.

Pinto, R.M., Dragileva, E., Kirby, A., Lloret, A., Lopez, E., St Claire, J., Panigrahi, G.B., Hou, C., Holloway, K., Gillis, T., et al. (2013). Mismatch repair genes Mlh1 and Mlh3 modify CAG instability in Huntington's disease mice: genome-wide and candidate approaches. PLoS Genet. *9*, e1003930.

Ripke, S., Sanders, A.R., Kendler, K.S., Levinson, D.F., Sklar, P., Holmans, P.A., Lin, D., Duan, J., and Ophoff, R.A.; Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium (2011). Genome-wide association study identifies five new schizophrenia loci. Nat. Genet. *43*, 969–976.

Rozmahel, R., Wilschanski, M., Matin, A., Plyte, S., Oliver, M., Auerbach, W., Moore, A., Forstner, J., Durie, P., Nadeau, J., et al. (1996). Modulation of disease severity in cystic fibrosis transmembrane conductance regulator deficient mice by a secondary genetic factor. Nat. Genet. *12*, 280–287.

Sanford, L.P., Kallapur, S., Ormsby, I., and Doetschman, T. (2001). Influence of genetic background on knockout mouse phenotypes. Methods Mol. Biol. *158*, 217–225.

Savic, D., Distler, M.G., Sokoloff, G., Shanahan, N.A., Dulawa, S.C., Palmer, A.A., and Nobrega, M.A. (2011a). Modulation ofTcf7l2 expression alters behavior in mice. PLoS ONE *6*, e26897.

Savic, D., Ye, H., Aneas, I., Park, S.-Y., Bell, G.I., and Nobrega, M.A. (2011b). Alterations in TCF7L2 expression define its role as a key regulator of glucose metabolism. Genome Res. *21*, 1417–1425.

Silva, A.J., Simpson, E.M., Takahashi, J.S., Lipp, H.-P., Nakanishi, S., Wehner, J.M., Giese, K.P., Tully, T., Abel, T., Chapman, P.F., et al. (1997). Mutant mice and neuroscience: recommendations concerning genetic background. Banbury Conference on genetic background in mice. Neuron *19*, 755–759.

Singh, P., Schimenti, J.C., and Bolcun-Filas, E. (2015). A mouse geneticist's practical guide to CRISPR applications. Genetics *199*, 1–15.

Sklar, P., Smoller, J.W., Fan, J., Ferreira, M.A.R., Perlis, R.H., Chambert, K., Nimgaonkar, V.L., McQueen, M.B., Faraone, S.V., Kirby, A., et al. (2008). Whole-genome association study of bipolar disorder. Mol. Psychiatry *13*, 558–569.

Smoller, J.W., Ripke, S., Lee, P.H., Neale, B.M., Nurnberger, J., Santangelo, S., Sullivan, P.F., Perlis, R.H., Purcell, S.M., and Fanous, A.; Cross-Disorder Group of the Psychiatric Genomics Consortium (2013). Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. Lancet *381*, 1371–1379.

Sorge, R.E., Martin, L.J., Isbester, K.A., Sotocinal, S.G., Rosen, S., Tuttle, A.H., Wieskopf, J.S., Acland, E.L., Dokova, A., Kadoura, B., et al. (2014). Olfactory exposure to males, including men, causes stress and related analgesia in rodents. Nat. Methods *11*, 629–632.

Threadgill, D.W., Dlugosz, A.A., Hansen, L.A., Tennenbaum, T., Lichti, U., Yee, D., LaMantia, C., Mourton, T., Herrup, K., Harris, R.C., et al. (1995). Targeted disruption of mouse EGF receptor: effect of genetic background on mutant phenotype. Science *269*, 230–234.

Wahlsten, D., Metten, P., Phillips, T.J., Boehm, S.L., 2nd, Burkhart-Kasch, S., Dorow, J., Doerksen, S., Downing, C., Fogarty, J., Rodd-Henricks, K., et al. (2003). Different data from different labs: lessons from studies of gene-environment interaction. J. Neurobiol. *54*, 283–311.

Winham, S.J., Cuellar-Barboza, A.B., Oliveros, A., McElroy, S.L., Crow, S., Colby, C., Choi, D.-S., Chauhan, M., Frye, M., and Biernacka, J.M. (2014). Genome-wide association study of bipolar disorder accounting for effect of body mass index identifies a new risk allele in TCF7L2. Mol. Psychiatry *19*, 1010–1016.