Identification of a Functional Non-coding Variant in the GABA$A$ Receptor $\alpha$2 Subunit of the C57BL/6J Mouse Reference Genome: Major Implications for Neuroscience Research

Megan K. Mulligan1*, Timothy Abreo1, Sarah M. Neuner2,3, Cory Parks1, Christine E. Watkins1, M. Trevor Houseal1, Thomas M. Shapaker1, Michael Hook1, Haiyan Tan4, Xusheng Wang4, Jesse Ingels4, Junmin Peng4, Lu Lu1, Catherine C. Kaczorowski5, Camron D. Bryant6, Gregg E. Homanics6 and Robert W. Williams1

1 Department of Genetics, Genomics and Informatics, University of Tennessee Health Science Center, Memphis, TN, United States, 2 Department of Anatomy and Neurobiology, University of Tennessee Health Science Center, Memphis, TN, United States, 3 The Jackson Laboratory, Bar Harbor, ME, United States, 4 Departments of Structural Biology and Developmental Neurobiology, Center for Proteomics and Metabolomics, St. Jude Children's Research Hospital, Memphis, TN, United States, 5 Laboratory of Addiction Genetics, Department of Pharmacology and Experimental Therapeutics and Psychiatry, Boston University School of Medicine, Boston, MA, United States, 6 Departments of Anesthesiology and Perioperative Medicine, Neurobiology, and Pharmacology and Chemical Biology, University of Pittsburgh, Pittsburgh, PA, United States

GABA$A$ receptors containing the $\alpha$2 subunit (GABRA2) are expressed in most brain regions and are critical in modulating inhibitory synaptic function. Genetic variation at the GABRA2 locus has been implicated in epilepsy, affective and psychiatric disorders, alcoholism and drug abuse. Gabra2 expression varies as a function of genotype and is modulated by sequence variants in several brain structures and populations, including F2 crosses originating from C57BL/6J (B6J) and the BXD recombinant inbred family derived from B6J and DBA/2J. Here we demonstrate a global reduction of GABRA2 brain protein and mRNA in the B6J strain relative to other inbred strains, and identify and validate the causal mutation in B6J. The mutation is a single base pair deletion located in an intron adjacent to a splice acceptor site that only occurs in the B6J reference genome. The deletion became fixed in B6J between 1976 and 1991 and is now pervasive in many engineered lines, BXD strains generated after 1991, the Collaborative Cross, and the majority of consomic lines. Repair of the deletion using CRISPR-Cas9-mediated gene editing on a B6J genetic background completely restored brain levels of GABRA2 protein and mRNA. Comparison of transcript expression in hippocampus, cortex, and striatum between B6J and repaired genotypes revealed alterations in GABA$A$ receptor subunit expression, especially in striatum. These results
suggest that naturally occurring variation in GABRA2 levels between B6J and other substrains or inbred strains may also explain strain differences in anxiety-like or alcohol and drug response traits related to striatal function. Characterization of the B6J private mutation in the Gabra2 gene is of critical importance to molecular genetic studies in neurobiological research because this strain is widely used to generate genetically engineered mice and murine genetic populations, and is the most widely utilized strain for evaluation of anxiety-like, depression-like, pain, epilepsy, and drug response traits that may be partly modulated by GABRA2 function.

Keywords: GABRA2, C57BL/6J, GABA-A receptor, neuroscience, genetics, CRISPR (clustered regularly interspaced short palindromic repeats)/Cas9, BXD mice strains, inbred mouse strains

INTRODUCTION

GABA-A receptors are pentameric ligand gated chloride channels responsible for fast inhibitory neurotransmission. The sequence, structure, and chromosomal location of receptors and their cognate subunits are highly conserved among mammals (Olsen and Sieghart, 2009). Receptors are typically composed of two α, two β, and a γ or δ subunit, with the most common receptor subtypes containing α1 or α2 subunits. Subunit composition can alter receptor pharmacology and localization (Sieghart, 1995). Receptors containing the α2 subunit (GABRA2) are abundantly expressed in mouse hippocampus, frontal cortex, amygdala, dorsal and ventral striatum, and hypothalamus—brain regions important for motivation, reward, anxiety, depression, and fear. Over 500 common human variants have been identified at the GABRA2 locus, including numerous intronic variants, a handful localized to the UTRs, and one synonymous variant in exon 4. Genetic variation at this locus has been linked to alcohol dependence (Covault et al., 2004; Edenberg et al., 2004; Lappalainen et al., 2005; Agrawal et al., 2006; Enoch et al., 2006; Fehr et al., 2006; Lind et al., 2008; Luyckx et al., 2008; Villafuerte et al., 2013; Li et al., 2014), the subjective effects of alcohol (Pierucci-Lagha et al., 2005; Roh et al., 2011; Uhr et al., 2013; Arias et al., 2014), excess EEG fast activity (Lydall et al., 2011), cocaine reward (Dixon et al., 2010), substance dependence (Enoch et al., 2010), epilepsy (International League Against Epilepsy Consortium on Complex Epilepsies, 2014), and impulsivity and insula cortex activity during reward or loss anticipation (Villafuerte et al., 2013). Furthermore, interactions between early life stress and GABRA2 polymorphisms may influence vulnerability to addiction (Enoch, 2008; Dick et al., 2009; Enoch et al., 2010). However, the impact of these variants on expression, isoform usage, or function of GABRA2 protein or mRNA is unclear (Myers et al., 2007; Heinzen et al., 2008; Gibbs et al., 2010; Liu et al., 2010; Kim et al., 2014; Lieberman et al., 2015).

Preclinical models can be useful in dissecting enigmatic relationships between human behavior and genetic variation. Genetically engineered mouse models in which Gabra2 expression or sensitivity to modulation has been altered (knock-out and knock-in) suggest a role for GABA-A receptors containing the α2 subunit in depression (Vollenweider et al., 2011), alcohol intake and behavioral response to alcohol (Boehm et al., 2004; Blednov et al., 2011, 2013; Liu et al., 2011), anxiety (Dixon et al., 2008, and cocaine-conditioned reinforcement and behavioral sensitization (Morris et al., 2008; Dixon et al., 2010). Genetic deletion of Gabra2 is also associated with a decrease in miniature inhibitory postsynaptic current in the nucleus accumbens core (Dixon et al., 2010), suggesting an important role of GABRA2 in modulating inhibitory signaling in this region. However, no studies to date have characterized naturally occurring variants segregating among murine populations that could alter GABRA2 levels or function, and provide insight into the role of human variants.

In this study, we leverage naturally occurring variation in murine populations of varying genetic complexity in order to better understand the role of Gabra2 variants on gene function. We previously profiled gene expression data from multiple strains and crosses, including nearly isogenic C57BL/6 substrains and the well-characterized BXD family of strains derived from a cross between B6J and DBA/2J (D2) inbred mice, to characterize genetic variation in Gabra2 expression (Mulligan et al., 2012). Here we identify a spontaneous intronic deletion in B6J that underlies variation in GABRA2 protein and mRNA expression across diverse murine populations that incorporate B6J as a progenitor strain. We have quantified molecular consequences of the private mutation in B6J by CRISPR-Cas9 genome editing. This work demonstrates the functionality of often ignored non-coding variants in the mouse genome and characterizes the downstream consequences of Gabra2 variation on molecular traits.

MATERIALS AND METHODS

Mice

B6(Cg)-Tyr<sup>e<sup>−<sup>2J/J mice were purchased from Jackson Laboratory and C57BL/6NCrl mice were purchased from Charles River Laboratories at 8-weeks-of-age and allowed to acclimate for 9 days at UTHSC before tissue collection. C57BL/6J, DBA/2J, C57BL/6NJ, C57BL/6JeiJ, and C57BL/6ByJ were maintained in-house up to generation 6 at UTHSC from Jackson Laboratory stock. BXD mice were maintained at UTHSC. The B6-Gabra2<sup>2<sup>e<sup>mi6<sup>6<sup>6<sub>J CRISPR repair line (homozygous Gabra2<sup>2<sup>e<sup>mi6<sup>6<sub>J/KJ) was generated at the University of Pittsburgh (as described below) and is maintained as a heterozygous
breeding colony by the Mulligan laboratory at UTHSC. All mice were housed in same-sex cages, allowed ad-lib access to food and water and maintained on a 12h/12h light/dark cycle in a climate-controlled facility. Only adult mice between 70 and 200 days of age were included in the study. All animal care and handling procedures were approved by the UTHSC and the University of Pittsburgh Institutional Animal Care and Use Committees.

Data Sets Used
The following data sets available at GeneNetwork.org were used in the analysis: GN76 OHSU/VA B6D2F2 Brain mRNA M430 (Aug05) RMA (Hitzemann et al., 2004), GN84 OHSU/VA B6D2F2 Striatum M430v2 (Sep05) RMA (Hitzemann et al., 2004), GN281 INIA Hypothalamus Affy MoGene 1.0 ST (Nov10) (Mozhui et al., 2012), GN323 INIA Amygdala Cohort Affy MoGene 1.0 ST (Mar11) RMA, GN110 Hippocampus Consortium M430v2 (Jun06) RMA (Overall et al., 2009), GN175 UCLA BHBF2 Brain (2005) mirtrio (Orozco et al., 2009; van Nas et al., 2010), GN171 UCLA CTB6/B6CTF2 Brain (2005), mirtrio, INIA LCM (11 Regions) BASELINE RNA-seq Transcript Level, GN206 UMTA Hippocampus Exon (Feb09) RMA, and GN732 GTEXv5 Human Brain Hippocampus RefSeq (Sep15) RPKM log2 (Lonsdale et al., 2013). In addition to references provided, additional information on these data sets can also be found at GeneNetwork.org.

Tissue Collection
The cortex, hippocampus, and striatum (dorsal and ventral) were dissected from up to three males and three females for each strain (effective strain n = 5 or 6). At time of tissue harvest, purchased mice were ~9 weeks and in-house strains were ~8 weeks of age. Tissue was immediately frozen in liquid nitrogen and stored at −80°C until further use. Additional peripheral tissue, including the liver and spleen, was also dissected from each strain and stored at −80°C until further use. Whole brains from wild-type mice were housed in same-sex cages, allowed a breeding colony by the Mulligan laboratory at UTHSC. All in-house mice were dissected from each strain and stored at −80°C until further use. Whole brains from wild-type mice were housed in same-sex cages, allowed a breeding colony by the Mulligan laboratory at UTHSC. All in-house mice were dissected from each strain and stored at −80°C until further use.

qPCR to Validate mRNA Levels
Five biological replicates were used for each strain and consisted of both males and females (two or three of each sex). RNA was isolated from brain tissue using the RNeasy Lipid Tissue Kit (Qiagen), with cDNA then being generated from 1 µg of RNA utilizing the First Strand Reverse Transcription kit (Roche). Quantitative PCR was performed to assess target gene expression using the KAPA SYBR Fast qPCR Master Mix, probes from Roche’s Universal ProbeLibrary, custom primers from Integrated DNA Technologies (IDT), and a Roche Light Cycler 480 II. Universal probes and custom primers were selected and designed, respectively, using the Universal Probe Library (LifeScience). For Gabra2 (NM_008066.3) we used Universal Probe #103 and the following primers: Left–ACAAAAAGGGATGGCTTG and Right–TCATGACGGAG CCTTTCTCCT. Cyclophilin D (NM_026352.3) was selected as the endogenous control for all tissues (Schmittgen and Livak, 2008). Data was analyzed using the ΔΔCt method as described in Livak and Schmittgen (2001), with B6j selected as the reference or calibrator strain to calculate fold-change expression. For gene expression in the CRISPR mice, the coefficient of variation was quantified using the ΔCt method. All statistical tests were conducted in R using R Studio (RStudio: Integrated Development for R. RStudio, Inc., Boston, MA, United States). One-way ANOVA was performed when comparing more than two independent samples and Tukey’s HSD post hoc test subsequently applied to assess pairwise significance between genotypes.

Validation of Genomic Variants
Variants that were polymorphic between the B6j reference genome and B6NJ and D2 were identified using 100X coverage of the D2 genome1 (Wang et al., 2016) and the Welcome Trust Sanger Institute (sequenced genomes available for 36 inbred mouse strains, including the B6NJ substrain²). Initially we identified two high quality indels located on Chr 5 at 71,014,638 bp (deletion) and 71,031,384 bp (rs225241970, deletion), and a highly conserved intergenic SNP located at 70,931,531 bp (all coordinates given based on the GRCm38/mm10 assembly). Genomic DNA was extracted from B6J, D2, B6NJ, B6EiJ, B6ByJ, BXD29, and BXD40 using the DNeasy Blood & Tissue Kit (Qiagen) from frozen liver or spleen tissue. Variants were genotyped using primers designed against the reference genome using the Primer3Plus web tool and purchased from IDT. Primers targeting each variant include: (1) SNP rs29547790@70,931,531 bp Left–AAAAGTGGGTTGGTTGG and Right–GGAGTGCACT TCTCTTCTTTGG, (2) indel@71,014,638 bp Left–TCAGGA GCCTCAGTTTTGTGCTG and Right–TCTTCTAGCTCGTTCTGTT CTGTAA, and (3) indel rs225241970@71,031,384 bp Left–AG ACCCCTTGGGAAGAAGAG and Right–GGTCTCTACAGGA AATGAAACGG. Genomic intervals (~100–300 bp) containing the variants were amplified by PCR and unincorporated nucleotides and primers removed using ExoSAP-IT (Thermo Fisher Scientific). Traditional Sanger capillary sequencing

1http://ucscbrowser.genenetwork.org/
2http://www.sanger.ac.uk/cgi-bin/modelorgs/mousegenomes/snps.pl
was performed by the UTHSC Molecular Resource Center Institutional Core using the ABI Prism 3130 Genetic Analyzer system. The only variant that could not be confirmed was the indel rs225241970 located on Chr5 at 71,031,384 bp as none of the seven strains tested, including B6J, were polymorphic at this position. This genomic region includes a microsatellite repeat (CA)n that may have led to sequencing errors in the reference genome.

**Traditional Western Analysis**

Lysates were prepared from frozen tissue and protein concentration was determined using a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific). 50 μg of total protein was loaded and separated on a 7.5% SDS-PAGE gel. Proteins were transferred using the Bio-Rad Turbo Transfer system and blocked for 30 min at room temperature. For chemiluminescence western blots, blots were incubated in the Super Signal West Pico Chemiluminescent Substrate kit (Thermo Fisher Scientific #34080) on a BioRad ChemiDoc imaging system. Blots were then incubated in a stripping solution (Restore Western Blot Stripping Buffer, Thermo Fisher Scientific #21059) for 8 min, re-probed overnight at 4°C with a GAPDH antibody (1:5000, mouse monoclonal, Fitzgerald, Scientific #21059) for 8 min, re-probed overnight at 4°C. Blots were washed three times (5 min each) with blocking buffer and incubated with horseradish-peroxidase conjugated antibody for 1 h at room temperature. Blots were developed using the Odyssey image scanner. Blot intensities across all blots were quantified using ImageJ. Original blots can be found in Supplementary Figures S1, S2.

**Proteomics Analysis by 6-Plex TMT-LC/MS/MS**

The analysis was performed essentially based on the previously reported protocol (Bai et al., 2017; Niu et al., 2017). In brief, hippocampus was dissected from male B6J and D2 mice to create pooled samples consisting of three animals each. Pooled samples represented B6J mice aged approximately 60 days (two pooled samples) or 1 year (one pooled sample) and D2 mice aged 90 days (two pooled samples) or 1 year (one pooled sample). Tissue from pooled samples was lysed and digested into peptides. After desalting, the peptides were labeled with TMT reagents and filtered to achieve a 1% protein false discovery rate. TMT intensities were extracted, filtered, normalized, and summarized into peptide and protein quantification results. A total of 22,005 proteins representing 7,074 protein groups were identified and 7,014 protein groups were quantified. Statistical analysis was performed to determine cutoff for altered proteins and to evaluate associated false discovery rate. In this case the significance threshold was set at \( p < 0.005 \) with a minimum of 3 peptides detected. A total of 8 peptides and 34 spectral counts were detected for GABRA2 (full data set provided in Supplementary Table S1).

**Generation of CRISPR-Engineered Mice**

Knockin mice (B6-Gabra2\textsubscript{em1Geh/J} or Gabra2\textsubscript{KI/KI}) were produced with CRISPR/Cas9 using technologies as described (Blednov et al., 2017). Briefly, a sgRNA targeting Gabra2 was generated at the intron/exon junction near Chr 5 at 71,014,638 bp (GRCm38/mm10 assembly) was identified using the WTSI Genome Editing database (Hodgkins et al., 2015). Two partially overlapping PCR primers (F:GAAATTAATACGACTATAGGCAGTTTTAGAGCTAGAAATAGC; R:AAAAGCACGGAGCTGTTTTCAAAGGAATTGTAAATTTATTTATACTGGCTATTTCTAGCTCTAAAAC) were used to generate a T7 promoter based sgRNA template (sgRNA target sequence underlined in above sequence) as described (Bassett et al., 2014). The sgRNA and Cas9 mRNA were produced by \textit{in vitro} transcription, purified using the MEGAclean Kit (Ambion), ethanol precipitated, and re-suspended in DEPC treated water. A 121 nucleotide single stranded DNA repair template oligo (TTTTATAGCCCTACTATCTTCAAACAGATGACATCTCTTTCAAAGGAATTGTAAATTTATTTATTTAGAGCTAGAAATAGC; R:AAAAGCACGGAGCTGTTTTCAAACAGATGACATCTCTTTCAAAGGAATTGTAAATTTATTTATTT) harboring the desired T insertion (bold, italic in above sequence) in the intron of Gabra2 was purchased as Ultramer DNA (Integrated DNA Technologies, Coralville, IA). sgRNA (100 ng/µl), Cas9 mRNA (75 ng/µl), and repair oligo (100 ng/µl) were injected into the cytoplasm of B6J one-cell embryos as described (Yang et al., 2014; Blednov et al., 2017). Pups resulting from injected embryos were screened for DNA sequence changes in intron 3 of the Gabra2 gene by PCR/DNA sequence analysis. Briefly, a 447 bp ampiclon spanning the knockin site was PCR amplified with forward (CAACCGAGGAGGAAAGAGATGACGCTGTTTTCTAGCTCTAAAAC) and reverse (TTCGAAGCAGCTGTTTTCAAAGGAATTGTAAATTTATTTATTTA) primers. PCR products were sequenced directly or subcloned into pCR2.1-TOPO (Invitrogen) and sequenced. All mice were genotyped by DNA sequence analysis (Sanger sequencing) using this approach.

The male founder (F0) was subsequently crossed to a female B6J mouse to generate F1 progeny. F1 mice were crossed to generate F2 mice. The colony is maintained through heterozygous breeding and all molecular phenotyping was performed in generations F2 and higher.

**Analysis of CRISPR/Cas9 Off-Target Effects**

The CRISPOR program v4.0\textsuperscript{3} was used to search for off-target sites (Haeussler et al., 2016). For precise genome editing, \textit{cis} and \textit{trans} effects were analyzed by T7 promoter containing sgRNA template (sgRNA target sequence underlined in above sequence) as described (Bassett et al., 2014). The sgRNA and Cas9 mRNA were produced by \textit{in vitro} transcription, purified using the MEGAclean Kit (Ambion), ethanol precipitated, and re-suspended in DEPC treated water. A 121 nucleotide single stranded DNA repair template oligo (TTTTATAGCCCTACTATCTTCAAACAGATGACATCTCTTTCAAAGGAATTGTAAATTTATTTATTT) harboring the desired T insertion (bold, italic in above sequence) in the intron of Gabra2 was purchased as Ultramer DNA (Integrated DNA Technologies, Coralville, IA). sgRNA (100 ng/µl), Cas9 mRNA (75 ng/µl), and repair oligo (100 ng/µl) were injected into the cytoplasm of B6J one-cell embryos as described (Yang et al., 2014; Blednov et al., 2017). Pups resulting from injected embryos were screened for DNA sequence changes in intron 3 of the Gabra2 gene by PCR/DNA sequence analysis. Briefly, a 447 bp ampiclon spanning the knockin site was PCR amplified with forward (CAACCGAGGAGGAAAGAGATGACGCTGTTTTCTAGCTCTAAAAC) and reverse (TTCGAAGCAGCTGTTTTCAAAGGAATTGTAAATTTATTTA) primers. PCR products were sequenced directly or subcloned into pCR2.1-TOPO (Invitrogen) and sequenced. All mice were genotyped by DNA sequence analysis (Sanger sequencing) using this approach.

The male founder (F0) was subsequently crossed to a female B6J mouse to generate F1 progeny. F1 mice were crossed to generate F2 mice. The colony is maintained through heterozygous breeding and all molecular phenotyping was performed in generations F2 and higher.

\textsuperscript{3}crispor.tefor.net
the sgRNA must match a 20 nucleotide target sequence in the genomic DNA (protoscaler sequence) that includes a protoscaler adjacent motif (PAM) with sequence NGG (where N is any nucleotide). The sgRNA target sequence (TGAATCAT AAACCTATATTAGG) and PAM used to generate CRISPR-engineered mice was predicted to generate 245 off-targets using CRISPOR. However, all predicted off-targets included two (8 predicted off-target sites), three (57 predicted off-target sites), or four (380 predicted off-target sites) mismatches in the 12 bp against the PAM site. Perfect matches are more likely to generate off-target effects compared to mismatches. All off-targets were given an off-target score [cutting frequency determination or CFD score (Doench et al., 2016)] in CRISPOR. We ranked candidates by the CFD score and selected the top 15 predicted off-target sites for confirmation using traditional capillary sequencing (described above). All of these sites contained at least 4 mismatches and were located in intergenic or intronic regions (Supplementary Table S2) with high sequence complexity and repetitive sequences. As a result, two of the primer pairs failed to produce a PCR product, even when using B6J genomic DNA.

Analysis of Global Gene Expression in CRISPR-Engineered Gabra2 Mice

Cortex, hippocampus, and striatum were harvested from three male and three female Gabra2<sup>B6</sup>/B6 and Gabra2<sup>K1/K1</sup> mice aged 100 to 130 days-of-age as described above. RNA was extracted using the RNeasy Lipid Tissue kit (Qiagen) and contaminating genomic DNA was removed following treatment with RNase-free DNase (Qiagen). RNA quantity was measured using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific) and quality was assessed using the RNA nano chip and Bioanalyzer (Agilent). All samples had RNA integrity values (RIN) over 8. Gene expression was measured using the mouse Clariom D Assay (Thermo Fisher Scientific). Sample preparation and hybridization were performed according to the manufacturer’s protocol and performed by the UTHSC MRC Institutional Core. Affymetrix Expression Console Software (Thermo Fisher Scientific) was used to annotate and process raw probe cell intensity files. The Gene Level-SST-RMA algorithm was used to perform signal space transformation (SST) and guanine cytosine count correction prior to robust multi-array average (RMA) normalization on the batch of 36 samples (3 brain regions × 2 sexes × 3 biological replicates × 2 genotypes). There were no outliers following normalization. The data set was filtered to include only “Main” category probe sets corresponding to annotated coding and non-coding transcripts with a variance > 0.1 across all samples. This reduced the data set from 72,688 to 12,699 probe sets. This data set was used for enrichment analysis and to investigate possible changes in other subunits following reduction of Gabra2. For each analysis, genes with differential expression between Gabra2<sup>K1/K1</sup> and Gabra2<sup>B6</sup>/B6 were detected using Student’s t-test in R to compare expression of each gene by genotype. A nominal p-value was used for both analyses (no adjustment for multiple testing) because of the exploratory nature of enrichment analysis and in order to perform the comparative analysis across GABA<sub>α</sub> receptor subunit genes. Male and female samples of each genotype were pooled for the analysis.

Enrichment analysis was performed for differentially expressed genes (p < 0.01, n = 684) in the striatum using WebGestalt (Zhang et al., 2005; Wang et al., 2013). Overrepresentation of functional terms in the differentially expressed gene set compared to the reference set (consisting of the 12,699 filtered probe sets) was calculated using the Overrepresentation Enrichment Analysis (ORA) method using the following functional databases: Gene Ontology (GO) biological process, GO cellular component, GO molecular function, KEGG, and mammalian phenotype ontology (MPO). Significant enriched terms were identified using the following criteria: (1) minimum number of three genes in a category and (2) a False Discovery Rate (FDR) of <0.05 or nominally significant p-value < 0.01. For analysis of compensatory changes, the expression of all 19 major subunits was compared between genotypes in the cortex, striatum, and hippocampus.

RESULTS

Reduced Levels of Gabra2 Occur Only in the B6J Substrain

The B6J strain showed the lowest levels of Gabra2 among 16 common and wild-derivative inbred strains, including the closely related C57BL/6ByJ strain (Figure 1A). This difference was robustly detected across expression platforms and brain regions. For example, average log2 expression of Gabra2 in the hippocampus was 9.3 for B6J and 11.2 for D2 with over 3.6-fold variation across inbred strains of mice (Affymetrix M430 array, Figure 1A). The decrease in Gabra2 mRNA levels was global, and this reduction relative to D2 was also detected in 10 mesocorticollimbic regions (RNA-seq, Figure 1B). To evaluate whether reduced levels of Gabra2 were only evident in B6J, we examined expression in a small pedigreed of C57BL/6 substrains with defined dates of separation—B6J, B6(Cg)-Tyr<sup>f<sup>−</sup>2J/J (B6C2J), C57BL/6EiJ (B6EiJ), C57BL/6ByJ (B6ByJ), C57BL/6NJ (B6NJ), and C57BL/6NCrJ (B6NCrJ) (Figure 1C). Of these substrains, B6EiJ was the most recently separated from B6J in 1976. The albino strain (B6C2J) was maintained as an inbred colony since the discovery of the albino mutation around 1970 (Townsend et al., 1981) but over the past decade has been repeatedly backcrossed to B6J stock and is therefore now listed as a congenic (Cg) strain of B6 rather than as an independent strain. Expression of Gabra2<sub>C</sub> in cortex, striatum, and hippocampus was significantly higher (p < 0.05) in B6EiJ, B6ByJ, B6NJ, B6NCrJ, and D2 as compared to B6J and B6C2J (Figure 1D). These results explain the remarkable variation in Gabra2 levels (over 4-fold) observed both within sets of BXD strains based on time of generation between the late 1970’s and the early 2000’s (Taylor et al., 1999; Peirce et al., 2004) (Figures 1D–E and Supplementary Figure S3) and recent crosses between B6J and C3H/HeJ or B6J and CAST/EiJ (Mulligan et al., 2012). Lower expression in these crosses was always associated with inheritance of the B6J parental allele of Gabra2. F2 intercrosses can also be...
FIGURE 1 | Continued
lowest expressing individuals from a panel of inbred strains B6J and derivative congenic strains, such as B6C2J, are the result of the wild derived strains (e.g., CAST/EiJ, PWK/PhJ, and MRL). Expression to both the D2 strain and the B6NJ substrain (0.77 ± 0.07) less GABRA2 normalized peptide counts were detected in B6J versus D2, indicating a global reduction in B6J and all known congenic strains relative to B6J shown at left. The B6N lineage (separated in 1951 at generation F2 from B6J) includes C57BL/6ByJ (B6ByJ), C57BL/6NCrl (B6NCrl), and C57BL/6NJ (B6NJ) and all showed higher expression of GABRA2 relative to B6J. The CAST/EiJ, PWK/PhJ, and MRL strains showed lower expression compared to the D2, B6NJ, and BXD40 substrains. Intradividual variability was observed among the strains, with wild derived strains (e.g., CAST/EiJ, PWK/PhJ, and MRL) showing lower expression compared to the D2, B6NJ, and BXD40 substrains.

To identify the causal variant, we genotyped D2, C57BL/6 substrains, and BXD strains that differed in the level of Gabra2 expression based on inheritance of the presumed wild type allele prior to the late 1970s (D2, B6ByJ, B6NJ, B6EiJ and BXD29) or the mutant allele which was fixed in the B6J substrain between 1976 and 1991. The intronic expression level (measured by qPCR) in the cortex, hippocampus, and striatum of D2, C57BL/6 substrains, and congenic strains relative to B6J shown at left. The B6N lineage (separated in 1951 at generation F2 from B6J) includes C57BL/6ByJ (B6ByJ), C57BL/6NCrl (B6NCrl), and C57BL/6NJ (B6NJ) and all showed higher expression of GABRA2 relative to B6J. The CAST/EiJ, PWK/PhJ, and MRL strains showed lower expression compared to the D2, B6NJ, and BXD40 substrains. Intradividual variability was observed among the strains, with wild derived strains (e.g., CAST/EiJ, PWK/PhJ, and MRL) showing lower expression compared to the D2, B6NJ, and BXD40 substrains.

Identification of Candidate Non-coding Sequence Variants Modulating Gabra2 Expression

The Gabra2 gene resides in a genomic region that is largely identical by descent among inbred strains, with the exception of the wild derived strains (e.g., CAST/EiJ, PWK/PhJ, and SWR/EiJ) and contains very few known sequence variants, most of which are located within intronic or intergenic regions. B6J and derivative congenic strains, such as B6C2J, are the lowest expressing individuals from a panel of inbred strains and C57BL/6 substrains (Figure 1). Thus, the causal sequence variant must share a common pattern that differentiates B6J from all other inbred strains and substrains. To identify candidate variants near the Gabra2 locus, we used existing resources generated by our group [100X coverage of the D2 genome4 (Wang et al., 2016) and the Wellcome Trust Sanger Institute (>30X average coverage of the genomes of 36 inbred mouse strains, including the B6NJ substrain5)]. Using these resources, we identified one SNP and two insertion/deletions (indels) that were of high sequence quality. All three were private to B6J compared to B6NJ and D2. We were able to independently validate one SNP and one indel by capillary-based dideoxynucleotide sequencing. Validated candidate polymorphisms were located either ~30 Kb downstream of the Gabra2 gene locus in an intergenic region with high conservation among the mammalian lineage (SNP rs29547790 located on Chr 5 at 70,931,531 bp) or in an intron (the third nucleotide from exon 4) adjacent to a splice acceptor site (indel located on Chr 5 at 71,014,638 bp) (Figure 2A).

To identify the causal variant, we genotyped D2, C57BL/6 substrains, and BXD strains that differed in the level of Gabra2 expression based on inheritance of the presumed wild type allele prior to the late 1970s (D2, B6ByJ, B6NJ, B6EiJ and BXD29) or the mutant allele which was fixed in the B6J substrain between 1976 and 1991 (B6J and BXD40). Importantly, only the genotype at the intronic variant was perfectly associated with Gabra2 expression levels (Figure 2A). Strains with the mutant allele (single base deletion), such as B6J and BXD40, had low expression of Gabra2 compared to the five strains with the wild type allele. The intronic variant was located near a splice acceptor site but the specific

4http://uscbrowser.genenetwork.org/
5http://www.sanger.ac.uk/cgi-bin/mologs/mousegenomes/snps.pl

---

FIGURE 1 | Strain variation in Gabra2 expression. (A) Low expression of Gabra2 in B6J brain relative to other inbred strains and substrains. B6JJ demonstrated the lowest expression of Gabra2 in all brain regions (data set GN110, probe set 1421738_at) compared to 15 inbred strains, including the D2 strain, and a closely related B6J congenic strain (C57BL/6ByJ). (B) Gabra2 expression is reduced in 10 mesocorticolimbic regions relative to D2, indicating a global reduction in B6J. (C) Approximate time line sequestering C57BL/6 substrains. (D) Average Gabra2 expression level (measured by qPCR) in the cortex, hippocampus, and striatum of D2, C57BL/6 substrains, and congenic strains relative to B6J shown at left. The B6N lineage (separated in 1951 at generation F2 from B6J) includes C57BL/6ByJ (B6ByJ), C57BL/6NCrl (B6NCrl), and C57BL/6NJ (B6NJ) and all showed higher expression of Gabra2 relative to B6J. Even the C57BL/6NJ (B6NJ) substrain that diverged most recently from C57BL/6 (B6J) in 1976 showed higher brain expression of Gabra2. The only substrain with low Gabra2 expression similar to C57BL/6 (B6J) is the B6C2J (B6C2J), which was derived using an unbiased shotgun proteomics approach. Relative expression of Gabra2 protein shown as normalized relative intensity for Gabra2 in B6J and D2 hippocampus. Eight different peptide sequences that matched Gabra2 were used to generate normalized relative intensity counts for each sample. Each sample is 3 pooled animals (males only). (H) Hippocampal protein levels of Gabra2 are reduced in B6J (n = 3) compared to the closely related B6NJ (n = 3) substrain and D2 (n = 2) strain (Western blot analysis). The experiment was repeated twice and both B6J sample 1 and B6N sample 1 were run each time. Tissue from Gabra2 KO mice and their wild type control strain (WT) were kindly provided by Y. Blednov at The University of Texas. These mice were originally maintained on a mixed C57BL/6 and 129/SvEv background but were backcrossed twice to B6J upon arrival at the University of Texas, the WT control strain has low Gabra2 protein expression similar to B6J. ANOVA for effect of strain is significant (F(2,5) = 103.9, p < 0.001). Post hoc tests show B6J is different from both B6N and D2 (p = 0.008) but B6NJ and D2 are not significantly different from each other (p = 0.15).
FIGURE 2 | Identification of a B6J private variant in Gabra2. (A) Position of variants near the Gabra2 gene locus. An intergenic SNP (yellow arrowhead) and a single nucleotide indel (red arrowhead) in the third intron were the only confirmed variants among B6J and other inbred mouse strains. Genotyping of strains with high (red) or low (blue) expression of Gabra2 revealed that the likely causal variant is the intronic indel (single base pair deletion in B6J). (B) Evidence for disruption of mRNA processing is shown as an accumulation of intronic reads in the affected intron three (designated with an asterisk) and downstream introns (four and five) in B6J (blue) compared to D2 (red). Normalized and binned read pileup of poly-A enriched striatal RNA-seq data from 10 to 11 B6 and D2 individuals, respectively, generated by Bottomly and colleagues (Bottomly et al., 2011) and hosted for viewing and analysis at the GeneNetwork mirror of the UTHSC Genome Browser (http://ucscbrowser.genenetwork.org/). Gabra2 gene model is shown above (5’ to 3’ orientation with dashed line representing introns and solid vertical bars representing exons).

(Continued)
impact on splicing is not known. However, striatal RNA-seq data (Bottomly et al., 2011) generated for B6J (n = 10) and D2 (n = 11) revealed a higher number of reads mapping both to the intron containing the enigmatic deletion and two downstream introns in the B6J strain (Figure 2B) which could indicate errors in mRNA processing. Significantly higher intronic expression (probe 4462485; \( p < 0.001 \)) was also detected among BXD strains (BXD33 through 100) that inherited the B6J allele (6.1 ± 0.22) compared to those that inherited the D2 allele (5.06 ± 0.13) using an exon microarray platform (GeneNetwork data set GN206, UMIUTaffy Hippocampus Exon (Feb09) RMA; Figure 2C). All Gabra2 exons are expressed at the mRNA level in BXD strains regardless of allele, suggesting that the variant does not pose any mRNA isoforms. Instead, all coding exons have higher expression in strains that inherited the D2 (D) allele, and three introns containing or immediately downstream of the variant have higher expression in strains that inherited the B6J (B) allele (Figure 2C). In sum, a non-coding intronic deletion is the sole candidate variant responsible for decreased GABRA2 protein and mRNA expression in B6J.

Repair of the Candidate Non-coding Variant Is Sufficient to Restore GABRA2 Levels in the B6J Genetic Background

To test whether the intronic deletion is sufficient to globally reduce brain mRNA and protein in B6J we used the CRISPR (clustered regularly interspaced short palindromic repeats) and CRISPR associated (Cas9) system to repair the mutation on a pure B6J genetic background (B6-\(Gabra2^{em1Geh}\); Figure 3A). Insertion (knockin; KI) of a single nucleotide was sufficient to fully restore GABRA2 protein (Figure 3B and Supplementary Figure S2) and mRNA levels (Figure 3C) in brain tissue of B6-\(Gabra2^{em1Geh}\) homozygous knockin (KI, \(Gabra2^{K1/K1}\)) and heterozygous (HET, \(Gabra2^{K1/B6}\)) F2 offspring (Figure 3D). These results are consistent with the recessive mode of inheritance of the mutant and hypomorphic B6J allele (\(Gabra2^{B6}\)) that is associated with reduced expression due to the presence of the intronic deletion. The use of CRISPR engineering can produce both off-target effects and founders that are mosaic for the introduced mutation (e.g., some cells may have the mutation, some may not, and other cells may have different mutations). However, the expression level of GABRA2 protein and mRNA in the brain tissue of the founder mouse was equivalent to strains without the Gabra2 deletion (e.g., ByJ and N); Figures 3B,C indicating correction in most, if not all, neurons of the founder mouse. In addition, no off-target modifications were detected in the top 15 predicted off-target sites in the F0 mouse and his progeny compared to the B6J reference genome (see section “Materials and Methods” and Supplementary Table S2).

Intronic Gabra2 Deletion Is Associated With Alterations in GABAergic Striatal Gene Expression and Signaling

To precisely identify biological processes or pathways altered by variation in Gabra2 levels in the striatum, hippocampus, and cortex, we performed enrichment analysis on genes differentially expressed \((p < 0.01; 684 \text{ genes})\) between homozygous \(Gabra2^{K1/K1}\) mice with “normal” Gabra2 levels and “mutant” \(Gabra2^{B6}/B6\) mice with low levels of Gabra2. No significantly enriched categories were observed for differentially expressed genes from the cortex or hippocampus. However, a number of enriched categories (FDR adjusted \(p\)-value <0.05) related to GABAergic synaptic signaling were detected in striatum, including ion transport, localization to the synapse, GABA receptor activity, retrograde endocannabinoid receptor signaling, dopaminergic synapse, cAMP signaling pathway, and GABAergic synapse (Supplementary Tables S3, S4). Mammalian Phenotype Ontology terms related to emotional behavior, cognition, anxiety, sensorimotor gating, fear, and reward were also enriched, albeit only at a nominally significant level. These results indicated that alterations in GABAergic signaling, specifically in the striatum, might be a direct downstream consequence of reduced Gabra2 levels.

Alterations in GABAergic signaling could result from the loss of Gabra2-containing receptors or changes in the composition of GABA-A receptors to reflect inclusion of other alpha subunits in the absence of GABRA2. To begin to address these issues, we compared the transcript levels of all major GABA-A receptor subunits (\(\alpha 1-6; \beta 1-3, \gamma 1-3, \text{and} \delta\)) in multiple brain regions between homozygous \(Gabra2^{K1/K1}\) mice with normal Gabra2 levels and mice harboring the mutant B6J allele (\(Gabra2^{B6}/B6\)). In addition to a large and significant reduction in Gabra2 mRNA, the transcript levels of several major GABA-A receptor alpha subunits (\(Gabra1, Gabra3, Gabra5\)) were also significantly reduced in striatum of \(Gabra2^{B6}/B6\) mice relative to \(Gabra2^{K1/K1}\) mice (Figure 4). In contrast, the level of Gabra4 was modestly increased in \(Gabra2^{B6}/B6\) mice. In striatum, alterations in alpha subunits were accompanied by significant decreases in beta (1 and 2) and gamma (1 and 2) subunits and a significant increase in the delta subunit (B6J allele relative to KI allele;
**FIGURE 3** | Repair of the private deletion in B6J is sufficient to restore GABRA2 protein and mRNA levels. (A) Site of repair of the private deletion in B6J (green), guide RNA (red) and PAM site (blue) are shown. The exon is highlighted in yellow. (B) Protein levels of GABRA2 are also restored in the hippocampus (top panel) and cortex (bottom panel) of the original Gabra2\(^{K1/K1}\) founder mouse (\(n = 1\)) relative to B6J (\(n = 3\)) and are similar to that of B6NJ (\(n = 3\)). Western blot shown to the left and adjusted density bar graph shown to the right for each region. (C) Hippocampal and cortical Gabra2 mRNA levels were restored in the original Gabra2\(^{K1/K1}\) founder after insertion of a single nucleotide relative to B6J. Expression of Gabra2 in the founder mouse was similar to that of other B6 substrains. Expression was measured via qPCR and shown relative to B6J. Males and females were combined for the statistical analysis (\(n = 5\) to 6 per strain). Note that CRISPR founders can often be mosaic for the introduced mutation such that some cells may have the mutation, some may not, and other cells may have different mutations. We include molecular validation in the founder in order to demonstrate that our engineering approach worked, and that mosaicism is not evident in brain tissue of the Gabra2\(^{K1/K1}\) founder mouse. Tissue from Gabra2 knockout mice (KO, predominately B6J background mixed with 129/SvEv, see section Materials and Methods) was kindly provided by Y. Blednov from The University of Texas and used as a negative control. (D) Hippocampal, striatal, and cortical Gabra2 mRNA levels were also restored in heterozygous Gabra2\(^{K1/K0}\) and homozygous knockin Gabra2\(^{K1/K1}\) mice after insertion of a single nucleotide relative to homozygous Gabra2\(^{B6J/B6J}\) mice. Expression measured via qPCR and data were analyzed at the A.Ct level (relative to the expression of the control gene Cyclin D). There is a significant main effect of genotype \([F(2,32) = 47.1, p = 2.9e-10]\) and brain region \([F(2,32) = 27.89, p = 9.7e-08]\). The interaction was not significant. Pairwise significance between genotypes and within brain region were determined using Tukey’s HSD test. Significance defined as: #\(p < 0.1\), *\(p < 0.05\), **\(p < 0.01\), ***\(p < 0.001\).

**DISCUSSION**

Using a combination of genetic, genomic, functional, and bioinformatics approaches we have identified a private and non-coding single nucleotide deletion in the Gabra2 gene in the B6J strain that results in substantial reduction of brain mRNA and protein levels (Figures 1, 2). Repair of this deletion using CRISPR-Cas9 genome editing fully restores GABRA2 levels (Figure 3) on the B6J genetic background. The resulting genetically engineered rodent model system (Gabra2\(^{B6J/B6J}\) and Gabra2\(^{K1/K1}\) mice) is thus, a valuable resource for probing the molecular and behavioral consequences of alterations in GABRA2 expression. These results have

Figure 4). Highly significant alterations in alpha subunits were not observed in the hippocampus or cortex between genotypes, but there was a trend for increased expression of both the α1 subunit in the cortex and the α3 subunit in the hippocampus (B6J allele relative to K1 allele). In the cortex there was an associated trend for increased expression of the β2 subunit and decreased expression of the γ1 subunit. In the hippocampus, there was a significant decrease in γ1 subunit levels and a trend for decreased expression of the β1 and γ3 subunits. Taken together, these results indicate that decreased levels of GABRA2 associated with the B6J allele are likely to be associated with profound alterations in the level and composition of GABA-A receptor subunits, especially in the striatum.

**FIGURE 4** | Highly significant alterations in alpha subunits were not observed in the hippocampus or cortex between genotypes, but there was a trend for increased expression of both the α1 subunit in the cortex and the α3 subunit in the hippocampus (B6J allele relative to K1 allele). In the cortex there was an associated trend for increased expression of the β2 subunit and decreased expression of the γ1 subunit. In the hippocampus, there was a significant decrease in γ1 subunit levels and a trend for decreased expression of the β1 and γ3 subunits. Taken together, these results indicate that decreased levels of GABRA2 associated with the B6J allele are likely to be associated with profound alterations in the level and composition of GABA-A receptor subunits, especially in the striatum.
important implications for any investigators using B6J as their background strain for molecular genetic studies of neurobiology and behavior, in particular those involving study of the GABAergic system.

Unlike previous studies in which function of Gabra2 was probed through complete genetic deletion, the mutation associated with the B6J allele of Gabra2 is hypomorphic and not a complete null. Complete deletion of function represents an extreme case in biological systems and there are often many adaptive changes that can make interpretation of results difficult. Substitution by other α subunits has not been profiled extensively in Gabra2 KO mouse lines and some differences in behavior could be due to enhanced expression of other subunits and differences in genetic background because existing Gabra2 KO mice differ in the proportion of B6J and 129/SvEv genetic background. In contrast, our study probed the role of natural variation in the level of Gabra2 in the context of an isogenic B6J background and profiled the expression of all major GABA-A subunits in cortex, striatum, and hippocampus—regions involved in different aspects of memory, mood disorders and addiction. Surprisingly, even in the context of an isogenic B6J background, we detected many alterations in GABAergic signaling at the transcriptional level and observed changes in GABA-A receptor subunit expression between Gabra2^{B6J/B6J} and Gabra2^{K1/K1} in all three regions with the most profound alterations detected in the striatum. Regardless of brain region, the major alteration (fold change > 2.5) was the reduction in Gabra2 expression. However, the levels of other alpha subunits were also altered (Figure 4), albeit at much lower levels (<1.5 fold). Exceptions include the levels of Gabra5 (1.7 fold decrease relative to Gabra2^{K1/K1}) and Gabra3 (2.5 fold decrease relative to Gabra2^{K1/K1}) in the striatum. There are no known mutations in Gabra3 or Gabra5 in the B6J background, so the reduction in their levels in striatum is almost certainly a direct result of the reduction in Gabra2. Based on the available data, our hypothesis is that there are simply fewer α2 containing GABA-A receptors in B6J derived strains, and possibly fewer GABA-A receptors in some regions. This is supported by the trend toward decreased cortical α2:γ1 transcript levels, decreased hippocampal α2:β1:γ1 and α2:β1:γ3 transcript levels, and decreases in multiple alpha, beta, and gamma transcript levels in the striatum of Gabra2^{B6J/B6J} relative to Gabra2^{K1/K1}. The pronounced impact of the B6J mutation on GABA-A receptor subunit mRNA expression in the striatum, combined with the involvement of the striatum in anxiety and behavioral sensitization to alcohol and other drugs of abuse, strongly suggests that alterations in the level of Gabra2 in this region drive related behavioral differences between mice with the low-expressing B6J allele and the high-expressing wild type allele. Future functional studies exploring behavior and inhibitory neurotransmission across multiple brain regions in B6-Gabra2^{em1Geh} / J genotypes will test this hypothesis.

Overall, the findings of our study have broad implications for both mouse and human research. Despite numerous associations between human GABRA2 variants and addiction-related traits, the mechanisms underlying these associations are unclear. To date, there is no direct link between variants in GABRA2 and alterations in gene expression or function in human brain. However, if functional variants in GABRA2 exist, they are almost certainly non-coding variants that exert a functional impact through alteration of splicing or gene regulation, or mRNA stability as opposed to variants resulting in a complete loss of function. This may be similar to the function
of the non-coding variant controlling variation in *Gabra2* levels in crosses involving B6J and in our B6-*Gabra2*^2em1Geh/J preclinical model system and suggests that non-coding variants in GABRA2 in human populations might be functional and causal. Evaluating and confirming the functional impact of these variants in human populations is extremely difficult due to the complexity of human genetics and environmental factors, as well as the limited ability to assay brain gene expression and circuitry. Therefore, preclinical models such as ours are of vital importance for generating new hypotheses, endophenotypes, and underlying molecular mechanisms to test in human populations. This might include the generation of humanized mice in which candidate functional GABRA2 variants are introduced into the *Gabra2* corrected B6J genetic background that is permissive for detection of variants that modulate *Gabra2* expression.

We discovered a functional variant causing reduced *Gabra2* expression in B6J, the mouse reference genome and the most widely used inbred strain in biomedical research. It is now critical to re-evaluate gene deletion and other studies in light of the naturally occurring mutation in *Gabra2*. For example, the *Gabra2* locus was recently posited to be a modifier of genetic deletion of Scn1a, a voltage-gated sodium channel gene implicated in a spectrum of seizure-related disorders in humans (Hawkins et al., 2016). Hawkins and colleagues observed that the genetic background of Scn1a+/− mutations impacted the severity of the epilepsy phenotype such that mutations generated on a B6J background had increased seizures and premature death compared to the same mutation on a 129S6/SvEvTac background, which had a normal phenotype. Our study suggests a specific hypothesis, namely that reduced GABRA2 expression and alterations in GABAergic signaling comprise the causal mechanism underlying enhanced seizure susceptibility induced by Scn1a deletion on the permissive B6J background. Likewise, because the B6J strain is also the most widely used strain for biomedical research, several findings in this strain may be confounded by alterations in GABRA2 and other GABA-A subunits. However, as long as investigators are aware of gene variants with a large impact on expression and function, such as the non-coding variant in *Gabra2*, these naturally occurring mutations can be vital tools for discovery and systems genetics (Li et al., 2010; Mulligan and Williams, 2015; Wang et al., 2016). The CRISPR-Cas9 B6J *Gabra2* repaired line (B6-*Gabra2*^2em1Geh/J) will serve as a crucial resource for directly testing the role of *Gabra2* variation related to phenotypic differences in B6J derived lines and substrains.

**AUTHOR CONTRIBUTIONS**

MM designed the studies, ran the experiments, performed the analysis, and wrote the manuscript. RW designed the studies and along with GH, CK, and CB provided intellectual support and helped in editing the manuscript. TA, CW, MTH, TS, and JI performed the animal and molecular work. CP and MH completed the data analysis. SN and CK provided the western blot analysis and reagents. LL, HT, XW, and JP provided resources for proteomics. GH provided the CRISPR engineered mice.

**FUNDING**

This research was partially supported by NIAAA INIA grants U01AA13499 and U01AA016662 to MM, LL, and RW, NIH grant R01AG047928 to JP and XW, NIH grants R21AG048446 to CK and F31AG050357 to SN, NIDA grant R01DA039168 and NCI grant R01CA221260 to CB, and NIAAA grants AA010422 and AA020889 to GH.

**ACKNOWLEDGMENTS**

We thank Carolyn Ferguson (University of Pittsburgh), Casey Bohl (University of Tennessee), and Melinda McCarty (University of Tennessee) for superb technical support.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.00188/full#supplementary-material
using marker rs13478320 located on Chr 5 at 70,742,059 bp. X Chromosome

databases; PSM#: the total number of peptide-spectrum matches for each protein. Gene Symbol:

samples from 3- and 12-month-old C57BL/6J and DBA/2J mice were used in this study. Three hippocampal samples were lysed and pooled into 6 groups (3 groups per strain). The samples were analyzed by 6-plex TMT-LC/LC-MS/MS. TMT report ion intensities were extracted and summarized for each protein. Gene Symbol: extracted from the UniProt database; protein accession: from the UniProt databases; PSM#: the total number of peptide-spectrum matches for each protein; peptide#: the total number of peptides identified for each protein; Relative protein abundance: the ratio of the ion intensity of each sample and the average of 6 samples; p-value: p value calculated by t statistics. A total of 7,014 unique proteins were quantified.

TABLE S4 | Summary of Predicted Off-Targets. Predictions made using the CRISPR program and prioritized based on CFD score.

TABLE S3 | Overrepresented functional terms for differentially expressed genes in the striatum of CRISPR engineered Gabra2 Ki mice. Gene members are shown for categories with less than 20 members. All gene members for each category are shown in Supplementary Table S4.

TABLE S4 | Summary of genes with significantly enriched functional terms. Summary of results are provided in the first worksheet (C = total number of genes in category; O = number of observed genes in category; E = number of genes in category expected by chance). Remaining worksheets contain all enriched categories and genes.

REFERENCES

Agrawal, A., Edenberg, H. J., Foroud, T., Bierut, L. J., Dunne, G., Hinrichs, A. L., et al. (2006). Association of GABRA2 with drug dependence in the collaborative study of the genetics of alcoholism sample. Behav. Genet. 36, 640–650. doi: 10.1007/s10519-006-9089-9

Arias, A. J., Covault, J., Feinn, R., Pond, T., Yang, B. Z., Ge, W., et al. (2014). A GABRA2 variant is associated with increased stimulation and ‘high’ following alcohol administration. Alcohol Alcohol. 49, 1–9. doi: 10.1093/alcalc/agt163

Bai, B., Tan, H., Pagala, V. P., Hendershot, L., et al. (2017). Deep profiling of proteome and phosphoproteome by isobaric labeling, extensive liquid chromatography, and mass spectrometry. Methods Enzymol. 585, 377–395. doi: 10.1016/bse.2016.10.007

Bassett, A. R., Tibbit, C., Ponting, C. P., and Liu, J.-L. (2014). Highly efficient targeted mutagenesis of Drosophila with the CRISPR/Cas9 system. Cell Rep. 4, 220–228. doi: 10.1016/j.celrep.2013.06.020

Blednov, Y. A., Benavidez, J. M., Black, M., Chandra, D., Rudolph, U., et al. (2013). Linking GABA(A) receptor subunits to alcohol-induced conditioned taste aversion and recovery from acute alcohol intoxication. Neuropharmacology 67, 46–56. doi: 10.1016/j.neuropharm.2012.10.016

Blednov, Y. A., Boorghese, C. M., McCracken, M. L., Benavidez, J. M., Geil, C. R., Osterndorf-Kahanek, E., et al. (2011). Loss of ethanol conditioned taste aversion and motor stimulation in knockin mice with ethanol-insensitive alpha2-containing GABA(A) receptors. J. Pharmacol. Exp. Ther. 336, 154–165. doi: 10.1124/jpet.110.171645

Blednov, Y. A., Rosahl, T. W., and Stephens, D. N. (2008). Targeted deletion of the Gabra2 gene encoding alpha2-subunits of GABA(A) receptors facilitates performance of a conditioned emotional response, and abolishes anxiolytic effects of benzodiazepines and barbiturates. Pharmacol. Biochem. Behav. 90, 1–8. doi: 10.1016/j.pbb.2008.01.015

Doench, J. G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E. W., Donovan, K. F., et al. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat. Biotechnol. 34, 184–191. doi: 10.1038/nbt.3437

Dixon, C. I., Rosahl, T. W., and Stephens, D. N. (2008). Targeted deletion of the Gabra2 gene encoding alpha2-subunits of GABA(A) receptors facilitates performance of a conditioned emotional response, and abolishes anxiolytic effects of benzodiazepines and barbiturates. Pharmacol. Biochem. Behav. 90, 1–8. doi: 10.1016/j.pbb.2008.01.015

Enoch, M. A. (2008). The role of GABA(A) receptors in the development of alcoholism. Pharmacol. Biochem. Behav. 90, 95–104. doi: 10.1016/j.pbb.2008.03.007

Enoch, M. A., Hodgkinson, C. A., Yuan, Q., Shen, P. H., Goldman, D., and Roy, A. (2010). The influence of GABA2, childhood trauma, and their interaction on alcohol, heroin, and cocaine dependence. Biol. Psychiatry 67, 20–27. doi: 10.1016/j.biopsych.2009.08.019

Enoch, M. A., Schwartz, L., Albaugh, B., Virkkunen, M., and Goldman, D. (2006). Dimensional anxiety mediates linkage of Gabra2 haplotypes with alcoholism. Am. J. Med. Genet. B Neuropsychiatr. Genet. 141B, 599–607. doi: 10.1002/ajmg.b.30336

Fehr, C., Sander, T., Tadic, A., Lenzen, K. P., Anghelescu, I., Klawe, C., et al. (2006). Confirmation of association of the GABA2 gene with alcohol dependence by subtype-specific analysis. Psychiatr. Genet. 16, 9–17. doi: 10.1097/01.ygg.0000185027.89816.d9

Gibbs, J. R., van der Brug, M. P., Hernandez, D. G., Traynor, B. J., Nalls, M. A., Laird, N. S., et al. (2010). Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. PLoS Genet. 6.e1000952. doi: 10.1371/journal.pgen.1000952

Haesler, M., Schonig, K., Eckert, H., Eschstruth, A., Miane, J., Reuda, J. B., et al. (2016). Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. Genome Biol. 17:148. doi: 10.1186/s13059-016-1012-2

Hawkins, N. A., Zachwieja, N. J., Miller, A. R., Anderson, L. L., and Kearney, J. A. (2016). Fine Mapping of a dravet syndrome modifier locus on mouse chromosome 5 and candidate gene analysis by RNA-seq. PLoS Genet. 12.e1006398. doi: 10.1371/journal.pgen.1006398

Heinzen, E. L., Gerson, D., Cronin, K. D., Maia, J. M., Shianna, K. V., Gabriel, B. W., et al. (2010). Tissue-specific genetic control of splicing: implications for the study of complex traits. PLoS Biol. 6:e1. doi: 10.1371/journal.pbio.1000913

Hitzemann, R., Reed, C., Malmberg, B., Lawler, M., Hitzemann, B., Cunningham, B., et al. (2004). On the integration of alcohol-related quantitative
Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression. 

Frontiers in Genetics | www.frontiersin.org 14 March 2019 | Volume 10 | Article 188

Mulligan, K. M., and Williams, R. W. (2012). Sex-specific modulation of gene expression networks in murine hypothalamus. Front. Neurosci. 6(3), doi: 10.3389/fnins.2012.00063

Mulligan, M. K., Wang, X., Adler, A. L., Mozhui, K., Lu, L., and Williams, R. W. (2012). Complex control of GABA(A) receptor subunit mRNA expression, variation, coexpression, and genetic regulation. PLoS One 7:e34586. doi: 10.1371/journal.pone.0034586

Mulligan, M. K., and Williams, R. W. (2015). Systems genetics of behavior: a prelude. Curr. Opin. Behav. Sci. 2, 108–115. doi: 10.1016/j.cobeha.2015.01.014

Meyers, A. J., Gibbs, J. R., Webster, J. A., Rohrer, K., Zhan, A., Marlowe, L., et al. (2007). A survey of genetic human cortical gene expression. Nat. Genet. 39, 1494–1499. doi: 10.1038/ng.2007.16

Niu, M., Cho, J. H., Kodali, K., Pagala, V., High, A. A., Wang, H., et al. (2017). Extensive peptide fractionation and y1 ion-based interference detection method for enabling accurate quantification by isobaric labeling and mass spectrometry. Anal. Chem. 89, 2956–2963. doi: 10.1021/acs.analchem.6b04415

Olsen, R. W., and Sieghart, W. (2009). GABA A receptors: subtypes provide diversity of function and pharmacology. Neuropharmacology 56, 141–148. doi: 10.1016/j.neuropharm.2008.07.045

Orozco, L. D., Cokus, S. J., Ghazalpour, A., Ingram-Drake, L., Wang, S., van Nas, A., et al. (2009). Copy number variation influences gene expression and metabolic traits in human. Hum. Mol. Genet. 18, 4118–4129. doi: 10.1093/hmg/ddp360

Overall, R. W., Kempermann, G., Peirce, J., Lu, L., Goldowitz, D., Gage, F. H., et al. (2009). Genetics of the hippocampal transcriptome in mouse: a systematic survey and online neurogenomics resource. Front. Neurosci. 3:55. doi: 10.3389/.

Pierucci-Lagha, A., Covault, J., Feinn, R., Nellistery, M., Hernandez-Avila, C., Oncken, C., et al. (2005). GABRA2 alleles moderate the subjective effects of alcohol, which are attenuated by finasteride. Neuropharmacology 30, 1193–1203. doi: 10.1016/sij/.

Ployn, J. F., and Sieghart, W. (2009). GABA A receptors: subtypes provide diversity of function and pharmacology. Neuropharmacology 56, 141–148. doi: 10.1016/j.neuropharm.2008.07.045

Schoen, T. D., and Livak, K. J. (2008). Analyzing real-time PCR data by the comparative C(T) method. Nat. Protoc. 3, 1108–1108. doi: 10.1038/nprot.

Sieghart, W. (1995). Structure and pharmacology of gamma-aminobutyric acidA receptor subtypes. Pharmacol. Rev. 47, 181–234.

Soyka, M., Preuss, U. W., Hesselbrock, V., Zill, P., Koller, G., and Bondy, B. (2008). GABA-A2 receptor subunit gene (GABRA2) polymorphisms and risk for alcohol dependence. J. Psychiatr. Res. 42, 184–191. doi: 10.1016/j.

Taylor, B. A., Wnek, C., Kotlus, B. S., Roemer, N., MacTaggart, T., and Phillips, S. J. (1999). Genotyping new BXD recombinant inbred lines from advanced intercrosses in mice. BMC Genet. 5:7. doi: 10.1186/1471-2156-5-7

Villafuerte, S., Strumba, V., Stoltenberg, S. F., Zucker, R. A., and Burmeister, M. (2013). Impulsiveness mediates the association between GABRA2 SNPs and lifetime alcohol problems. Genes Brain Behav. 12, 525–531. doi: 10.1111/gbb.12039

Vollenweider, I., Smith, K. S., Krist, R., and Rudolph, U. (2011). Antidepressant-like properties of alpha2-containing GABA(A) receptors. Behav. Brain Res. 217, 77–80. doi: 10.1016/j.bbr.2010.10.009

trait loci and gene expression analyses. Alcohol Clin. Exp. Res. 28, 1437–1448. doi: 10.1097/01.ALC.0000139827.86749.DA

Hodgkins, A., Farne, A., Perera, S., Gregory, T., Parry-Smith, D. J., Skarnes, W. C., et al. (2015). WGE: a CRISPR database for genome engineering. Bioinformatics 31, 3078–3080. doi: 10.1093/bioinformatics/btv308

International League Against Epilepsy Consortium on Complex Epilepsies (2014). Genetic determinants of common epilepsies: a meta-analysis of genome-wise association studies. Lancet Neuro. 13, 893–903. doi: 10.1016/S1474-4221(14)

70171–1

Kim, Y., Xia, K., Tao, R., Giusti-Rodriguez, P., Vladimirov, V., van den Oord, E., et al. (2014). A meta-analysis of gene expression quantitative trait loci in brain. Transl. Psychiatry 5:e459. doi: 10.1038/tp.2014.96
Wang, J., Duncan, D., Shi, Z., and Zhang, B. (2013). WEB-based gene set analysis toolkit (WebGestalt): update 2013. *Nucleic Acids Res.* 41, W77–W83. doi: 10.1093/nar/gkt439

Wang, X., Pandey, A. K., Mulligan, M. K., Williams, E. G., Mozhui, K., Li, Z., et al. (2016). Joint mouse-human phenome-wide association to test gene function and disease risk. *Nat. Commun.* 7:10464. doi: 10.1038/ncomms10464

Yang, H., Wang, H., and Jaenisch, R. (2014). Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. *Nat. Protoc.* 9, 1956–1968. doi: 10.1038/nprot.2014.134

Zhang, B., Kirov, S., and Snoddy, J. (2005). WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Res.* 33, W741–W748. doi: 10.1093/nar/gki475

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Mulligan, Abreo, Neuner, Parks, Watkins, Houseal, Shapaker, Hook, Tan, Wang, Ingels, Peng, Lu, Kaczorowski, Bryant, Homanics and Williams. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.