Genetic variation and sex differences are missed opportunities for addiction biology.

Authors:

Michael C. Saul[1,2], Jared R. Bagley[2,3], Lauren S. Bailey[2,3], Udita Datta[1,2], Price E. Dickson[1,2], Rainy Dodd[1,2], Leona H. Gagnon[1,2], Violet M. Kimble[4], Michael Leonardo[1,2], Sam-Moon Kim[2,5], Ashley Olson[1,2], Tyler Roy[1,2], Sarah A. Schoenrock[2,6], Troy Wilcox[1,2], J. David Jentsch[2,3], Ryan W. Logan[2,5], Colleen A. McClung[2,5], Vivek M. Philip[1,2], Laura G. Reinholdt[1,2], Stacey J. Sukoff Rizzo[1,2,7], Lisa M. Tarantino[2,6,8], and Elissa J. Chesler[1,2,*].

Affiliations: [1]The Jackson Laboratory of Mammalian Genetics, Bar Harbor, ME USA; [2]Center for Systems Neurogenetics of Addiction at The Jackson Laboratory, Bar Harbor, ME USA; [3]Binghamton University Department of Psychology, Binghamton, NY USA; [4]Drew University Department of Neuroscience, Madison, NJ USA; [5]University of Pittsburgh School of Medicine Department of Psychiatry, Pittsburgh, PA USA; [6]University of North Carolina-Chapel Hill Department of Genetics, Chapel Hill, NC USA; [7]University of Pittsburgh School of Medicine Department of Neurobiology, Pittsburgh, PA USA; [8]University of North Carolina-Chapel Hill Eshelman School of Pharmacy Division of Pharmacotherapy and Experimental Therapeutics, Chapel Hill, NC USA.

[*]To whom correspondence should be addressed: elissa.chesler@jax.org

Keywords: addiction, sex differences, genetics, cocaine, genomics.

Abstract

Though risk for cocaine use disorder is subject to substantial inter-individual variation, the sources of that variation – including, genetics and sex – are too often ignored in studies of this phenomenon. In genetically diverse laboratory mice, these individual differences explain the majority of variance in important cocaine-related behavioral, physiological, and striatum transcriptional responses traits. This individual variation represents a missed opportunity for discovery and translation of addiction mechanisms.

Main Text

Addictions are highly prevalent, leading to a public health crisis associated with substantial morbidity and mortality. Illicit substance use disorders afflict 1 in 14 young adults in the United States[1] and drug overdoses are now the leading cause of accidental death among American adults under 55[2]. Genetic variation and sex differences are both known to influence addiction vulnerability; cocaine use disorder is highly heritable ($H^2 = 0.71$)[3] and substance use behaviors show sex differences in both humans and other animals[4,5]. Consequently, the neurobiology underlying addiction cannot be understood completely without consideration of genetic background and sex.

Behavioral and genomic studies of addiction-related phenotypes often ignore individual differences by studying only males of a single inbred mouse strain[6] or outbred rat populations confounded by vendor[7]. These experimental methodologies ignore sex differences and genetic variation without reducing experimental noise[8,9]. Though rodent genetic variation may not capture precise human variants, it can be used to determine underlying mechanisms in addiction-relevant processes. Thus, inclusion of genetic variability and both males and females in rodent studies has the power to identify addiction-relevant targets, delineate variants' effects on co-regulation and co-expression networks, and define druggable network nodes. As a further benefit, genotypic and phenotypic precision allows genetics studies in rodents to be performed at orders of magnitude lower cost than human GWAS.

To assess genetics, sex, and their interaction on cocaine-related phenotypes in mice, we undertook a large-scale comprehensive evaluation of behavioral, physiological, and brain transcriptomic measures in both male and female mice from the eight inbred founder strains of the Diversity Outbred (DO) mouse heterogenous stock[10]. Together, these strains capture approximately 90% of the genetic diversity in *Mus musculus*[11]. We examined behavioral and physiological correlates of future cocaine use – including multiple novelty response behaviors, circadian rhythm phenotypes observed in cells, and reversal learning as a measure of reward learning and impulsivity. We directly measured cocaine-related behaviors such as initial locomotor sensitivity to cocaine and intravenous self-administration (IVSA) of cocaine. Finally,

we measured the striatum transcriptome response to cocaine in all eight founder strains with RNAseq. Our work represents the first time that many of these behavioral, physiological, and molecular traits have been studied using methods powered to detect sex and genotype effects and the extent to which these interact by moderating one another.

Response to novelty predicts psychostimulant addiction-related phenotypes in both humans[12] and mice[13]. We first assessed differences in behavioral traits related to exploration and response to novelty. These behavioral risk measures were moderately to strongly heritable among these strains (H^2 = 0.16-0.55, **Figure 1A-D**, **Supplementary Table S1**). Further, for the phenotype of total nose pokes in the hole board, there was a significant sex difference ($F_{1,597}$ = 10.69, p = 0.0014). We detected strain-by-sex interactions in the total number of hole board nose pokes ($F_{7,597}$ = 2.74, p = 0.0083) and in the proportion of distance traveled in the center of the open field ($F_{7,684}$ = 2.09, p = 0.043), indicating that sex differences in these novelty response traits are moderated by genetic background. Total nose pokes in the hole board (**Figure 1A**) and novel place preference (**Figure 1B**) showed the largest differences in wild-derived strains, demonstrating the importance of their expanded phenotypic range[14]. Transitions between the sides of the light-dark box also showed extreme phenotypes in wild-derived mouse strains (**Figure 1D**).

Circadian rhythm and reward-related behavioral phenotypes are co-inherited and rhythm disruptions are linked to development and progression of substance use disorders[15]. The molecular clock system directly influences expression of dopamine receptors in the striatum[16]. A cell-based assay on fibroblasts derived from each of the founder strains in which a *Bmal1-dLuc* reporter was utilized for circadian measurement of luciferase bioluminescence to assess differences in circadian rhythmicity[17,18]. We found very high heritability of the amplitude of these rhythmic patterns (H^2 = 0.59, **Figure 1E**), but no significant sex differences or strain x sex interactions. These results suggest that genetic differences in the molecular clock are one potential mechanism for individual differences in addiction-related phenotypes.

Reversal learning tasks evaluate impulsive and compulsive behaviors that predict addiction liability[19]. Within reversal learning paradigms, one measure of impulsivity is the number of premature responses produced during the reversal phase of the task. Premature reversal

responses in these showed moderate heritability across this panel of mice (H^2 = 0.29, **Figure 1F**), but no significant sex differences or strain x sex interactions. We observed the lowest rate of premature responding in the wild-derived PWK/PhJ strain.

We next sought to assess individual differences in initial locomotor response to cocaine. Individual differences in initial sensitivity to cocaine predict subsequent drug use in humans[20]. In these diverse mouse strains, initial cocaine sensitivity showed strong heritability (H^2 = 0.48 for initial sensitivity, **Figure 1G**). Wild-derived strains, particularly WSB/EiJ and PWK/PhJ, exhibited the highest initial sensitivity to cocaine.

Operant drug self-administration procedures directly quantify reinforced responding for drug[13]. Heritability of cocaine IVSA traits in the founders of the DO was very strong (H^2 = 0.47 for sessions to acquisition of IVSA; H^2 = 0.60 for total infusions at FR-1 1.0 mg/kg, **Figure 1H-I**). The high heritability of infusions earned is quite similar to the observed heritability of human cocaine use disorder[3]. Of note, 129S1/SvImJ mice do not acquire IVSA and do not take any cocaine during the acquisition phase. The largest phenotypic range manifests in wild-derived mice; PWK/PhJ mice acquire IVSA behavior very quickly and self-administer the most cocaine – about four times as much as C57BL/6J – while, of strains that acquire IVSA behavior, WSB/EiJ mice self-administer the least amount of cocaine. The broad phenotypic range of volitional cocaine-taking behaviors in these genetically diverse mouse strains represents a clear opportunity to study mechanisms underlying variation in the physiology of addiction.

Because addiction-related phenotypes are highly heritable and sometimes exhibit sex differences in a manner dependent on genetics, we next sought to assess heritable differences in the molecular response to cocaine. Male and female mice of the eight DO founder strains were given multiple injections of cocaine or saline over 19 days. Samples for bulk RNAseq were collected from striatum at least 24 hours after the final injection of repeated administration of either cocaine or vehicle (sham). First, we documented that in the absence of cocaine, most transcripts show moderate to high baseline heritability across the founder strains (median H^2 = 0.29) and that subspecies of origin explains the greatest amount of variation in expression (**Figure 2A**). Using a linear modeling approach[21,22], we found few expression differences attributable solely to the effect of cocaine treatment. Instead, significant effects of cocaine

arose in interaction with individual differences such as genetic background. The greatest quantity of differentially expressed transcripts was observed in the strain-by-sex-by-drug treatment three-way interaction (1,282 genes at $q < 0.01$, **Figure 2B, Supplementary Table S2**). Some strains responded more strongly than others; 129S1/SvImJ and PWK/PhJ had many expressed genes influenced by cocaine and sex-by-cocaine interactions (PWK/PhJ: 89 genes at $q < 0.01$, **Figure 2C**) while the commonly used C57BL/6J strain had very few (4 genes at $q < 0.01$, **Figure 2C**). Because statistical power was approximately equal for all strains involved in this study, these differences likely reflect real individual differences in the brain's sex-specific responses to cocaine.

Many significant effects of cocaine on gene expression would have been missed in had only single strain been analyzed. For instance, we identified differential expression of the serotonin receptor 1D gene *Htr1d* in the strain-by-sex-by-cocaine interaction factor. This gene is upregulated in males of two strains but downregulated in males of another (**Figure 2D**).

Some of these heritable transcriptome differences we observed recapitulate previous work showing heritable differences in pharmacokinetics of cocaine[23] – for example, mice closely related to the high cocaine taking PWK/PhJ strain are known to have some of the highest brain concentrations of cocaine shortly after injection. These pharmacokinetic differences do not account for the strong differences between strains such as 129S1/SvImJ and A/J, whose brain cocaine pharmacokinetic profiles are very similar[23]. Further research on highly diverse mice may resolve the mechanisms driving individual differences between strains with closely matched pharmacokinetic profiles.

Genetic variation is a valuable resource for the discovery of biological mechanisms of addiction[24,25]. As in humans, individual differences among mice greatly influence behavioral, physiological, and transcriptomic cocaine-related traits. For some of these traits, individual differences explain more than half of the variation. This variation in addiction related traits is a largely untapped resource that would have otherwise remained unknown had we only examined a single inbred strain. Because the vast majority of the biomedical literature is built on limited genetic diversity, such individual differences represent a missed opportunity to find new, clinically relevant genes and mechanisms of addiction.

Online Methods

Standard Operating Procedures

All methodologies used in this work are documented in depth in the Center for Systems Neurogenetics of Addiction's Standard Operating Procedures (SOPs) (https://www.jax.org/research-and-faculty/research-centers/systems-neurogenetics/data-resources). The brief methods appearing below summarize these SOPs and reference specific SOPs.

Animals

Mice from the following strains were surveyed in these experiments: A/J (JAX stock #000646), C57BL/6J (JAX stock #000664), 129S1/SvImJ (JAX stock #002448), NOD/ShiLtJ (JAX stock #001976), NZO/HlLtJ (JAX stock #002105), CAST/EiJ (JAX stock #000928), PWK/PhJ (JAX stock #003715), and WSB/EiJ (JAX stock #001145). These strains are the founders of the DO heterogeneous stock and Collaborative Cross recombinant inbred strains. Surveys of these eight strains can demonstrate statistical heritability patterns that justify further dissection using the derived resources[10]. The mice used in the Research Animal Facility at The Jackson Laboratory came from breeding colonies maintained in the Research Animal Facility. These colonies were derived from production colonies at The Jackson Laboratory and breeders were replaced with animals from The Jackson Laboratory's production colony at least every five generations. The mice used in the Jentsch Lab at Binghamton University were shipped to the Jentsch Lab from either the Chesler Lab colonies or from production colonies in The Jackson Laboratory. The studies described utilized a total of 1,085 mice (**Supplementary Table S3**).

Mouse Husbandry and Housing

All procedures were approved by the Jackson Laboratory of Mammalian Genetics institutional animal care and use committee. Mice were housed according to the CSNA animal housing SOP (https://www.jax.org/-/media/jaxweb/files/research-and-faculty/tools-and-resources/system-neurogenetics/csna-animal-housing.pdf).

Cocaine

Cocaine hydrochloride was provided by the National Institute on Drug Abuse Drug Supply Program Division of Therapeutics and Medical Consequences (catalog number: 9041-001). Cocaine was stored in powder form until it was formulated into 0.9% Saline (100 mg/mL clear solution) in various concentrations specific to each experiment according to their individual SOPs.

Novelty Response Behavioral Phenotypes

Open field, light-dark box, hole board, and novel place preference behavioral paradigms were conducted in order with one test per day on consecutive days during the light phase of the light:dark cycle. Open field data were collected for 60 minutes according to the SOP (https://www.jax.org/-/media/jaxweb/files/research-and-faculty/tools-and-resources/system-neurogenetics/open-field-assay.pdf?la=en&hash=32DDAFF2B17B2D4961C136C5616C4982AC23EC3B). Light-dark data were collected for 20 minutes with the mouse starting in the light side of the chamber facing the dark side according to the SOP (https://www.jax.org/-/media/jaxweb/files/research-and-faculty/tools-and-resources/system-neurogenetics/light-dark-assay.pdf?la=en&hash=A63CF8D22EB7936CF6C69A3178373981F4016675). Hole board data were collected for 20 minutes according to the SOP (https://www.jax.org/-/media/jaxweb/files/research-and-faculty/tools-and-resources/system-neurogenetics/hole-board-assay.pdf?la=en&hash=EC343A797D37209CF64D34E6031608A511D8E15D). Novel place preference included a five minute acclimation period to a center chamber, a 10 minute exposure period to a randomized exposure side, a five minute acclimation period, then testing for 20 minutes for preference for novel or exposure side of the novel place preference apparatus according to the SOP (https://www.jax.org/-/media/jaxweb/files/research-and-faculty/tools-and-resources/system-neurogenetics/novelty-place-preference-assay.pdf?la=en&hash=B5D2D0FC9028B408E84729C0C8832C580AB8E039). All mice from the Center for Systems Neurogenetics of Addiction were tested through this novelty pipeline prior to any other test and were then randomized and assigned into either reversal learning, cocaine

locomotor sensitization, or cocaine intravenous self-administration. The novelty study produced observations from a total of 783 mice.

Bmal1-dLuc Circadian Rhythm Data

Data for circadian rhythm were measured in primary fibroblast cultures generated from skin biopsies in the founders. To isolate fibroblasts, ear biopsies (1 mm in diameter) were digested in Dulbecco's Modified Eagle's Medium (DMEM, HyClone) containing 2.5 mg/ml collagenase D (Gibco) and 1.25 mg/ml pronase (Millipore) for 90 mins and then plated in DMEM growth media containing 10% Fetal Bovine Serum (FBS, HyClone), 292 μg/ml L-glutamine (HyClone), 100 units/ml penicillin (Hyclone) and 100 μg/ml streptomycine (HyClone). *Bmal1-dLuc* reporter was delivered to fibroblasts by lentiviral-mediated gene delivery (VectorBuilder). Following synchronization of rhythms by 15 μM forskolin (Sigma) for 2 hrs, the temporal patterns of *Bmal1-dLuc* bioluminescence was recorded for ~70 secs at intervals of 10 mins over 6-7 days from fibroblast cultures in DMEM recording media containing 15 μM forskolin, 25 mM HEPES (Gibco), 292 μg/ml L-glutamine, 100 units/ml penicillin, 100 μg/ml streptomycine, and 10μM luciferin (Promega) by an automated 32-channel luminometer (Lumicycle, ActiMetrics) in a standard tissue culture incubator at 32°C. The amplitude of bioluminescence rhythms was determined from baseline-subtracted data using the damped sine fit and Levenberg-Marquardt algorithm[26]. The circadian study produced observations from a total of 205 mice.

Reversal Learning

Data for reversal learning were collected using the SOP (https://www.jax.org/-/media/jaxweb/files/research-and-faculty/tools-and-resources/system-neurogenetics/reversal-learning-assay.pdf?la=en&hash=8484E47B170462960E11C1FAEEE6FF3CE6FDFC08). The reversal learning data produced observations from a total of 202 mice.

Initial Locomotor Sensitivity to Cocaine

Data for initial locomotor sensitivity were collected as described previously[27] using data from days 1-3 in the SOP for locomotor behavioral sensitization to cocaine (https://www.jax.org/-/media/jaxweb/files/research-and-faculty/tools-and-resources/system-neurogenetics/cocaine-

locomotor-sensitization-assay.pdf?la=en&hash=9E5D4C248C3BCCAD947C164AE81663C13A77EB0D). Briefly, mice were placed into the open field arena for 30 minutes, removed, and injected i.p. with either saline (days 1-2) or 10 mg/kg cocaine (day 3) and returned to the open field arena for 60 minutes. Distance moved after injection on day 3 minus day 2 was uses as a measure of initial locomotor sensitivity to cocaine. The sensitization study produced observations from a total of 230 mice.

Cocaine Intravenous Self-Administration

Prior to cocaine intravenous self-administration, mice were implanted with a jugular catheter and allowed 10 days for post-operative recovery. In an operant conditioning paradigm, mice were allowed to acquire cocaine self-administration at 1.0 mg/kg, then evaluated for dose-response effects at eight different doses. After a final stabilizing dose at 1.8 mg/kg, responses during seven days of withdrawal were recorded. Finally, cued reinstatement was recorded for two days. Self-administration in these eight mouse strains was performed according to v1.0 of the CSNA's SOP (https://www.jax.org/-/media/jaxweb/files/research-and-faculty/tools-and-resources/system-neurogenetics/intravenous-self-administration-ivsa-paradigm.pdf?la=en&hash=FA64135F219C7DF65937A1CF9270301B0E771836). The intravenous self-administration study produced observations from a total of 217 mice.

Data Deposit

Data for each phenotype will be deposited in the Mouse Phenome Database (MPD)[28] upon publication.

Heritability Calculations

For each trait, heritability was calculated from linear models using the isogenic strain as the independent categorical variable using the following equation:

$$h^2 = \frac{MS_{strain}}{MS_{strain} + (n_{mean} - 1) * MS_{resid}}$$

where MS_{strain} is the mean square of the strain effect, n_{mean} is the mean number of samples within each strain, and MS_{resid} is the mean square of the residuals. For the reversal learning

data, an additional additive covariate of site was included in the model to account for inter-lab variation. This term was not utilized in the heritability calculation.

For some traits such as number of infusions at FR-1 1.0 mg/kg cocaine self-administration, a single strain such as 129S1/SvImJ showed little to no variation, which may upwardly bias heritability calculations. For these traits, heritability was calculated both with and without the low variance strain. The data reported in the paper rely upon the inclusive calculation, but results of both methods of calculation are reported for completeness (see **Supplementary Table S1**).

RNAseq

Striatum tissue was collected during the light stage of the light:dark cycle between 24 and 48 hours after the final injection in the cocaine behavioral sensitization protocol according to the SOP (https://www.jax.org/-/media/jaxweb/files/research-and-faculty/tools-and-resources/system-neurogenetics/post-sensitization-tissue-collection.pdf?la=en&hash=9E6CD8DEB39606B791A5D25F6CD0611EF14D96A7). Tissue was collected for both sexes of each founder strain exposed to either sham (saline) or 10 mg/kg IP cocaine.

RNA was isolated from striatum tissue using the MagMAX mirVana Total RNA Isolation Kit (ThermoFisher) and the KingFisher Flex purification system (ThermoFisher). Tissues were lysed and homogenized in TRIzol Reagent (ThermoFisher). After the addition of chloroform, the RNA-containing aqueous layer was removed for RNA isolation according to the manufacturer's protocol, beginning with the RNA bead binding step.

RNA concentration and quality were assessed using the Nanodrop 2000 spectrophotometer (Thermo Scientific) and the RNA Total RNA Nano assay (Agilent Technologies). 2µl of diluted 1:1000 diluted ERCC Spike-in Control Mix 1 (Ambion by Life Technologies) was added to 100ng of each RNA sample prior to library construction. Libraries were prepared by the Genome Technologies core service at The Jackson Laboratory using the KAPA RNA Hyper Prep Kit with RiboErase (HMR) (KAPA Biosystems), according to the manufacturer's instructions. Briefly, the protocol entails depletion of ribosomal RNA (rRNA), RNA fragmentation, first and second strand

cDNA synthesis, ligation of Illumina-specific adapters containing a unique barcode sequence for each library, magnetic bead size selection, and PCR amplification. Libraries were checked for quality and concentration using the D5000 ScreenTape assay (Agilent Technologies) and quantitative PCR (KAPA Biosystems), according to the manufacturers' instructions.

RNAseq libraries were pooled and sequenced by Novogene in 150 bp paired-end format on an Illumina NovaSeq 6000 sequencer targeting 90 million read pairs per sample. Sequencing achieved a median read depth of 132 million reads. The resultant reads were determined to be of consistently high quality using fastqc v0.11.3 and MultiQC v1.2.

Reads were generated from raw data and demultiplexed using BCL2Fastq v2.18.0.12, concatenated by sample, and aligned with the STAR aligner v2.6.1[29] to the GRCm38 mouse reference genome with v94 of the Ensembl transcriptome. Transcript-level quantification was estimated using RSEM v1.3.0[30] on a transcriptome BAM file produced as an output of this alignment. The data were imported into R v3.5.1 and summarized to the gene level using tximport v1.10.1[31], TMM-normalized using edgeR v3.24.3[21], and imported into limma v3.38.3[32] using the $\log_2$-transformation function voom. We compared multivariate approaches modeling with interaction factors between edgeR and voom+limma approaches and found that voom+limma performs better than edgeR for controlling false negatives. Upon initial examination of the findings, we identified intermittent contamination with choroid plexus, which potentially derives from the ventricular aspect of the dorsal striatum. Correcting for this contamination necessited an additive covariate for choroid plexus consisting of log-mean CPM values of *Kl* and *Ttr* expression, unambiguous markers for choroid plexus[33]. These values were $\log_2$ transformed for work in limma. For An overall model for all strains included this choroid plexus factor as a nuisance variable plus the main effects of strain, sex, and cocaine injection and all of their interactions. Individual models included the choroid plexus nuisance variable plus sex, cocaine injection, and sex:cocaine injection interaction. Correction for local false discovery rates utilized the qvalue package in R v2.14.1[34]. Because brain transcriptional changes are subtle[35], all results reported are at $q < 0.01$ with no fold-change cutoff (**Supplementary Table S2**).

Raw data and transcript-level expression estimates will be deposited in the Gene Expression Omnibus[36] upon publication (accession number: GSEXXXXX).

Figures

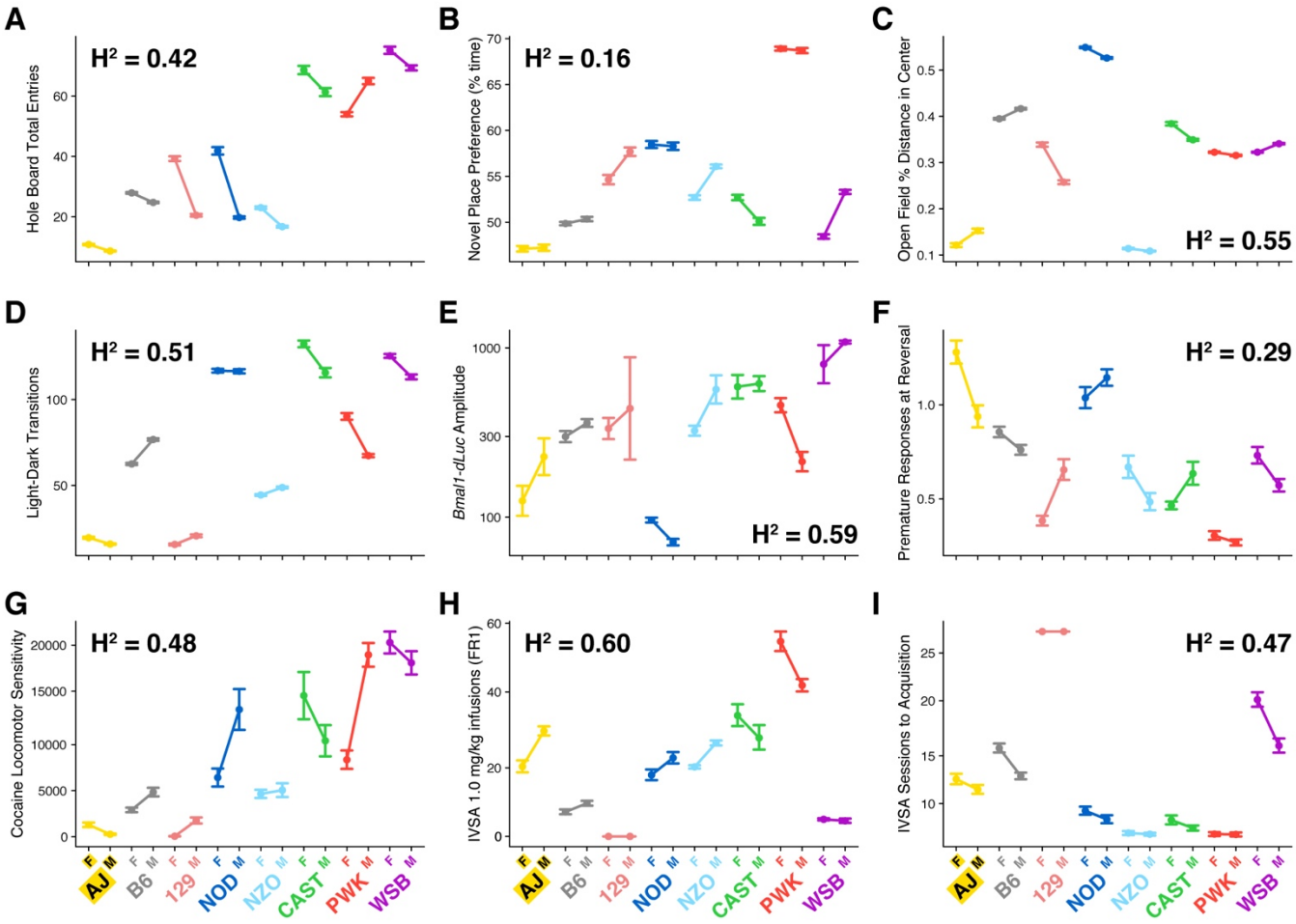

Figure 1: Heritable differences in cocaine-related behavioral and physiological traits. All plots show the mean ± the standard error in both sexes with females on the left. **A)** Hole board total entries, **B)** Novel place preference (% time in novel zone), **C)** Open field (% time in center of open field), **D)** Transitions between light and dark in the light-dark box, **E)** Circadian rhythm fibroblast *Bmal1-dLuc* luminescence amplitude, **F)** Reversal learning premature responses to the hole rewarded during acquisition, **G)** Initial locomotor sensitivity to cocaine (cm moved on Day 3 – Day 2), **H)** Cocaine intravenous self-administration number of infusions at 1.0 mg/kg FR1, and **I)** Sessions to acquisition of cocaine intravenous self-administration.

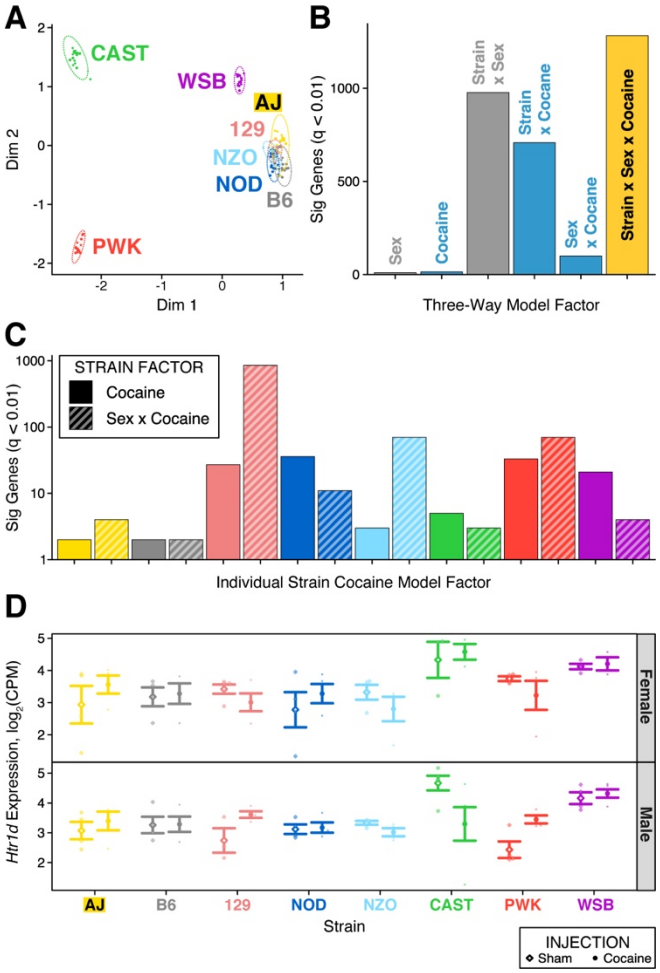

Figure 2: Individual differences in striatum cocaine response transcriptomes. **A)** High heritability of baseline transcriptome partitions the subspecies in a multidimensional scaling plot, **B)** In a three-way analysis of strain, sex, and cocaine, many genes showed significant expression changes in response to cocaine in interaction factors; the strain x sex x cocaine interaction factor identified 1,282 DEGs. In contrast, there were very few significant DEGs in the main effect of cocaine. **C)** Two-way models show the magnitude of transcriptional responses to cocaine differs by strain. B6 has few differentially expressed genes while 129 and PWK have many. The y-axis is $\log_{10}$. **D)** The serotonin 1D receptor gene *Htr1d* is differentially expressed in males of 129S1/SvImJ, PWK/PhJ, and CAST/EiJ strains. CAST/EiJ males show the opposite directionality of differential expression to PWK/PhJ and 129S1/SvImJ males.

Supplemental Material

Supplementary Table S1: Statistical test results for heritability and genotype-by-sex interactions.

Supplementary Table S2: Differential expression results for all tests discussed.

Supplementary Table S3: Sample sizes for all strain and sex combinations for studies reported.

Raw behavioral data will be deposited in the Mouse Phenome Database upon publication.

Gene expression data will be deposited in the Gene Expression Omnibus upon publication.

All scripts, code, and metadata used for analysis are deposited in GitHub (repository: github.com/msaul/csna_founders_survey_2020)

Author Contributions

EJC, PED, LMT, SAS, JDJ, RWL, CAM, LGR, VMP, SJSR, and CSNA conceived the studies. SJSR, PED, JRB, LSB, SAS, RD, ML, AO, TR, TW, and LHG designed and implemented the behavioral experiments. SMK designed and implemented the circadian transcriptional experiment. JRB, LSB, UD, PED, ML, SMK, AO, TR, SAS, TW, LHG, VMP, and MCS analyzed the behavioral data. SMK, MCS, and VMP analyzed the circadian transcriptional data. MCS, VMK, and VMP analyzed the RNAseq data. MCS, UD, PED, JRB, LSB, SAS, TW, LHG, VMP, and EJC interpreted the behavioral results. SMK, MCS, and VMP interpreted the results of the circadian transcriptional experiment. MCS, VMK, VMP, and EJC interpreted the RNAseq results. MCS and EJC wrote the manuscript.

Acknowledgements

We gratefully acknowledge Technologies service at The Jackson Laboratory Genome for their expert assistance on the RNAseq experiment, Surgical Services at The Jackson Laboratory for their work on the intravenous self-administration paradigm, and the NIDA Drug Supply for providing cocaine. We would further like to thank Stephen Krasinski for critical reading and comments during manuscript preparation, Rob Williams for helpful discussion on the analysis,

and C. Herbert Pratt for scientific program management. This work was funded by NIH P50 DA039841 (Center for Systems Neurogenetics of Addiction) to EJC, LMT, JDJ, RWL CAM, VMP, LGR, and SJSR as well as NIH R01 DA037927 to EJC and K99 DA043573 to PED.

Literature Cited

1. SAMHSA. *The 2017 National Survey on Drug Use and Health (NSDUH)*. https://www.samhsa.gov/data/nsduh/reports-detailed-tables-2017-NSDUH (2017).

2. Kochanek, K. D., Murphy, S. L., Xu, J. & Arias, E. Deaths: Final Data for 2017. *Natl. Vital Stat. Rep.* **68**, 1–77.

3. Goldman, D., Oroszi, G. & Ducci, F. The genetics of addictions: uncovering the genes. *Nat. Rev. Genet.* **6**, 521–532 (2005).

4. Becker, J. B. & Chartoff, E. Sex differences in neural mechanisms mediating reward and addiction. *Neuropsychopharmacology* **44**, 166–183 (2019).

5. Becker, J. B., Perry, A. N. & Westenbroek, C. Sex differences in the neural mechanisms mediating addiction: a new synthesis and hypothesis. *Biol. Sex Differ.* **3**, 14 (2012).

6. Walker, D. M. *et al.* Cocaine self-administration alters transcriptome-wide responses in the brain's reward circuitry. *Biol. Psychiatry* (2018).

7. Fitzpatrick, C. J. *et al.* Variation in the Form of Pavlovian Conditioned Approach Behavior among Outbred Male Sprague-Dawley Rats from Different Vendors and Colonies: Sign-Tracking vs. Goal-Tracking. *PLoS ONE* **8**, e75042 (2013).

8. Prendergast, B. J., Onishi, K. G. & Zucker, I. Female mice liberated for inclusion in neuroscience and biomedical research. *Neurosci. Biobehav. Rev.* **40**, 1–5 (2014).

9. Tuttle, A. H., Philip, V. M., Chesler, E. J. & Mogil, J. S. Comparing phenotypic variation between inbred and outbred mice. *Nat. Methods* **15**, 994–996 (2018).

10. Saul, M. C., Philip, V. M., Reinholdt, L. G. & Chesler, E. J. High-Diversity Mouse Populations for Complex Traits. *Trends Genet.* **0**, (2019).

11. Roberts, A., De Villena, F. P.-M., Wang, W., McMillan, L. & Threadgill, D. W. The polymorphism architecture of mouse genetic resources elucidated using genome-wide resequencing data: implications for QTL discovery and systems genetics. *Mamm. Genome* **18**, 473–481 (2007).

12. Ersche, K. D., Turton, A. J., Pradhan, S., Bullmore, E. T. & Robbins, T. W. Drug addiction endophenotypes: impulsive versus sensation-seeking personality traits. *Biol. Psychiatry* **68**, 770–773 (2010).

13. Dickson, P. E. *et al.* Association of novelty-related behaviors and intravenous cocaine self-administration in Diversity Outbred mice. *Psychopharmacology (Berl.)* **232**, 1011–1024 (2015).

14. Wahlsten, D., Metten, P. & Crabbe, J. A rating scale for wildness and ease of handling laboratory mice: results for 21 inbred strains tested in two laboratories. *Genes Brain Behav.* **2**, 71–79 (2003).

15. Logan, R. W., Williams III, W. P. & McClung, C. A. Circadian rhythms and addiction: Mechanistic insights and future directions. *Behav. Neurosci.* **128**, 387–412 (2014).

16. Ozburn, A. R. *et al.* Direct Regulation of Diurnal Drd3 Expression and Cocaine Reward by NPAS2. *Biol. Psychiatry* **77**, 425–433 (2015).

17. Ramanathan, C. *et al.* Cell Type-Specific Functions of Period Genes Revealed by Novel Adipocyte and Hepatocyte Circadian Clock Models. *PLoS Genet.* **10**, e1004244 (2014).

18. Kim, S.-M., Neuendorff, N., Chapkin, R. S. & Earnest, D. J. Role of Inflammatory Signaling in the Differential Effects of Saturated and Poly-unsaturated Fatty Acids on Peripheral Circadian Clocks. *EBioMedicine* **7**, 100–111 (2016).

19. Izquierdo, A. & Jentsch, J. D. Reversal learning as a measure of impulsive and compulsive behavior in addictions. *Psychopharmacology (Berl.)* **219**, 607–620 (2012).

20. de Wit, H. & Phillips, T. J. Do initial responses to drugs predict future use or abuse? *Neurosci. Biobehav. Rev.* **36**, 1565–1576 (2012).

21. Chen, Y., Lun, A. T. L. & Smyth, G. K. Differential Expression Analysis of Complex RNA-seq Experiments Using edgeR. in *Statistical Analysis of Next Generation Sequencing Data* (eds. Datta, S. & Nettleton, D.) 51–74 (Springer International Publishing, 2014). doi:10.1007/978-3-319-07212-8_3.

22. Phipson, B., Lee, S., Majewski, I. J., Alexander, W. S. & Smyth, G. K. Robust Hyperparameter Estimation Protects Against Hypervariable Genes and Improves Power to Detect Differential Expression. *Ann. Appl. Stat.* **10**, 946–963 (2016).

23. Wiltshire, T. *et al.* Initial locomotor sensitivity to cocaine varies widely among inbred mouse strains: Initial locomotor sensitivity to cocaine in inbred mice. *Genes Brain Behav.* **14**, 271–280 (2015).

24. Kumar, V. *et al.* C57BL/6N mutation in cytoplasmic FMRP interacting protein 2 regulates cocaine response. *Science* **342**, 1508–1512 (2013).

25. Ruan, Q. T. *et al.* A Mutation in Hnrnph1 That Decreases Methamphetamine-Induced Reinforcement, Reward, and Dopamine Release and Increases Synaptosomal hnRNP H and Mitochondrial Proteins. *J. Neurosci.* **40**, 107–130 (2020).

26. Izumo, M., Johnson, C. H. & Yamazaki, S. Circadian gene expression in mammalian fibroblasts revealed by real-time luminescence reporting: Temperature compensation and damping. *Proc. Natl. Acad. Sci.* **100**, 16089–16094 (2003).

27. Schoenrock, S. A. *et al.* Characterization of genetically complex Collaborative Cross mouse strains that model divergent locomotor activating and reinforcing properties of cocaine. *Psychopharmacology (Berl.)* (2020) doi:10.1007/s00213-019-05429-3.

28. Bogue, M. A. *et al.* Mouse Phenome Database: a data repository and analysis suite for curated primary mouse phenotype data. *Nucleic Acids Res.* gkz1032 (2019) doi:10.1093/nar/gkz1032.

29. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

30. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. 16 (2011).

31. Soneson, C., Love, M. & Robinson, M. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research* **4**, (2016).

32. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47–e47 (2015).

33. Sathyanesan, M. *et al.* A molecular characterization of the choroid plexus and stress-induced gene regulation. *Transl. Psychiatry* **2**, e139 (2012).

34. Storey, J. D. & Tibshirani, R. Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci.* **100**, 9440–9445 (2003).

35. Hitzemann, R. *et al.* Introduction to Sequencing the Brain Transcriptome. in *International Review of Neurobiology* vol. 116 1–19 (Elsevier, 2014).

36. Barrett, T. *et al.* NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* **41**, D991–D995 (2012).