Copyright

by

Rayna Michelle Harris

2017

HIDDEN TEXT: Optional—If you do not include a copyright page, delete entire page and the following page break.

HIDDEN TEXT: NOTE: this page in hard copy with all original signatures must be submitted with the dissertation to the Graduate School; this is required whether the document is in electronic format or on paper. Whereas, the page included in the electronic document will be unsigned unless it is scanned in.

The Dissertation Committee for Rayna Michelle Harris Certifies that this is the approved version of the following dissertation:

Transcriptional Plasticity in the Hippocampus and its Role in Conditioned Avoidance Learning

**Committee:**

|  |
| --- |
|  |
| Hans A. Hofmann, Supervisor |
| Laura L. Coglin |
| André A. Fenton |
| Mikhail V. Matz |
| Boris V. Zemelman |

HIDDEN TEXT: The top line is for the Supervisor’s signature. There should be as many lines as there are members on the committee. Lines must be solid, not dotted. To delete signature lines, select the line you want to delete, go to the Table menu, select Table Properties, click on the Table tab, and click on the Borders and Shading button, then remove the bottom border of the table. Use the professor's name without titles or degrees

Transcriptional Plasticity in the Hippocampus and its Role in Conditioned Place Avoidance Learning

by

Rayna Michelle Harris

HIDDEN TEXT: Given first name, and previous academic degrees (B.A. or higher) B.A., B.S., etc. Your official name is the name which appears on your UT transcript.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

HIDDEN TEXT: The degree sought must be worded in the form given in the Graduate Catalog, such as Doctor of Philosophy, Doctor of Musical Arts, Doctor of Education.

The University of Texas at Austin

December 2017

Dedication

I dedicate this dissertation to my mother who loves and supports me even though I talk too infrequently for her taste, to my father who would have been so proud of this accomplishment if he had lived long enough to see this day, and to all the people in my extended Harris and Hunt families who encourage and inspire my academic endeavors.

Preface

This preface describes a personal and professional journey studying the molecular biology of brain function and complex behavior. Even with rapid advances in technology, we still don't understand the genetic and environmental interactions that give rise to complex diseases and behavior with small effect size. Approaches that use behavioral genomics or data science methods to answer questions are often criticized for lacking explicit hypotheses and predictions. Neuroscience experiments are criticized for lacking behaviorally relevant manipulations and being too reductionists. Senior scientists everywhere are saying that inadequate computational training is the most significant impediment to scientific advances. For my thesis, I attempted to overcome all of the problems mentioned above in modern biology to identify molecular links between brain and behavior.

The ecosystem surrounding my research was multi-disciplinary and multi-institutional. Because a broad audience my influenced research, I have a strong motivation to write my research for a broad audience. I hope that all my colleagues and future researchers can find a least one thing of use in my body research, whether it be related to a biological phenomena or technical innovation.

Before I devolve into the domain-specific details of research and its implications, let me first describe the ecosystem where I worked and thank all the people who have made significant contributions to my growth and development as a scientist.

First and foremost, this thesis would not exist without my, **Dr. Hans Hofmann**. We built many academic programs and things that enriched our academic careers (and hopefully those of our colleagues). We carried out numerous successful and unsuccessful projects, but Hans in his eternal optimism always insisted that 'failures' were successes in one shape or form.

Though not technically part of my graduate research, the three years I worked as a lab manager in the Hofmann lab (2009-2012) played a defining role in shaping my brain and my view of the world. **Lauren O'Connell** was an incredible mentor who taught me numerous molecular protocols and gave great advice that jump-started my academic journey. **Misha Matz, Carly Kenkle, and Sarah Davis** adopted me into their evolutionary genomics journal club where I discovered some of the analytical methods I still use today. **Steve Phelps, Kim Hoke, and Ron Oldfield** supported my professional development those early years by writing numerous letters of recommendation for travel awards and fellowships. I'll never forget how many great ideas were sparked during the Friday "Brain, Behavior, and Evolution Seminars" and the "Friday Afternoon Rooftop Socials" which lead to funded research proposals and published papers.

Eventually, I enrolled as a graduate student in the Department of Cell and Molecular Biology in the fall of 2012. For the next five years, **Rachel Wright** was my "partner-in-crime" when it came to studying for exams, writing grant proposals, doing yoga, seeing live bands, and navigating life's hurdles. In the classroom, **Rick Russel, Jeff Gross, Ruth Buskirk, and Ed Marcotte** helped me develop a deep understanding of genetics, molecular biology, systems biology, and evolution that I've carried with me throughout my career. **Claus Wilke** never formally taught or mentored me, but his online blog and virtual books were a constant information and advice. I am grateful to **Vishy Iyer, Andrea Gore, and John Mihic, Nancy Moran** for their service on my qualifying examination and thesis proposal committees.

The summer of my first year in grad school, Hans introduced me to a beautiful community of neuroscientists that are united by the historic "Neural Systems and Behavior Course", which takes place for eight weeks every summer at the Marine Biological Laboratory in Woods Hole Massachusetts. It was here that I developed my thesis project and gained valuable experience mentoring students. There was undeniable synergy between teaching and doing science in the highly productive Neural Systems and Behavior environment. Proximity to my collaborator was critical for interpreting experimental results while questions from students pushed me to understand my approach even better.

I want to thank the other Co-Director and my collaborator, **André Fenton**, for all the wisdom, creativity, and patience he brought to my thesis research projects over the years. I thank **Hsin-Yi (Maddy) Kao, Ain Chung, and Juan Marcos Alarcon** for their tireless dedication to these collaborative research projects we conducted each summer. I think it is incredible that we were able to walk into an empty room and build a laboratory designed specially to carry out the research described in my thesis (**Fig. 0.1**).

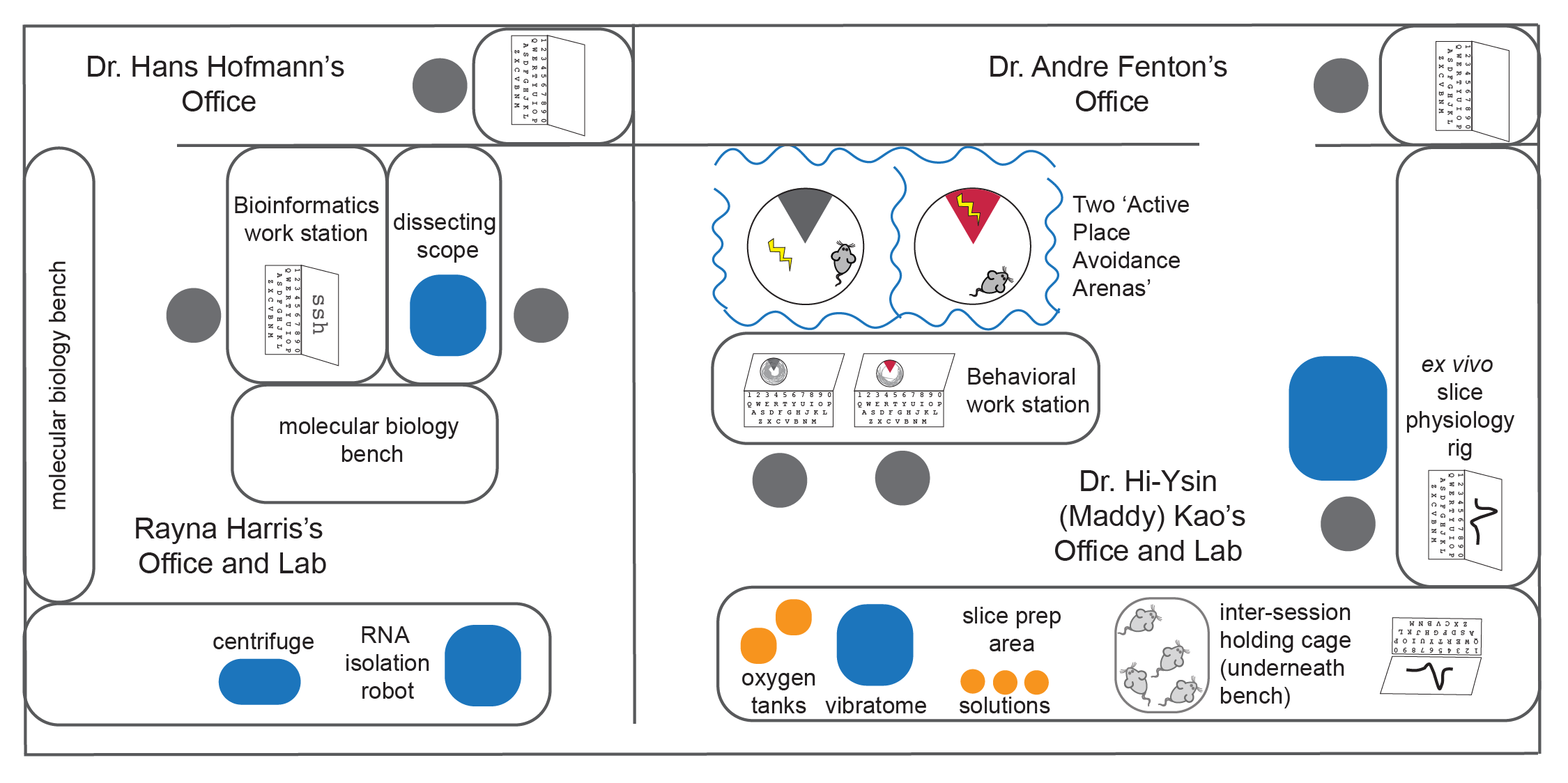
****

Fig. 0.1: Schematic illustration of my research and teaching laboratory the Marine Biological Laboratories.

As for teaching, my goal at the Neural Systems and Behavior Course was to build a new 2-week module where students could learn molecular approaches to complement behavioral and electrophysiological methods used in non-traditional species. Primarily, I taught students how to use quantitative real-time PCR and bioinformatics for RNA sequencing analysis1,2. I thank **David Shultz** and **Eva Fischer** for being a 'shining beacons of humanity' and demonstrating how to foster learning in a molecular biology teaching lab. There are too many to name individually, but I want to thank **all the NS&B students, faculty, and staff** who spent nearly every waking moment in the laboratory in pursuit of scientific knowledge and discovery.

The spring of my second year in graduate school, Hans hired me to be the training and outreach coordinator for the **Center for Computational Biology and Bioinformatics**, and it was here that I met a diverse community of scientists who use and love bioinformatics. I thank all the graduate students who volunteered their time leading the peer-led working groups where I shared our newfound knowledge and helped each other troubleshoot problems in R and Python. I thank **Scott Hunicke Smith, Dhivya Arrassappan, Anna Battenhouse, Benni Goetz, and Dennis Wylie** for developing the bioinformatic consulting group, which helped me build a strong foundation in data-driven discovery. I especially thank **Laurie Alvarez and Nicole Elmer** for orchestrating all the social events that brought us together.

I'm so grateful **April Wright** introduced me to the wonderful **Software Carpentry and Data Carpentry** Communities in the winter of 2015 at the exact moment when I needed and wanted it most. I was looking for role models who taught reproducible bioinformatics, and I hit the jackpot. I'm especially grateful to **Greg Wilson and Tracy Teal** who recognized my potential and opened doors that I didn't even know existed. **Kate Hertweck** became my constant companion for teaching, mentoring, and organizational governance (sometimes in real life but usually through virtual meetings). **Christina Koch, Erin Becker, and Sue McClatchy** helped me learn educational pedagogy, which allowed me to start thinking about the science of learning in the classroom and how it relates to the neuromolecular basis of learning and memory. I can't wait to see how the Carpentry communities continue to grow and spread the best-practices for research and teaching to people all around the world.

I want to close by thanking the thesis committee members and the Hofmann lab members who have guided me across the finish line. I thank **Laura Colgin** for the most elegant descriptions and illustrations of how the hippocampus works. I thank **Boris Zemelman** for keeping me up to date on cutting-edge research in molecular neuroscience. I thank **Mikhail Matz** always responding to my request for new functions or helping me troubleshoot errors associated with various R packages. Again, I thank André Fenton for the wisdom, creativity, patience, and dedication he brought to this research. I want to thank **Rebecca Young-Brim, Tessa Solomon Lane, Mariana Rodriguez-Santos, Caitlin Friesen, Eric Brenner, and Isaac Miller-Crews** for helping me craft perfect figures at our 'figure it out' meetings and for making the Hofmann lab such a friendly and productive environment.

Finally, I want to thank my family, friends, and colleagues who attended my public thesis defense. Your support and insightful comments inspire me to continue exploring scientific unknowns and sharing my findings.

Transcriptional Plasticity in the Hippocampus and its Role in Conditioned Place Avoidance Learning

Rayna Michelle Harris, Ph.D.

The University of Texas at Austin, 2017

Supervisor: Hans A. Hofmann

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis id aliquam quam, sed accumsan ligula. Vestibulum urna nunc, lobortis ac tincidunt nec, placerat a urna.

Table of Contents

HIDDEN TEXT: If you choose to place the chapter number (“Chapter 1”) and the chapter title (“Introduction”) on different lines, the automatically generated table of contents will reflect that format. After creating a new table of contents, set them on the same line by deleting the page number and paragraph marker at the end of each chapter number line.

List of Tables xiv

List of Figures xv

Introduction 1

Chapter 1: Avoidance learning induces strong up-regulation of transcription in the DG subfield of the dorsal hippocampus 3

Abstract 3

Introduction 4

Material & Methods 7

The Active Place Avoidance Task 7

Statistical analyses of behavior 8

Tissue preparation from DG, CA3, and CA1 subfields 9

RNA-sequencing and bioinformatics 9

Statistics and data visualization of RNA-sequencing data 10

Archival of data, code, and figures 10

Results 10

Conflict- and consistently-trained mice exhibit place avoidance 10

Conflict-trained mice exhibit cognitive discrimination 12

Confirmation of subfield-specific gene expression patterns 12

Research question 1: How does memory-associated place avoidance alter gene expression in the dorsal hippocampus? 14

Research question 2: Does cognitive discrimination alter gene expression? 15

Research question 3: Does unavoidable punishment (in the form of random, mild foot-shocks) alter gene expression? 16

Discussion 17

Acknowledgments 20

Chapter 2: FMR1-KO knockdown alters gene expression of some autism disorder risk genes 21

Abstract 21

Introduction 22

Methods 23

Active Place Avoidance 23

Statistical analyses of behavior and synaptic physiology 24

Tissue preparation and electrophysiology 25

RNA-sequencing and bioinformatics 25

Statistics and data visualization of RNA-sequencing data 26

Archival of data, code, and figures 26

Results 26

Confirm that FMR1-KO does not impair learning or memory 27

Fail to replicate previous research showing that FMR1-KO impairs cognitive discrimination 28

Transcriptional response to consitutative gene knockdown 29

Discussion 32

Acknowledgments 33

Chapter 3: Analysis of hippocampal transcriptomic responses to technical and biological perturbations 35

Abstract 35

Introduction 36

Methods 38

Sample processing 38

Bioinformatics 39

Meta-analysis 39

Archival of data, code, and figures 40

Results 40

The effects of cellular dissociation 40

The effects of stressful experience 41

The effects of cognitive training 43

Identifying unique and general patterns of hippocampal genomic plasticity 44

Discussion 47

Conclusions 50

Acknowledgments 50

Concluding Thoughts and Future Directions 52

Appendix 1: Candidate gene expression and synaptic plasticity do not explain differences in WT and FMR1-KO mice in avoidance behavior 53

Abstract 53

Introduction 53

Methods 55

Experimental design 55

Tissue collection 56

RNA isolation, cDNA synthesis, and real-time PCR 56

Statistical analyses 57

Archival of data and code 58

Results 58

Discussion 59

Acknowledgments 60

Appendix 2: My publications with significance and contribution 61

Bibliography 69

Vita … 76

## List of Tables

Table 1.1: Differentially expressed genes (alpha = 0.1) by cognitive training and subfield. 16

Table A1: TaqMan assays used in the study 53

## List of Figures

Fig. 0.1: Schematic illustration of my research and teaching laboratory the Marine Biological Laboratories. vii

Fig. 0.2. Graphical abstract of Transcriptional Plasticity in the Hippocampus and its Role in Avoidance Learning. 1

Fig. 1.1: Experimental design. 8

Fig. 1.2: Cognitive training alters spatial approach and avoidance behavior 11

Fig. 1.3: Subfield differences in hippocampal gene expression 13

Fig. 1.4: Place avoidance is associated with up-regulation genes involved in regulation of transcription. 15

Fig. 1.5: Place avoidance is associated with increased expression of genes that regulate synaptic activity in CA1 16

Fig. 1.6: Additional punishment also influences gene expression in CA1 17

Fig. 2.1: Active place avoidance task with conflict training. 24

Fig. 2.2: WT and FMR1-KO mice exhibit place avoidance and cognitive discrimination 27

Fig. 2.3: WT and FMR1-KO show evidence of differential gene expression in the CA1 29

Fig. 2.4: Downregulation of ion transport in the the the CA1 subfield 31

Fig. 2.5: CA3-CA1 synaptic strength is not altered by genotype or place avoidance training. 32

Fig. 3.1: The effect of cellular dissociation 42

Fig. 3.2: The effects of a stressful experience 43

Fig. 3.3: The effects of learned avoidance behavior 44

Fig. 3.4: Overlapping responses to technical and biological perturbations 45

Fig. 3.5: Meta-analysis of primary and public data 46

Fig. A1: The collaborative ‘Integrative Molecular Neuroethology Laboratory’ at the Marine Biological Laboratories. 55

Fig A2: Experimental design, methods, and results 58

## 

## Introduction

This doctoral thesis described three independent by highly similar research projects. Chapter 1 lays the foundation thoroughly describing how a conditioned place avoidance paradigm and its associated stressors alter transcriptomic and synaptic activity in the hippocampus. In chapter two, add a genetic manipulation to understand how behavior, physiology, and gene expression respond to this type of perturbation. Finally, in chapter 3 I conduct a preliminary study to understand how new technological methods for RNA sequencing alter hippocampal transcriptomes. This multi-disciplinary research provides a through analysis of how the brain works in the context of learning and memory and beyond.

I conducted a qPCR experiment to ask if avoidance behavior is generated by molecular changes that alter the activity of synapses in neural circuits (Appendix 1). HIDDEN TEXT: The following sample text and headings are for information only. Delete them after browsing.

Fig. 0.2. Graphical abstract of Transcriptional Plasticity in the Hippocampus and its Role in Avoidance Learning.

During the first year of the Hofmann-Fenton collaboration, I conducted a qPCR experiment to ask if avoidance behavior is generated by molecular changes that alter the activity of synapses in neural circuits. This project was not officially part of my thesis, but I have included in in the Appendix 1 because this experiment laid the foundation by demonstrating feasibility.

## Chapter 1: Avoidance learning induces strong up-regulation of transcription in the DG subfield of the dorsal hippocampus[[1]](#footnote-1)

### Abstract

The hippocampus is widely known to be involved in spatial navigation, learning, and memory, but little is known about how dynamic activity of transcriptome supports these cognitive processes. I aimed to fill that gap by asking, “How do memory-associated place avoidance and cognitive discrimination alter gene expression in the dorsal hippocampus?”. This research required an interdisciplinary approach to utilize both the hypothesis-driven techniques of behavioral neuroscience and the discovery-driven methods of behavioral genomics to identify novel neuromolecular substrates of memory. I found that gene expression in CA1 and DG discriminates internal versus external variables because yoked and trained mice had the identical physical experience of the world but could interpret the experience differently. I observed that different gene functions and patterns of differential expression in the hippocampal subfields. DG most responsive to memory formation, and all the upregulated genes are known to be involved nuclear signaling; while changes in CA1 reflect membrane-level ion-channel regulation. The CA1 results also demonstrate sensitivity to the amount of unavoidable shock or stress that occurred 24h in the past, which could be a molecular signature for a form of stress-related memory. Finally, the results demonstrated that conflict learning does not cause additional gene expression changes relative to initial learning. The strength of this research is its collaborative, integrative, and reproducible approaches that provide a deeper understanding of the molecular changes that are or are not associated with robust behavioral output indicative of memory.

### Introduction

For centuries, scientists and philosophers have sought to understand how the brain uses memory to drive changes in behavior. One challenge is that memory cannot be physically isolated. Memory is an instance in which an organism's current behavior is determined by some aspect of its previous experience3. It an emergent property that researchers observe as an overt change at one or more levels of biological organization. Recent advances in molecular biology and neuroscience pushed our understanding past the concept of the neural doctrine (single neurons are the brain’s information processing unit of organization) and the central dogma of molecular biology (function is coded by the process DNA->RNA->Protein) which alone do not solve the problem of understanding how the brain stores and recalls memories that change animal behavior. Neurons differ from one another in many ways: structurally, functionally and genetically, as well as the connections they make with other cells. The extent to which gene expression accounts for the function and characteristics of cell types, neural systems, and behavioral expression is unknown despite being fundamental to many research programs in biology and medicine. One goal of modern neuroscience is to fully characterize the structure and function of all the molecules in all the neurons in all the circuits of the human brain (4,5 but see 6). Additionally, understanding how such variability and plasticity in the brain gives rise to emergent phenotypical variation is of utmost importance for advancing neuroscience and related fields. Generations of behavioral neuroscientists have developed sophisticated paradigms for assessing the biological correlates of memory and cognition7. However, it can be a challenge to interpret the significance of observed behavior differences. “The problem with rats is that I can't ask them if they remember what they had for breakfast,” Howard Eichenbaum once said8. Nevertheless, community-developed standards were established such that robust behavioral phenotypes could be used as indicators of memory-associated changes in behavior.

The hippocampus is widely known to be involved in spatial navigation, learning, and memory9. In 1909, Cajal illustrated the hippocampus with directional connections between neurons between the entorhinal cortex (EC), subregions of Cornu Ammonius' Ammon's horn (CA) and the dentate gyrus (DG), with information flowing from in the EC -> CA1 -> CA3 -> DG -> EC or through additional pathways that skip subfields or layers within subfields10. In 1971, John O'Keefe found that place cells in the hippocampus function as an internal global positions system (GPS) in our brain9,11. Different place cells are active when an animal is in different locations within the environment12. In the 2000s, May-Britt Moser and Edvard Moser demonstrated the place cells of the CA1 were connected to grid cells of the entorhinal cortex13. The CA1 pyramidal cells are often referred to as place cells because they activate when an animal is in a particular place in its environment. The CA3-CA1 synapse has received a lot of attention for long-term potentiation (LTP), long-term depression (LTD), and other forms of synaptic plasticity as it might relate to learning. The DG has been given some attention for adult-born neurogenesis and its role in learning and memory14–17.

Since the early 80s with discoveries that the NMDA receptor and Calcium-dependent protein kinase II (CaMKII) play crucial roles in LTP, neuroscientists have been using pharmacology to identify the molecular and neural substrates of synaptic plasticity and memory. These studies linked changes in behavior to long-term potentiation, long-term depression, and the expression of numerous candidate genes18,19. Molecular and cellular biologists have shed light on what molecules serve as biomarkers for cells with different structure and function20–22. This tools have highlighted the heterogeneity of the structure and function of neurons and have lead to new hypothesis regarding how different neuron classes contribute to cognition. Modern genomic techniques such as microarrays and next-generation sequencing have made it possible to examine the molecular underpinnings of plasticity in animal behavior23–26. Yet despite its successes in generating discovery-driven insights, behavioral genomics has been criticized for its lack hypotheses27. Additionally, even though databases and repositories of gene expression for mouse hippocampi do exists28–33, they provide little insight into how dynamic activity of transcriptome supports memory.

The goal of this research is to marry the hypothesis-driven techniques of behavioral neuroscience with the discovery-driven methods of behavioral genomics to identify novel neuromolecular substrates of memory. First, I analyzed the behavior of mice in a task explicitly designed to test how the brain uses spatial information for memory processing. My experiment used a variant of the active place avoidance task14,34–37 which allowed us to tease apart the effects of subtle differences in an animal's experience of the environment. Then, I characterized the transcriptional pattern of activity from tissue samples from the DG-CA3-CA1 tri-synaptic pathway that has been implicated in spatial learning and memory. The results of this work showed that this work confirms previously identified sub-field specific patterns of expression and confirms previous descriptions of candidate memory gene activity in CA1. Importantly, I also detected novel patterns of transcription factor activation in the DG and a surprising lack of transcriptional plasticity in CA3. These two findings suggest new mechanisms for how information flows through a DG-CA3-CA1 pathway that has been implicated in memory. The results provide a unique perspective on the role of transcriptional stability and plasticity about hippocampal-dependent learning and memory. In the spirit of reproducible research and open science38–42, all data, code, and results are publically available and licensed under a creative commons for reuse for the continued advancement of basic research.

### Material & Methods

All animal care and use complies with the Public Health Service Policy on Humane Care and Use of Laboratory Animals and were approved by the New York University Animal Welfare Committee and the Marine Biological Laboratory Institutional Animal Care and Use Committee. Male C57BL/6J mice were housed at the Marine Biological Laboratory on a 12:12 (light: dark) cycle with continuous access to food and water in home cages with up to five littermates.

#### The Active Place Avoidance Task

To examine spatial learning and memory,we used a well-established active place avoidance paradigm14,34,43. Littermates were randomly assigned to one of our treatment groups (nyoked-consistent=8, nconsistent=8, nyoked-conflict=9, nconflict=9). All mice were exposed to nine sessions in the active place avoidance arena (**Fig. 1.1A**). Mice were placed on an elevated circular 40-cm diameter arena that rotated at 1 rpm. The arena wall was transparent and thus contained the mice to the arena while allowing it to observe the environment. The location of the mouse in the arena was determined from an overhead television camera a PC-controlled tracking system (Bio-Signal Group). Consistently trained mice in the active place avoidance task are conditioned to avoid mild shocks (constant current 0.2 mA 500 ms 60 Hz) that can be localized by visual cues in the environment. Yoked mice are delivered a sequence of unavoidable shocks that reproduces the time series of shocks received by the trained mice; however, the shocks delivered to the mice cannot be localized by visual cues in the environment. All sessions in the arena last 10 minutes. Mice are allowed to become familiar with walking on a rotating arena during a 10 min pre-training session. Then mice undergo 3, 10-min avoidance training sessions separated by a 2 h inter-trial interval. Mice are returned to their home cage overnight. The next day, mice are subjected to a 10-min ‘Retest session’ where the shock is in the same location as before. For the next three training sessions, the shock zone remains in the same place for consistently trained animals, but it is rotated 180° for the conflict-trained mice. The next day, all mice are subjected to a 1-min “Retention session” with the shock off.

#### Statistical analyses of behavior

Place avoidance was evaluated by end-point measures output by TrackAnalysis software. Forty quantitative variables were measured that capture the animals’ use of space and time. The two estimates of place avoidance used were the reduction in entrances into the shock zone and the increased time to stay out of the shock zone. A repeated measures ANOVA with sphericity correction was used to identify group differences in behavioral measure across all training sessions44,45. A two-way ANOVA was used to compare the mean difference between conflict and consistently trained mice during a single training session44. The non-parametric Kendall's tau statistic was used to estimate a rank-based measure of correlation between the number of entrances and maximum avoidance time44.

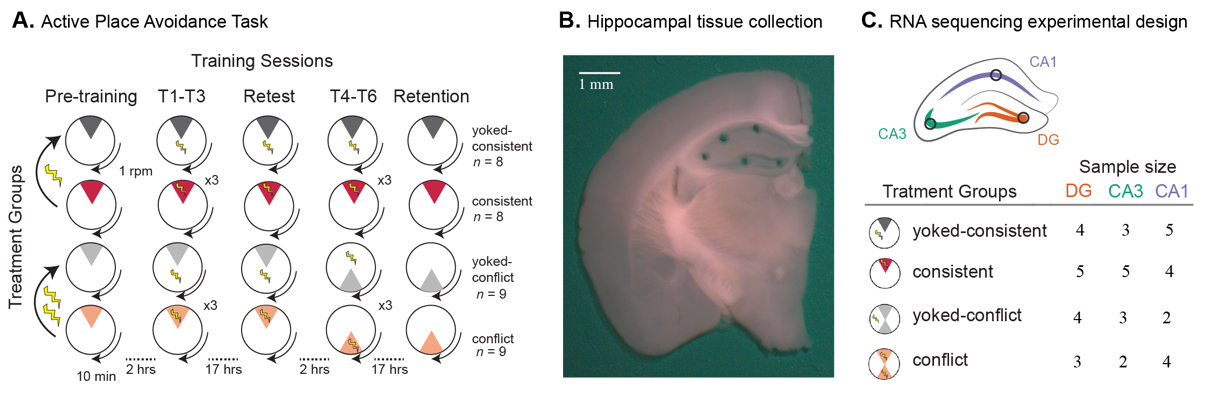


Fig. 1.1: Experimental design.

**A**) Mice were assigned to one of four groups: consistently-trained (red, n=8), yoked-consistent (dark grey, n=8), conflict-trained (peach, n=9), or yoked-conflict (light grey, n=9). Mice were placed on the rotating arena (1 rpm) for training sessions that lasted 10 min and was separated by 2-hour intersession interval or overnight (~17 hrs). Behavior was recorded during the Pre-training, Training (T1-T6), Retest, and Retention session. In the active place avoidance schematics, the shaded pie-shaped region is the behaviorally relevant region used for counting the number of entrances into the shock zone. The shocking of yoked mice is not spatially limited to the dark-grey pie-shaped zone, but consistent and conflict trained mice only receive shocks in the red and peach-shaded regions, respectively. **B**) A representative photo shows the size and location of tissue samples collected for RNA-sequencing. **C**) Graphical illustration of hippocampal tissues sequenced and sample sizes for each treatment group and hippocampal subfield.

#### Tissue preparation from DG, CA3, and CA1 subfields

Thirty minutes after the last cognitive training session, mice were anesthetized with 2% (vol/vol) isoflurane for 2 minutes and decapitated. Transverse 300 μm brain slices were cut using a vibratome (model VT1000 S, Leica Biosystems, Buffalo Grove, IL) and incubated at 36°C for 30 min and then at room temperature for 90 min in oxygenated artificial cerebrospinal fluid (aCSF in mM: 125 NaCl, 2.5 KCl, 1 MgSO4, 2 CaCl2, 25 NaHCO3, 1.25 NaH2PO4 and 25 Glucose) 46,47. The DG, CA3, CA1 subfields were microdissected using a 0.25 mm punch (Electron Microscopy Systems) and a Zeiss dissecting scope (**Figure 1B**). RNA was isolated using the Maxwell 16 LEV RNA Isolation Kit (Promega). RNA libraries were prepared by the Genomic Sequencing and Analysis Facility at the University of Texas at Austin and sequenced on the Illumina HiSeq platform.

#### RNA-sequencing and bioinformatics

Raw reads were from the transferred from the GSAF to the Stampede Cluster at the Texas Advanced Computing Facility (TACC) via Amazon Cloud. Quality of the data was checked using the program FASTQC48 and visualized using MultiQC49. Samples sizes for each treatment group and subfield are reported in **Figure 1C**. I obtained a median number of 7.3 million reads per samples, with a maximum of 37 million and a minimum of 1.5 million reads. Next, I used the program Kallisto50 for pseudo-alignment of raw reads to a mouse references transcriptome (Gencode version 7)51. A median number of 2.25 million reads were pseudo-aligned to the references transcriptome. To confirm that Kallisto performed well on raw reads, I also removed low-quality reads and contaminating adapter sequences using the program Cutadapt52, but these trimmed and filtered reads yield less than 1 million mapped reads for all but 3 of the 44 samples; therefore all subsequent analyses were conducted on the pseudo-aligned raw reads.

#### Statistics and data visualization of RNA-sequencing data

Transcript counts from Kallisto were imported into R44 and aggregated to yield gene counts using the ‘gene’ identifier from the Gencode transcriptome. I used DESeq253 to normalize and quantify gene counts with a false discovery corrected (FDR) p-value < 0.1. The DESeq2 models including subfield (DG, CA3, CA1), training group (yoked-consistent, consistent, yoked-conflict, and conflict) and their interaction. Hierarchical clustering by correlation and volcano plots were used to visualize the patterns of differential gene expression53–58. Principal component analysis (PCA) was conducted to reduce the dimensionality of the data, and ANOVAs were used to test for group differences in the principal components. A chi-squared goodness of fit test was used to test for equal distribution of up- and down-regulated gene expression between two-way contrasts44. I used GO\_MWU59 to identify gene ontology categories using a -log(p-value) as a continuous measure of significance that are significantly enriched with either up- or down-regulated genes for a given two-way contrast**.** I used ggplot258, cowplot57, pheatmap60, viridis61, ColorBrewer54, and colorblindr62 to make figures that are (hopefully) color-blind friendly. Multi-panel figures and illustrations were created using Adobe Illustrator.

#### Archival of data, code, and figures

I archived the raw sequence data and intermediate data files in NCBI's Gene Expression Omnibus Database (accession: GSE99765). The data and code are publically available on GitHub (https://github.com/raynamharris/IntegrativeProjectWT2015).

### Results

#### Conflict- and consistently-trained mice exhibit place avoidance

Active place avoidance is evidenced by a reduction in the mean number of entrances into the shock zone (F(24,240) = 5.140, p = 5.57e-08) (**Fig. 1.2A**). Place avoidance is also evidenced by an increase in the time the mouse stays out of the conditioned shock zone, as illustrated by time to the second entrance (F(24,240) = 5.30, p = 2.20e-09) (**Fig. 1.2B**).

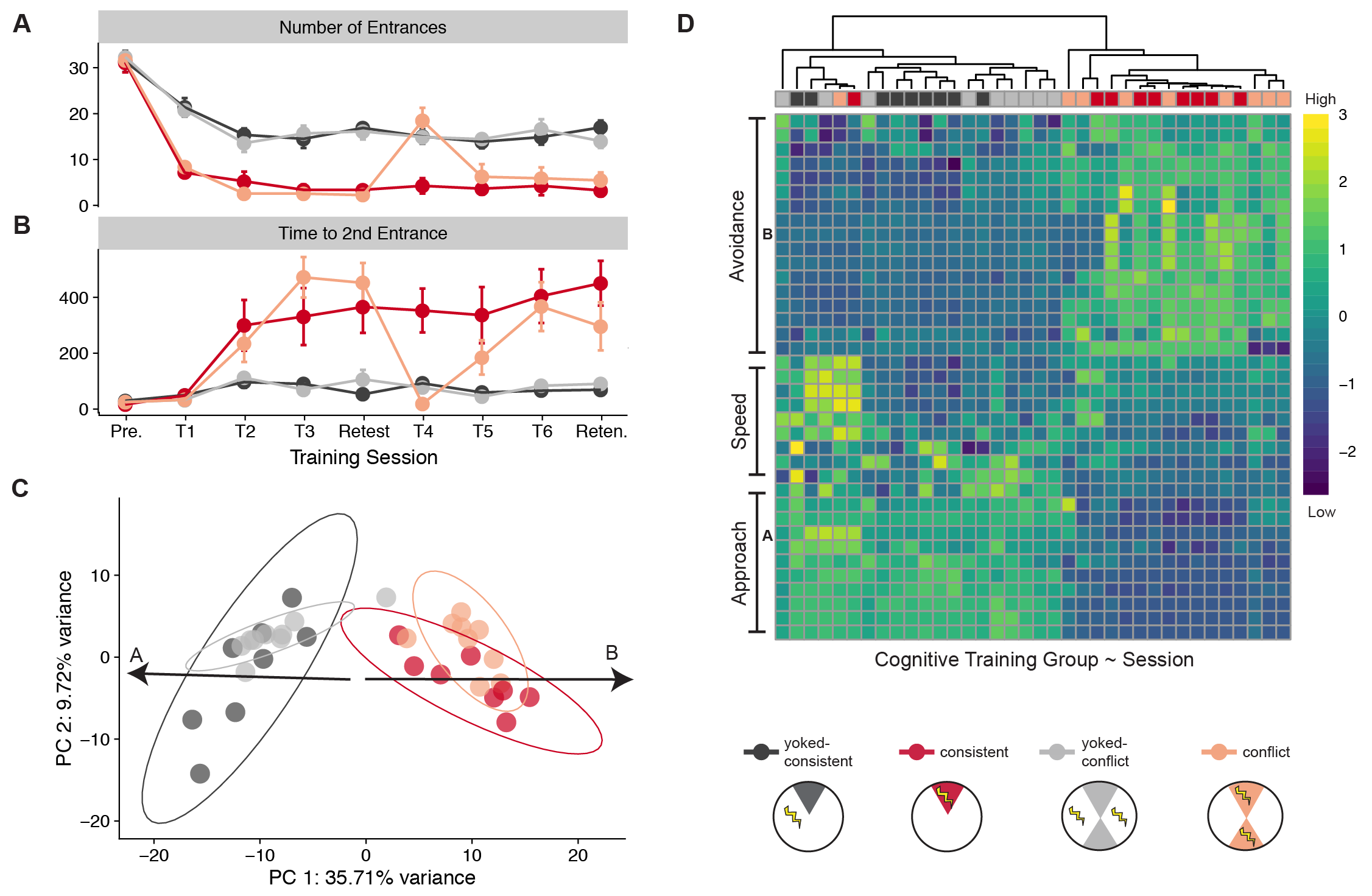


Fig. 1.2: Cognitive training alters spatial approach and avoidance behavior

**A**) Consistently trained (red lines) mice make fewer entrances into the shock zone than yoked-mice (dark grey lines) on all training (T1-T6), restest, and retention (Reten.) sessions but not during the pre-training session (Pre.). Conflict-trained mice (peach) and their yoked controls (light grey) show a similar pattern except for that mean number of differences between T1 and T4 do not differ between conflict-trained mice. **B**) Time to second shock zone entrance shows a pattern that is reciprocal to the mean number of entrances. **C**) A principal component analysis estimates that cognitive training explains 36% of the observed variation in behavior (red and peach versus dark grey and light grey). Among the top five contributing variables are the number of entrances and the max avoidance time. **D**) Hierarchical clustering by correlation of 40 behaviors shows that approach, latency to approach (or avoidance), and speed are primary behavioral variables captured by our video-tracking software. Clustering distinguishes trained and yoked animals but does not provide precise temporal resolution. The color scale shows centered z-s for high (yellow) and low (deep purple) values for each quantitative variable.

These two measures of behavior (number of entrances and time to second entrance) are inversely correlated (tauτ=-0.55, p < 2.2e-16). A principal component analysis of all the quantitative variables shows that the two measures load strongly onto PC1, which is significantly different between treatments groups (F3=70.92, p = 1.01e-13) (**Fig. 1.2C**). In fact, most of the quantitative measures captured by the video tracking software are positively or negatively correlated with the measures of shock zone entrances or avoidance, as evidenced by a hierarchical clustering analysis (**Fig. 1.2C**). Notably, measures of speed are not correlated with place avoidance, but there is a pattern that speed is highest during the pre-training session compared to the later sessions.

#### Conflict-trained mice exhibit cognitive discrimination

Cognitive discrimination requires distinguishing between similar but distinct experiences. I investigated cognitive discrimination by changing the location of shock. Conflict-trained and consistently-trained mice differ in the mean number of entrances during the T4/C1 session (F(1) =17.49, p=0.000801), but not during the T6/C3 session (F(1) = 0.265, p=0.614) (**Fig. 1.2A**). These results indicate that, with continued training, the conflict mice rapidly learned to discriminate between the memories of the old and new shock locations.

#### Confirmation of subfield-specific gene expression patterns

Large differences in subfield specific gene expression are well documented28–32, but the association of memory and gene expression is understudied. Thus, I examined broad patterns of gene expression variation in DG, CA3, and CA1 tissue samples from mice in each of the four cognitive training groups (**Fig. 1.1C, Table 1.2A**). My results confirm previous studies by showing significant differences in gene expression of thousands of genes between brain regions (**Fig. 1.3**). Hierarchical clustering of the top 250 differentially expressed genes at FDR 0.1 reveals a strong signature of subfield-specific expression, with all samples clustering by subfield (**Fig. 1.3B**). PC1 and PC2 account for 71% of the variation and are both significantly different between subfields (PC1 - F2:41=256.2, p << 0.001; PC2 - F2:41=1030, p << 0.001). PC6 explains only 1% of the variation in gene expression it does vary according to cognitive training (F3:40=12.01, p<<0.001).

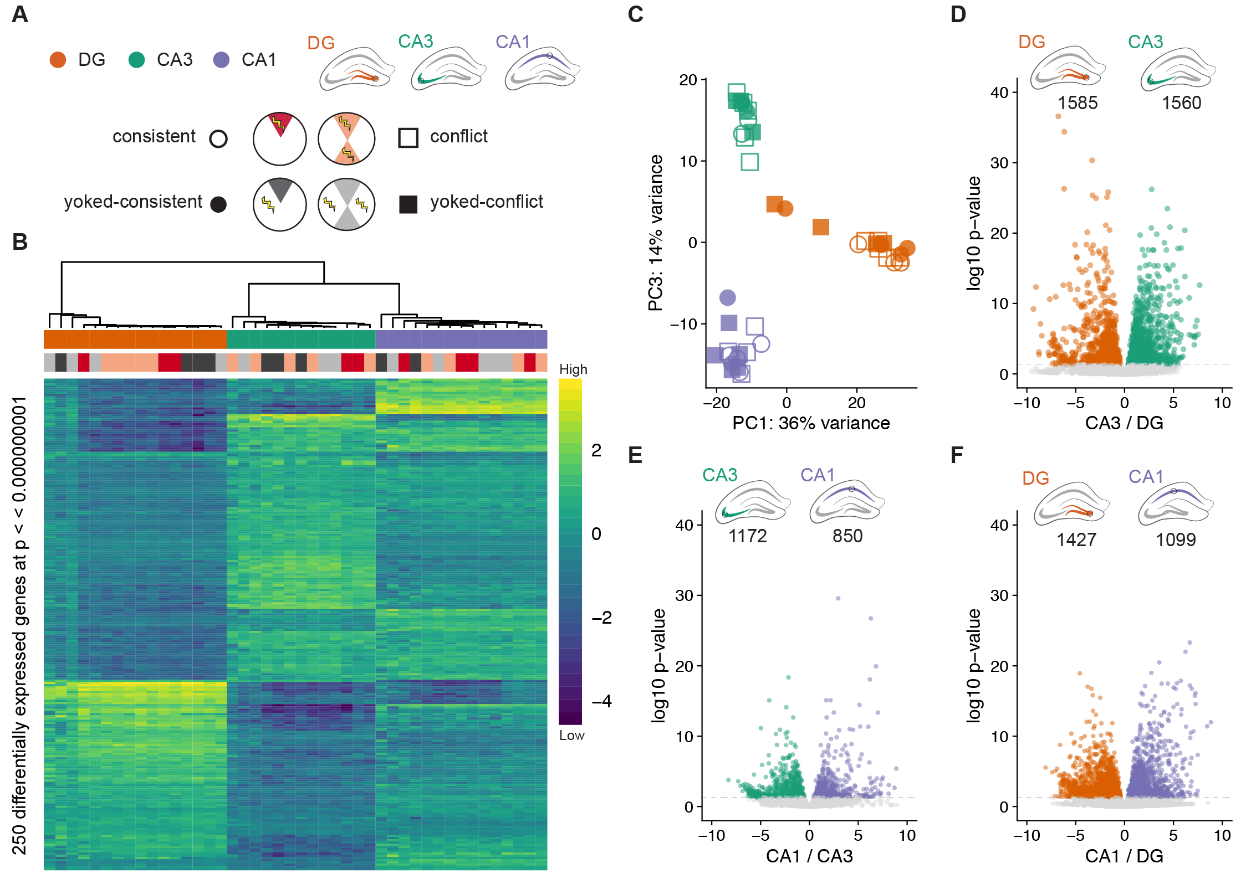


Fig. 1.3: Subfield differences in hippocampal gene expression

**A**) I compared gene expression in three hippocampal subfields from our four treatment groups (DG: orange, CA3: green, CA1: purple, yoked-consistent: filled circle, consistent: open square, yoked-conflict: filled square, conflict: open square). **B**) Hierarchical clustering of differentially expressed genes shows variation between subfields is much greater than variation induced by treatment. **C**) A principal component analysis estimates that over 50 % of the variation is capture in PC1 and P3, which visually separate the three hippocampal subfields. **D**) 3000 are differentially expressed in a symmetric pattern between DG and CA. **E, F**) Fewer genes are up-regulated in CA1 compared to both DG and CA3, but the magnitude of expression differences in greater between DG-CA1 than between CA3-CA1. For volcano plots, dots are partially transparent to aid visualization of density.

Hierarchical clustering and principal component analyses are a convenient way to visualize gene expression differences between all samples, but volcano plots are a convenient way to explore two-way contrasts in more detail. The contrast between brain regions highlights the magnitude of differential expression between subfields (DG vs. CA3: 3145, DG *vs*. CA1: 2526, CA3 *vs.* CA1: 2022 differentially expressed genes at FDR=0.1). The distribution differential gene expression is symmetrical between CA3 and DG (**Fig 1.3D**). Fewer genes that half the differentially expressed genes are up-regulated in CA1 compared to DG (**Fig 1.3E**) and CA3 (**Fig 1.3F**), but the magnitude of expression differences in greater between DG-CA1 (**Fig 1.3F**).

#### Research question 1: How does memory-associated place avoidance alter gene expression in the dorsal hippocampus?

All analyses of the effect of training on hippocampal gene expression were conducted independently for each subfield. In the DG, 116 genes were upregulated in the consistently training group while 0 were upregulated in the yoked group (**Fig 1.4, Table 1.2**). Among the top ten differentially expressed genes are those encoding transcription factors Egr4, Junb, SMAD. A gene ontology analysis shows an enrichment in molecular functions related to core promoter binding, nuclear localization sequence binding, signal sequence binding, poly(A) RNA binding, and heat-shock protein binding (**Fig 1.4**).

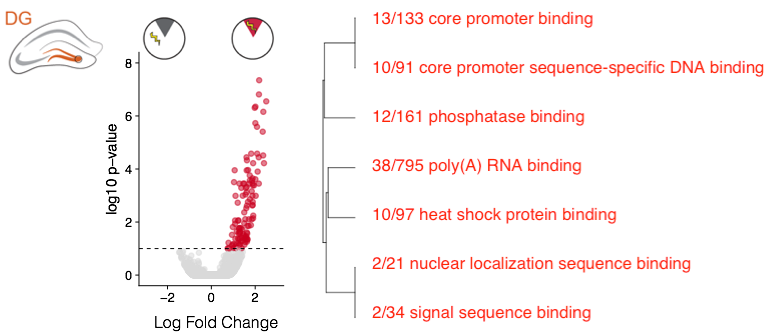


Fig. 1.4: Place avoidance is associated with up-regulation genes involved in regulation of transcription.

In the dentate gyrus (DG) 116 genes are upregulated in the consistently training group compared to the yoked samples (FDR = 0.1). An analysis of enrichment in gene ontology categories shows an enrichment in molecular function processes related to promoter binding and nuclear sequence binding (p < 0.05). Genes and GO categories are colored according to enrichment in trained (red) or yoked (black). The active place avoidance schematics, the shaded pie-shaped region is the behaviorally relevant region for counting a number of entrances into the shock zone. Trained (red) mice are shocked in this zone, but the shocking of yoked mice is not spatially limited to the dark-grey pie-shaped zone.

No genes in CA3 were differentially expressed in the consistently trained vs. yoked contrast (Table 1.2). This lack of activity is consistent with the Denny et al. study who expressed Cre-ERT2 under the direction of the activity regulated cytoskeletal-associated protein (Arc) promoter regions to compare activation of neuronal populations in the hippocampus during encoding and retrieval of memory17. However, in CA1, approximately 600 genes were differentially expressed between trained vs. yoked contrasts (Fig. 1.5, Table 1.2). The distribution of up and down-regulation of genes is symmetric in the CA1 subfield. An analysis of gene ontology categories shows an enrichment for molecular functions related to ion channel synthesis and activity (**Fig. 1.5**).

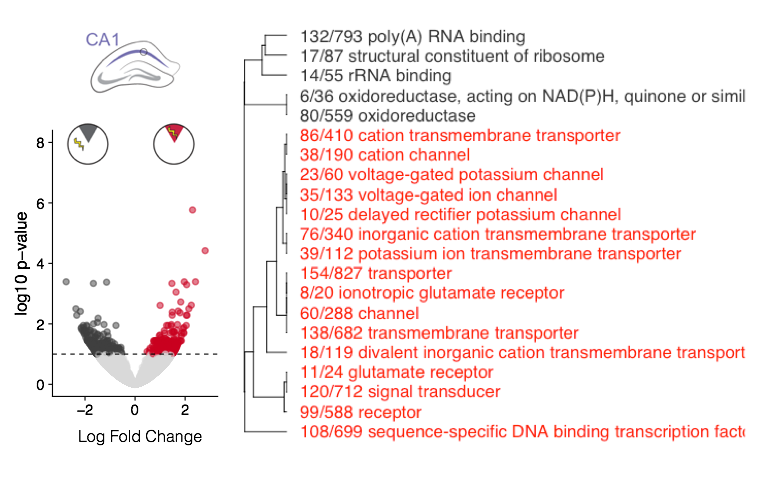


Fig. 1.5: Place avoidance is associated with increased expression of genes that regulate synaptic activity in CA1

In the CA1, 253 genes are upregulated in the consistently training group while only 255 are downregulated (FDR = 0.1). An analysis of enrichment in gene ontology (GO) categories shows an enrichment in molecular function processes related to ion channel transport and activity (p < 0.05). Genes and GO categories are colored according to enrichment in trained (red) or yoked (black)

#### Research question 2: Does cognitive discrimination alter gene expression?

Changing the shock zone location provides a test of cognitive discrimination that requires DG function14,63. Once an animal learns the new location of shock, cognitive discrimination enables judicious use knowledge of the current and former locations of shock. Yoked, consistently-trained, and conflict-trained animals vary in their degree of behaviors expressed that evidence cognitive discrimination. We asked if changes in gene expression are associated with cognitive discrimination. We found gene expression in the DG, CA3, and CA1 shows a remarkable lack of differential expression in response conflict training (**Table 1.1**). The results show that initial learning of the shock zone does initial learning of the shock zone does change synaptic function in the perforant path DG and in the pyramidal layer of CA1; however, less than 1/1000 of the transcriptome is differentially expressed in the animals that behaviorally demonstrated cognitive discrimination in relation.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Icons** | **Two-way contrasts between groups** | **DG** | **CA3** | **CA1** |
|  | consistent vs. yoked-consistent | 116 | 0 | 508 |
|  | conflict vs. consistent | 0 | 2 | 0 |
|  | yoked-conflict vs. yoked-consistent | 1 | 1 | 409 |
|  | conflict vs. yoked-conflict | 4 | 0 | 0 |

Table 1.1: Differentially expressed genes (alpha = 0.1) by cognitive training and subfield.

Consistent training alters the expression of ~100 and ~200 genes in DG and CA1, respectively. Conflict training has almost no effect on hippocampal subfield expression relative to its yoked counterpart nor to the consistently trained animals.

#### Research question 3: Does unavoidable punishment (in the form of random, mild foot-shocks) alter gene expression?

The yoked-conflict group received more foot-shocks on day two when the conflict animals were performing the cognitive discrimination task. The CA1 but not the DG or CA3 shows gene expression response to differing levels of punishment (**Fig. 1.6**). In the CA1, I identified differentially expressed of 409 genes between yoked groups that received different amounts of punishment (FDR = 0.1). The volcano plots show a near symmetric distribution of genes that are higher in the consistent yoked (left side, dark-grey) and conflict yoked (right side, light grey). Analysis of enrichment in gene ontology (GO) categories shows an enrichment in molecular function processes related to glutamate receptors, signal transduction, ion channel transport in the conflict-yoked group; whereas the consistent yoked group showed an enrichment in processes related to RNA binding and ribosomal activity (p < 0.05).

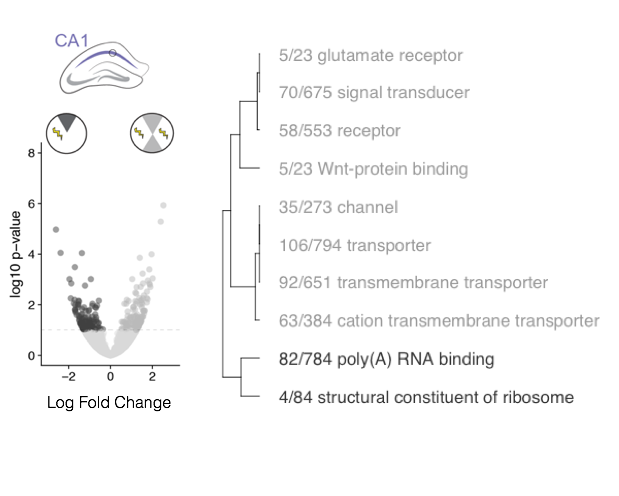


Fig. 1.6: Additional punishment also influences gene expression in CA1

In the CA1, 409 genes are differentially expressed in between yoked groups that received different amounts of punishment. (FDR = 0.1). An analysis of enrichment in gene ontology (GO) categories shows an enrichment in molecular function processes related to ion channel transport and synaptic activity (p < 0.05). Genes and GO categories are colored according to enrichment in yoked-conflict (light grey) or yoked-consistent (dark-grey).

### Discussion

My approach combines hypothesis-driven and data-driven techniques to yield new insight into the dynamic nature of the hippocampus. I observed behaviors subservient of learning, memory, and cognitive discrimination in our experimental groups of mice, and we identified associated changes in the expression of specific molecular pathways in subfields hippocampus using RNA sequencing. I filled the gap by being the first to conduct an unbiased RNA sequencing screen of genes related to subfields of the hippocampus. The brain-wide scans of the Allen Institute28,29 and Cembrowski et al.31,32 were unbiased in their molecular discovery, as they examined behaviorally naïve mice and so did not incorporate rigorous behavioral assays. The exquisitely detailed work of a cadre of hippocampal neuroscientists described single genes or pathways in detail, but they did not measure the activity of the entire transcriptome.

This approach led to the following finding and interpretations. Gene expression in was DG most responsive to memory formation, which supports previous findings of physiological plasticity in the region Park et. 2015. One can conclude that gene expression discriminates internal versus external variables because yoked and trained mice had the identical physical experience of the world but could interpret the experience differently – gene expression in CA1 and DG sensitive to this. Different types of genes are differentially regulated in DG and CA1. Change in DG is related to nuclear signaling while changes in CA1 reflect membrane-level ion-channel regulation. The CA1 results also demonstrate sensitivity to the amount of unavoidable shock or stress that occurred 24h in the past. Importantly, tissue was sampled at a time when cortisol levels are very likely to be equal64,65. This CA1 response could be a molecular signature for a form of stress ‘memory’ (i..e. a persistent (24-h) change in the system as a function of experience even if the mouse does not express it by long-term behavior or physiology). Additionally, I did detect changes in protein kinases, but I did not detect any differences in the variant PKMz or PKC, even thought Pastalkova et al. 2006 and others have shown that DG synaptic plasticity and place avoidance memory is crucially mediated by PKMz66. Finally, the results demonstrated that conflict learning does not cause additional gene expression changes relative to initial learning. Perhaps the changes associated with memory formation and behavioral change occur within a network of cells that were not captured in our tissue samples.

There are several limitations to our experimental design. First of all, gene expression was only measured at one time-point. Thus I cannot say what the activity looks like during different stages of memory acquisition and recall. Secondly, I measured activity within many neurons and other cell types in a tissue sample of the hippocampal subfields. This research provided more insight into the subfield specificity of the hippocampus, but single-cell techniques will be needed to tease apart differences among cell types or cells with differencing activity states. Thirdly, I did not conduct an unbiased screen of all subfields of the hippocampus. This study overlooked the CA2, CA4, and neighboring cortical regions that send and receive information from the hippocampus. Finally, when looking for small effects, it is essential to have a sufficient sample size, so we opted to sequence more biological replicates from a few brain regions than the other way around. In the end, a small sample size does limit the power of our experiment to identify gene expression changes associated with memory. A challenge for the future will be to capture and understand molecular activity in the hippocampus before, during, and after learning.

The strength of this research is its collaborative and integrative approach. We brought hypotheses and expertise from relating disciplines of biology to address current problems in neuroscience and genomics. Thus, we now have a deeper understanding of the molecular changes that are or are not associated with robust behavioral output indicative of memory. I also applied best practices in open and reproducible research to provide a robust pipeline that can be re-used. The raw and processed data, analysis pipelines, results, and interpretations are available for download, enabling the curious to export data into their environment for more specialized analyses. This dataset expands upon and complement other existing publicly available gene expression databases, mostly notably Lein et al., 2007 and Cembrowski et al., 2016.

### Acknowledgments

I thank Laura Colgin, Mariana Rodriguez, and Eric Brenner for comments on earlier versions of this manuscript. I thank Becca Young Brim, Caitlin Freisen, Tessa Solomon Lane, Mariana Rodriguez, Eric Brenner, and Issac Miller-Crews for helping advice regarding on figures and oral presentations of this research. Thanks to the general R scienticific community for openly sharing your software, for making useful tutorials, and for being responsive to user-requests for new functionality.

## Chapter 2: FMR1-KO knockdown alters gene expression of some autism disorder risk genes[[2]](#footnote-2)

### Abstract

Fragile X syndrome is the most prevalent form of genetically inherited intellectual disability. FMRP regulates the local translation of a subset of mRNAs at synapses. In the absence of FMRP, dysregulated mRNA translation can lead to altered synaptic function, loss of protein synthesis-dependent synaptic plasticity, and impaired cognitive discrimination. To better understand the effects silencing the *Fmr1* gene, we conducted an integrative analysis of learning behavior, cognitive discrimination, and hippocampal gene expression. Importantly, I confirmed that FMR1-KO does not impair learning or memory; however, I failed to replicate previous research showing that FMR1-KO impairs cognitive discrimination. This lack of difference in the expression of avoidance behavior may be a result of a “noisy” environment that interfere with the task, or it could indicate that cognitive deficits are not very robust. I identified 20 genes whose expression in the CA1 subfield was altered by constitutive knockdown of FMR1, including *Grin1*, *Cacna1g, Car4, Ccnd1.* These results suggest that knockdown causes dysregulation of multiple genes that contribute to autism spectrum disorder, but that this disruption does not have significant effects on a hippocampal-dependent place avoidance condition task. This research provides a unique contribution to diverse fields of studying including hippocampal function and human health research.

### Introduction

Fragile X syndrome is the most prevalent form of genetically inherited intellectual disability. This X-linked disorder is caused by mutations in the untranslated region of the Fmr1 gene repeat that give rise to hypermethylation and transcriptionally silencing. Thus, patients with fragile X syndrome do not make the fragile X mental retardation protein (FMRP). Autistic-like features and problems with working and short-term memory are frequent in patients with fragile X syndrome67. Autism spectrum disorder is a genetically heterogeneous condition. While many genes predisposing an individual to autism spectrum disorder, have been identified and understanding the causal disease mechanism remains elusive68. A deeper understanding of the molecular interaction between fmr1 in a brain region-specific manner might shed light on phenotypic variation.

FMRP regulates the local translation of a subset of mRNAs at synapses. In the absence of FMRP, dysregulated mRNA translation leads to altered synaptic function and loss of protein synthesis-dependent synaptic plasticity69. Voineagu et al. 2011 found that striking regional patterns of attenuated gene expression in the frontal and temporal cortex suggesting abnormalities in cortical patterning70. Fmr1 knock-out mice71 show cognitive discrimination deficits are prominent, but learning and memory appear unimpaired35.

To better understand the effects silencing the Fmr1 gene, we conducted an integrative analysis of learning behavior, cognitive discrimination, and hippocampal gene expression. Cognitive discrimination requires distinguishing between similar but distinct experiences. I investigated cognitive discrimination by changing the location of shock. I predicted that FMR1-KO would alter cognitive discrimination performance, CA3-CA1 synaptic strength, and the activity of hundreds of genes in the CA1 subfield of the hippocampus. To address this question, I used the active place avoidance task to measure the behavioral correlates of learning, memory, and cognitive discrimination. Then, I analyzed slice electrophysiological data to estimate changes in CA3-CA1 synaptic strength. Finally, I identified patterns of differential gene expression in the CA1 subfield as a result of constitutive Fmr1 knockout. This approach provides a blueprint for investigation of the behavior, physiology, and molecular consequences of FMR1 KO in a rodent model. The data, analysis tools, and results are publically available for readers to explore in more detail.

### Methods

All animal care and use complies with the Public Health Service Policy on Humane Care and Use of Laboratory Animals and were approved by the New York University Animal Welfare Committee and the Marine Biological Laboratory Institutional Animal Care and Use Committee. Male C57BL/6J wild-type and Fragile X mental retardation protein knockout mice (FMR1-KO) were housed on a 12:12 (light: dark) cycle with continuous access to food and water in home cages with up to five littermates.

#### Active Place Avoidance

To examine spatial learning and memory,we used a well-established active place avoidance paradigm14,34,43. Littermates were randomly assigned to one of our treatment groups (WT: nyoked-consistent=4, nconsistent=8, nyoked-conflict=3, nconflict=9; FMR1-KO: nyoked-consistent=7, nconsistent=9, nyoked-conflict=5, nconflict=5). All mice were exposed to nine sessions in the active place avoidance arena (**Fig. 2.1**). Mice were placed on an elevated circular 40-cm diameter arena that rotated at 1 rpm. The arena wall was transparent and thus contained the mice to the arena while allowing it to observe the environment. The location of the mouse in the arena was determined from an overhead television camera a PC-controlled tracking system (Bio-Signal Group). Consistently trained mice in the active place avoidance task are conditioned to avoid mild shocks (constant current 0.2 mA 500 ms 60 Hz) that can be localized by visual cues in the environment. Yoked mice are delivered a sequence of unavoidable shocks that reproduces the time series of shocks received by the trained mice; however, the shocks delivered to the mice cannot be localized by visual cues in the environment. All sessions in the arena last 10 minutes. Mice are allowed to become familiar with walking on a rotating arena during a 10 min ‘Pre-training session’. Then mice undergo 3, 10-min avoidance training sessions separated by a 2 h inter-trial interval. Mice are returned to their home cage overnight. The next day, mice are subjected to a 10-min ‘Retest session’ where the shock is in the same location as before. For the next three training sessions, the shock zone remains in the same place for consistently trained animals, but it is rotated 180° for the conflict-trained mice. The next day, all mice are subjected to a 1-min “Retention session” with the shock off.

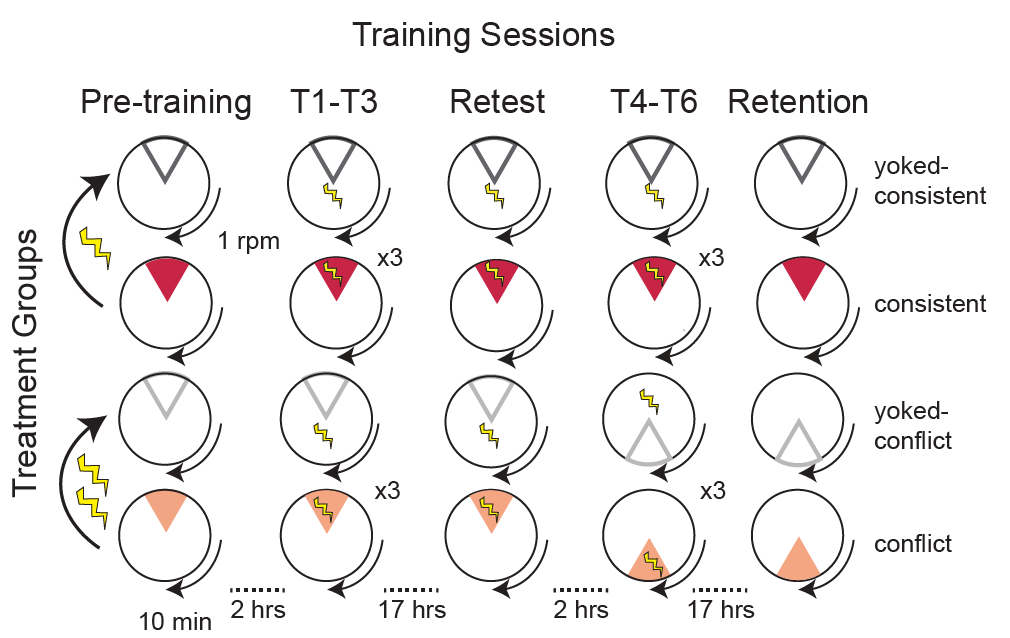


Fig. 2.1: Active place avoidance task with conflict training.

Mice were assigned to one of four groups: consistently-trained (red), yoked-consistent (dark grey), conflict-trained (peach), or yoked-conflict (light grey). Mice were placed on the rotating arena (1 rpm) for training sessions that lasted 10 min and was separated by 2-h intersession interval or overnight (~17 h).

#### Statistical analyses of behavior and synaptic physiology

Place avoidance was evaluated by end-point measures output by TrackAnalysis software. Forty quantitative variables were measured that capture the animals’ use of space and time. I evaluated reduction in entrances into the shock zone to estimate place avoidance. A three-way ANOVA with Tukey Honest Significant Differences (TukeyHSD) test was carried out to determine the influence of Genotype \* Treatment Group \* Training Session and the Genotype \* Treatment Group interaction on number of entrances. A two-way ANOVA was carried out to determine the influence of Genotype \* Treatment Group and their interaction on CA3-CA1 synaptic strength, as measured by the maximum eFPSP slope.

#### Tissue preparation and electrophysiology

Thirty minutes after the last cognitive training session, mice were anesthetized with 2% (vol/vol) isoflurane for 2 minutes and decapitated. Transverse 300 μm brain slices were cut using a vibratome (model VT1000 S, Leica Biosystems, Buffalo Grove, IL) and incubated at 36°C for 30 min and then at room temperature for 90 min in oxygenated artificial cerebrospinal fluid (aCSF in mM: 125 NaCl, 2.5 KCl, 1 MgSO4, 2 CaCl2, 25 NaHCO3, 1.25 NaH2PO4 and 25 Glucose)46,47.

Field excitatory postsynaptic potentials (fEPSP) from the CA3-CA1 input (Stratum Radiatum) were obtained via stimulation with bipolar electrodes. Stimulus-response relationships between input voltage stimulation and fEPSP slope amplitude were generated at increasing voltage stimulations at CA3-CA1 inputs47.

#### RNA-sequencing and bioinformatics

The CA1 subfields were microdissected using a 0.25 mm punch (Electron Microscopy Systems) and a Zeiss dissecting scope. RNA was isolated using the Maxwell 16 LEV RNA Isolation Kit (a donation from Promega). RNA libraries were prepared by the Genomic Sequencing and Analysis Facility at the University of Texas at Austin and sequenced on the Illumina HiSeq platform. Raw reads were from the transferred from the GSAF to the Stampede Cluster at the Texas Advanced Computing Facility (TACC) via Amazon Cloud. Quality of the data was checked using the program FASTQC48 and visualized using MultiQC49. Low-quality reads and contaminating adapter sequences were removed using the program Cutadapt52. I used Kallisto50 for pseudo-alignment of reads and transcript counting using the-the Gencode M11 mouse transcriptome51.

#### Statistics and data visualization of RNA-sequencing data

Transcript counts from Kallisto were imported into R44 and aggregated to yield gene counts using the ‘gene’ identifier from the Gencode transcriptome. I used DESeq253 to normalize and quantify gene counts with a false discovery corrected (FDR) p-value < 0.1. The DESeq2 model included only genotype (WT vs. FMR1KO) since all samples came from the CA1 subfield of the yoked-consistent treatment group. Hierarchical clustering by correlation was used to visualize the patterns of differential gene expression53–58. I used ggplot258, cowplot57, pheatmap60, viridis61, and colorblindr62 to make figures that are color-blind friendly. Multi-panel figures were created using Adobe Illustrator.

#### Archival of data, code, and figures

I archived the raw sequence data and intermediate data files in NCBI's Gene Expression Omnibus Database (accession GSE100225). The data, code, and results are publically available on GitHub (https://github.com/raynamharris/FMR1CA1rnaseq).

### Results

FMRP loss has been to affect memory discrimination but not the initial forming of memories. This lack of memory discrimination is thought to lead to exaggerated responses to environmental changes in autistic patients. The goal of this research was to identify transcriptional changes in FMR1-KO mice that might explain impaired memory discrimination. Given that FMRP is a translational modifier, little research has been done to investigate transcriptional changes upstream that might occur through regulatory feedback processes. We used the active place avoidance task with conflict learning to observe initial avoidance learning and cognitive discrimination in WT and FMR1-KO mice. The behavioral paradigm was designed specifically to constrain animal behavior to provide hypothesis-driven insight into to role of FMRP in cognition and cognitive discrimination. In this task, yoked mice and trained mice experience the same physical world, but they have a different internal representation of the world as expressed by their spatial use of the active place avoidance arena. The use of two training variation allows the experimenter to tease apart initial avoidance learning and cognitive discrimination.

#### Confirm that FMR1-KO does not impair learning or memory

In the Active Place Avoidance Task, learning evidenced by a reduction in the mean number of entrances into the shock zone. To determine if learning was impaired in FMR1-KO, I measured the mean number of entrances into the shock zone in animals that were consistently trained or yoked-consistent (**Fig. 2.2**).

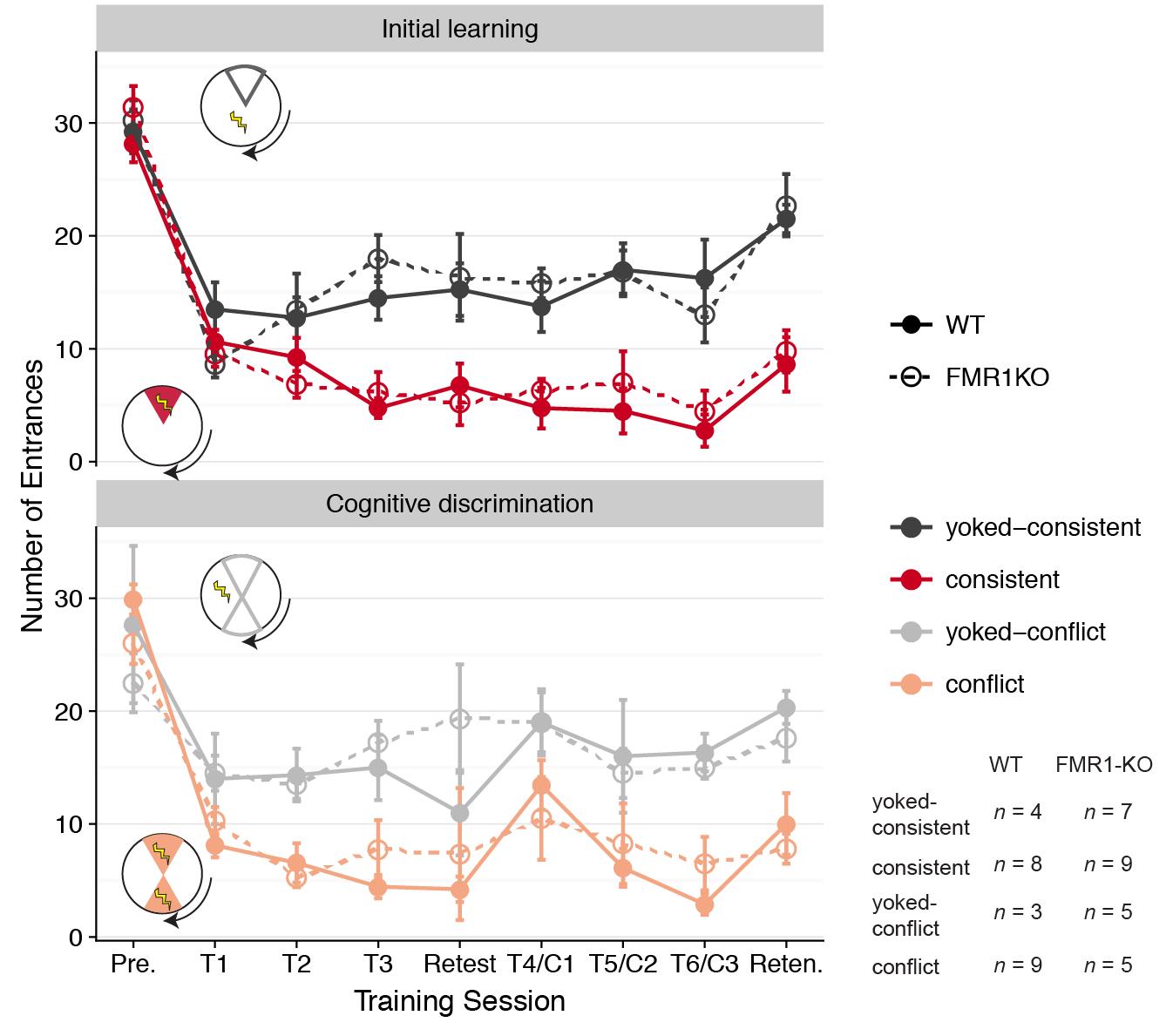


Fig. 2.2: WT and FMR1-KO mice exhibit place avoidance and cognitive discrimination

**Top**) Consistently trained mice (red) from both WT (solid lines, filled circles) and FMR1-KO (dashes lines, open circles) make fewer entrances into the shock zone than yoked-mice (grey). **Bottom**) Conflict trained mice (pink) from both WT (solid lines, filled circles) and FMR1-KO (dashes lines, open circles) learn to avoid the initial and the conflicting shock zones. Data are reported as mean +/- SEM. A Mann-Whitney statistical test shows that there are no significant differences due to genotype for any treatment at any time point (as evidenced by all overlapping standard error bars). **Inserts**) The pie-shaped shaded regions of the inserts highlight the region used to count the number of entrances. Sample sizes for each treatment group and genotype are shown on the bottom right.

A three-way ANOVA with Tukey Honest Significant Differences (TukeyHSD) test was carried out on the number of entrances by Genotype \* Treatment Group \* Training Session and the Genotype \* Treatment Group interaction. As expected Treatment Group had a highly significant effect on number of entrances [F(1, 197)= 116.354, p < 0.001]. The interaction between the effects of Genotype and Training Group was not significant [F(1, 197)= 0.025, p = 0.874], and there was no significant difference between WT consistent and FMR1-KO consistent (p = 0.92) or between WT yoked-consistent and FMR1-KO yoked-consistent (p = 0.69). These results are consistent with previous research showing that FMR1-KO does not impair the acquisition of spatial memory35.

#### Fail to replicate previous research showing that FMR1-KO impairs cognitive discrimination

In the Active Place Avoidance Task with conflict training, cognitive discrimination is evidenced by a reduction in the mean number of entrances after the location of the shock zone has been rotated 180°. To determine if cognitive discrimination was impaired in FMR1-KO, I measured the mean number of entrances into the shock zone in animals that were consistently trained or yoked (**Fig. 2.2**). A three-way ANOVA with Tukey Honest Significant Differences (TukeyHSD) test was carried out to determine the influence of Genotype \* Treatment Group \* Training Session and the Genotype \* Treatment Group interaction on the number of entrances during the active place avoidance task with the conflict training. As expected Treatment Group had a highly significant effect on number of entrances [F(1, 166)= 68.069, p < 0.001]. Unexpectedly, the interaction between the effects of Genotype and Training Group was not significant [F(1, 166)= 0.101, p = 0.7509]. Likewise, there was no significant difference between WT conflict and FMR1-KO conflict (p = 0.99) or between WT yoked-conflict and FMR1-KO yoked-conflict (p = 0.97). These results are inconsistent with previous research showing that FMR1-KO did impair cognitive discrimination35.

One explanation for the lack of cognitive impairment is a failure to tightly control the physical environment. The mean number of entrances into the shock zone by the yoked-consistent groups (mean entrances at T1= 13.5, see figure **Fig 2.2A**) is much lower than a previous study (mean entrances at T1= 21.5, see figure **Fig 1.2A**). This indicated that there might have been a confounding factor in the environment that cause mice to develop a preference for or aversion to one particular place in its environment.

#### Transcriptional response to constitutive gene knockdown

Given the lack of replication of a previous result showing impaired cognitive discrimination, I elected not to continue looking for the molecular underpinnings of impaired cognitive functions in the FMR1-KO mouse. Instead, I elected to identify molecular differences between the WT and FMR1-KO mouse when the external and internal representations of the world are, as far as I can tell, equivalent. Thus, I sequenced the transcriptome the CA1 subfield of the dorsal hippocampus from the mice in the yoked-consistent treatment group (**Fig 2.3A**).

RNA was isolated from a tissue sample (250 μm in diameter x 300 μm thickness) from the CA1 subfield of the dorsal hippocampus. Briefly, the transcriptome was constructed from mRNA-enriched Illumina libraries, transcript levels were estimated with Kalliso50 using the Gencode Mouse reference transcirptome51, the statistical significance of enriched genes and molecular functions was inferred using DESeq253 and GO\_MWU59, respectively. I identified 20 genes whose expression in the CA1 subfield was altered by constitutive knockdown of FMR1 (**Fig 2.3B**). About half of these genes are upregulated in FMR1-KO mice (*Apc2, Arel1, Brf1, Cry, Fibcd1, Grin1, Ncdn, Pnmal2, Prpf8, Sidt1, Slc8a2, Tnik,* and *Wipf3*) while the other half are down-regulated in FMR1-KO mice (*Cacna1g, Car4, Ccnd1, Cpne7, Dlx1, Efcab6, Fgfr1, Fmr1, Kcnt1, Mtus1, Plat, Serpina3n, Slc29a4, Sstr3,* and *Xbp1*).

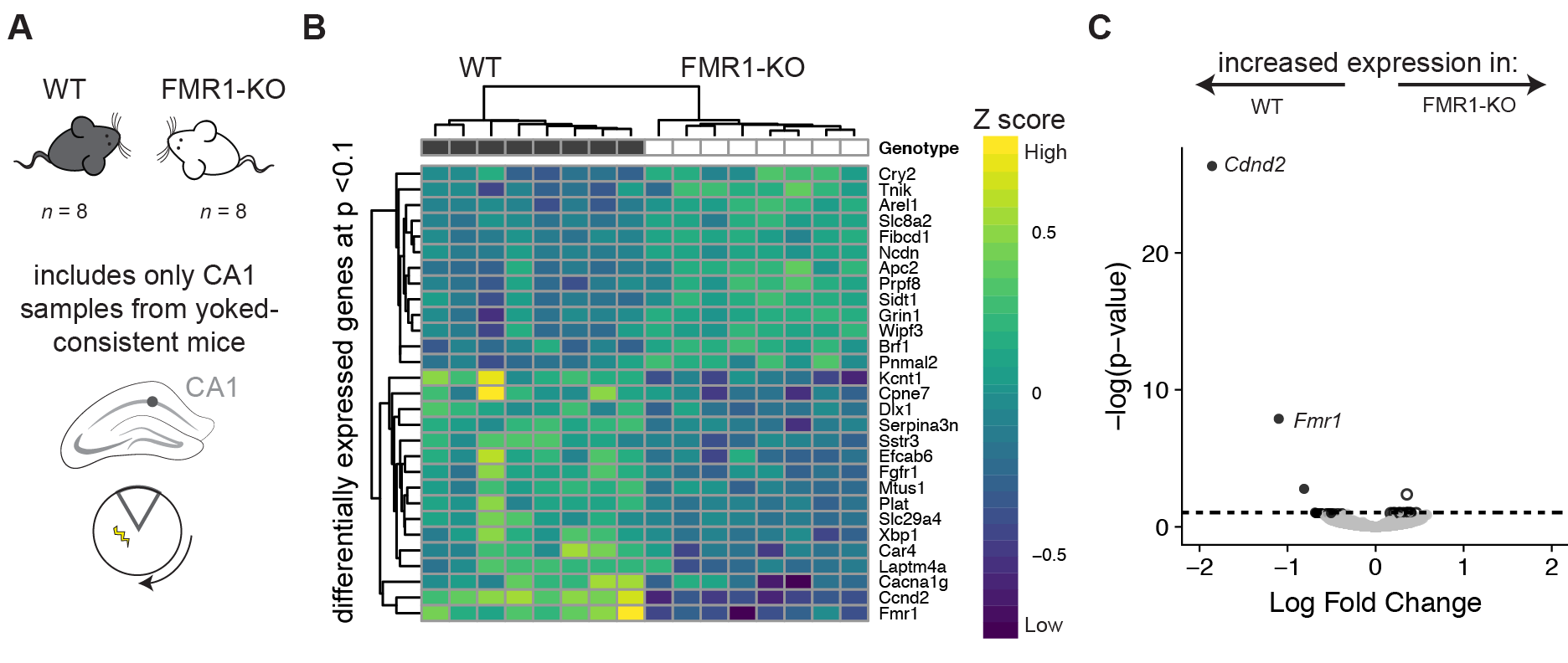


Fig. 2.3: WT and FMR1-KO show evidence of differential gene expression in the CA1

**A**) The sample size for RNA-sequencing is 8 WT and 8 FMR1-KO tissues from the CA1 subfield from only the consistent-yoke group. **B**) Hierarchical clustering of differentially expressed genes shows that only 13 genes are upregulated in response to FMR1KO while 16, including *Fmr1,* weredownregulated in the CA1 subfield of yoked-consistent mice. **C**) A volcano plot shows that expression *Ccnd2* and *Fmr1* are highly upregulated in WT compared to FMR1-KO mice. Genes with negative log fold change value are more highly expressed in WT (dark grey filled circles) while genes with a positive log fold change value are more highly expressed in the Fmr1 mice (open circles). Genes that are not significantly expressed are shown in light grey.

A volcano plot shows that expression *Ccnd2* and *Fmr1* are highly upregulated in WT compared to FMR1-KO mice (**Fig 2.3C**). Detection of a minimal *Fmr1* transcripts was expected given the method of gene knockout71. The protein encoded by *Ccnd2* is a highly conserved cyclin that regulates cyclin-dependent kinases. Co-localization of FMRP and *Ccnd2* mRNA has been shown in the developing brain72.

The gene encoding the pore-forming α1 subunit of the voltage-gated calcium channel (*Cacna1g*) is also among the list of down-regulated genes (**Fig 2.3B**). SNPs in *Cacna1g* and other calcium channel genes have been found to be associated with autism spectrum disorder73–75. Calcium binding and calcium transport are also notably on the list of molecular functions that are significantly downregulated in FMR1-KO mice (**Fig 2.4**).

Interestingly, the gene encoding the ionotropic glutamate Receptor (*Grin1*) was upregulated in WT mice (**Fig. 2.3B**). Using immuno-precipitation (IP) followed by microarray analysis of gene expression Brown et al. 2001 found that the association of Grin1a mRNA and the FMRP-bound large messenger-ribonucleoprotein (mRNP) complex was enriched in WT-IP compared with the FMR1-KO-IP76. Glutamate receptors are known to influence long-term synaptic modulation by stimulating the synthesis of synaptic proteins, including postsynaptic density 95 (PSD-95) and FMRP77–80.



Fig. 2.4: Downregulation of ion transport in the the the CA1 subfield

Down-regulation of ion channel binding, receptor binding, calcium binding, metal ion membrane transport, calcium ion transmembrane transporter, delayed rectifier potassium channel, channel, cation channel, channel regulator, calmodulin binding, PDZ domain binding, structural molecular, and structural constituent of ribosome. On the plot, different fonts are used to indicate significance (bold: p < 0.01, regular: p < 0.05) and color indicates enrichment with either up (red) or down (blue) regulated genes. The fraction next to GO category name indicates the fraction of "good" genes that exceed the p-value cutoff.

Calmodulin-dependent protein kinases are abundant in the postsynaptic density. Overexpression of can dramatically increase synaptic strength is a major regulator of long-term potentiation (LTP) and long-term depression (LTD) processes81,82. Studies have found that knockout mice without Calmodulin-dependent protein kinases CaMKIIA demonstrate a low frequency of LTP83. However, even thought calmodulin-binding activity is reduced in the FMR1-KO mice (**Fig. 2.4**) I did not observe a reduction in synaptic strength at the CA3-CA1 synapse (as measured by maximum fEPSP slope) with avoidance learning or with FMR1-KO **(Fig. 2.5**). The interaction between Genotype and Training Group was not significant [F(3, 48)= 0.304, p = 0.822], and Genotype alone was not significant [F(1, 48)= 0.133, p = 0.738]. Thus, even thought Calmodulin activity is reduced, synaptic strength is not diminished.



Fig. 2.5: CA3-CA1 synaptic strength is not altered by genotype or place avoidance training.

I did not observe reduction in synaptic strength at the CA3-CA1 synapse (as measured by maximum fEPSP slope) with avoidance learning or with FMR1-KO. The interaction between Genotype and Training Group was not significant, and Genotype alone was not significant. Violin plots are colored by treatment group and shaded according to genotype.

### Discussion

To better understand the effects silencing the Fmr1 gene, we conducted an integrative analysis of avoidance, cognitive discrimination, and hippocampal gene expression WT and FMR1-KO mice. Importantly, I confirmed that FMR1-KO does not impair learning or memory; however, I failed to replicate previous research showing that FMR1-KO impairs cognitive discrimination. This lack of difference in the expression of avoidance behavior may be a result of a “noisy” environment that interfere with the task, or it could indicate that cognitive deficits are not very robust.

Because of this inconsistent behavioral results, I focused on the naive animals (those without cognitive training to increase learning and memory) for subsequent analysis of the transcriptional effects of gene-knockdown better. I found that only 20 genes in the CA1 hippocampal subfield differed significantly between yoked-consistent WT and yoked-consistent FMR1-KO mice. Among the affected genes were those that genes with mRNA or protein products known to interact with FMRP directly or are known autism disorder risk genes. These data suggest that even thought FMRP is typically thought to disrupt translation, FMRP knock out also has effects on the upstream process of gene expression.

These results suggest that knockdown causes dysregulation of multiple genes that contribute to autism spectrum disorder, but that this disruption does not have significant effects on a hippocampal-dependent place avoidance condition task. This research provides a unique contribution to diverse fields of studying including hippocampal function and human health research.

### Acknowledgments

I am grateful to Brett Mensh, Konrad Kording, and the PLOS One Computational Biology Community for publishing a beautiful paper and graphical abstract for structuring papers. I'm grateful for discussions with Dr. James Noonam's lab at Yale University. I thank Becca Young Brim, Caitlin Friesen, Tessa Solomon Lane, Mariana Rodriguez, Eric Brenner, and Issac Miller-Crews for helping advice regarding on figures and oral presentations of this research. I thank members of the Boris Zemelman, Laura Colgin, and Misha Matz for helpful discussions. The bioinformatic workflow was inspired heavily by workshops and online resources from the BioITeam (https://wikis.utexas.edu/display/bioiteam), the Center for Computational Biology and Bioinformatics (http://ccbb.utexas.edu), and Software Carpentry Curriculum on the Unix Shell, Git for Version Control, and R for Reproducible Research84–86. This work is supported by a Society for Integrative Biology (SICB) grant and a UT Austin Graduate School Continuing Fellowship to RMH; a generous gift from Michael Vasinkevich to AAF; NIH-NS091830 to JMA, IOS-1501704 to HAH; NIMH-5R25MH059472-18, the Grass Foundation, and the Helmsley Charitable Trust.

## Chapter 3: Analysis of hippocampal transcriptomic responses to technical and biological perturbations[[3]](#footnote-3)

### Abstract

Single-neuron gene expression studies may be especially important for understanding nervous system structure and function because of the neuron-specific functionality and plasticity that defines functional neural circuits. Cellular dissociation is a prerequisite single-cell sequencing analysis, but the extent to which this process alters the molecular activity in neural tissues has not been determined. This information is necessary for interpreting the results of experimental manipulations that affect neural function such as learning and memory. Here, I quantified the transcriptomic response to cellular dissociation in tissue samples from the dentate gyrus (DG), CA3, and CA1 hippocampus subfields relative to tissues that have been homogenized for whole tissue analysis. I determined that 1% of the hippocampal tissue transcriptome is altered by the process of cellular dissociation. Next, I compared this 'dissociation-induced' gene-expression response to previously described changes in hippocampal subfield expression in response to stressful experience and cognitive training. I found that chemical dissociation response is greater than the transcriptomic response to a stressful experience but weaker than cognitive training. Finally, I compared our findings against a database of gene expression for single-cell population expression in DG, CA1, and CA1. This meta-analysis identified genes whose expression varies by subfield and by treatment. The important contribution of this paper to the literature is that I begin to describe concordant and discordant effects of technical and behavioral manipulations which will inform the design of future single-neuron transcriptome studies to facilitate a more comprehensive understanding of brain function.

### Introduction

Nervous systems are comprised of diverse cell types that include neurons, glia, and vascular cells, each serving distinct functions and thus expressing different genes. Consider the hippocampus, a structure central for spatial navigation and the processing of event memory in the mammalian brain. To date, distinct aspects of navigation and memory processing have been firmly correlated to activity of particular cellular subfields within the hippocampal formation. This subfield-specific understanding of hippocampal function has led to the notion that cells within a given subfield are homogeneous in their molecular blueprint and perform the same function. However, even within the anatomically-defined subfield of CA1, there are identifiable subclasses of pyramidal cells that belong to distinct functional circuits87,88. This diversity is even greater when I consider that specific cells within a functional class can be selectively altered by neural activity in the recent or distant past. For example, only a third of the pyramidal cells of the superficial CA1 sub-layer is expected to be activated by experience, and only a subset of those will undergo synaptic strengthening to trigger further gene expression changes89,90. All this diversity implies distinctive gene expression, very likely at the level of single neurons, and such considerations may curtail interpretations of gene expression studies that use mixtures of cells or microdissected tissue samples.

Fortunately, recent advances in tissue processing and sequencing technologies have allowed genome-wide analysis of single cells in the context of brain and behavior studies25,91,92. These approaches have led to systems-level insights into the molecular substrates of neural function, along with the discovery or validation of candidate pathways regulating physiology and behavior. While the complexity of some tissues complicates the interpretation of transcriptome data collected from samples containing hundreds to tens of thousands of cells representing numerous cellular subclasses at different levels of diversity, difficulties with interpretation can be minimized by careful experimental design governing both data collection and data analysis. To complement this effort, and optimize experimental designs, it is necessary to understand the extent to which the treatment of tissue samples before transcriptome analysis might confound interpretation of the results.

I examined the effect of cellular dissociation on the transcriptomes of specific hippocampal subfields (CA1, CA3, and DG) by comparing tissue homogenization (as a control) and cellular dissociation protocols. Next, I compared this 'dissociation-induced' gene-expression effect to the effect of prior stressful experience that accompanies many protocols to assess learning, memory and innate behaviors, and cognitive training on hippocampal subfield gene expression. Finally, I compared our results against a database of gene expression for single-cell population expression in DG, CA1, and CA1 to further validate the patterns of gene expression that I identified. Knowing how technical perturbations influence the ability to detect the molecular signature of differences in neural and behavioral variables is a significant step in calibrating the ability to mechanistically understand hippocampal function. In addition to understanding the impact of cell dissociation and stressful experience on hippocampus gene expression, the present findings allow evaluating the extent to which gene expression profiles of heterogeneous tissue samples compared with single neuron gene expression profiles. The important contribution of this paper to the literature is that I begin to describe concordant and discordant effects of technical and behavioral manipulations which will inform the design of future single-neuron transcriptome studies to facilitate a more comprehensive understanding of brain function.

### Methods

All animal care and use complies with the Public Health Service Policy on Humane Care and Use of Laboratory Animals and were approved by the New York University Animal Welfare Committee.

#### Sample processing

A 1-year old C57BL/6J mouse was taken from its cage, anesthetized with 2% (vol/vol) isoflurane for 2 minutes and decapitated. Transverse 300 μm brain slices were cut using a vibratome (model VT1000 S, Leica Biosystems, Buffalo Grove, IL) and incubated at 36°C for 30 min and then at room temperature for 90 min in oxygenated artificial cerebrospinal fluid (aCSF in mM: 125 NaCl, 2.5 KCl, 1 MgSO4, 2 CaCl2, 25 NaHCO3, 1.25 NaH2PO4 and 25 Glucose)46,47. Two adjacent tissue samples were collected from CA1, CA3, and DG, respectively in the dorsal hippocampus by a punch tool (0.25 mm, P/N: 57391; Electron Microscopy Sciences, Hatfield, PA).

The ‘control samples' were homogenized by pestle in the homogenization solution provided by Maxwell 16 LEV RNA Isolation Kit (Promega, Madison, WI). The ‘cellular dissociation samples' were incubated for 75 minutes in aCSF containing 1 mg/ml pronase at room temperature, vortexed, centrifuged, and terminated by replacing aCSF containing pronase with aCSF. The sample was then vortexed, centrifuged, and gently triturated by 200-μl pipette tip twenty times in aCSF containing 1% FBS. This sample was centrifuged and added to the homogenization solution using the Maxwell 16 LEV RNA Isolation Kit (Promega, Madison, WI). Frozen samples were sent arena to The University of Texas

for RNA isolation as per the manufacturer instruction (Maxwell 16 LEV RNA Isolation Kit (Promega, Madison, WI). RNA libraries were prepared by the Genomic Sequencing and Analysis Facility at the University of Texas at Austin using the Illumina HiSeq platform.

#### Bioinformatics

Raw reads were processed and analyzed on the Stampede Cluster at the Texas Advanced Computing Facility (TACC). Quality of the data was checked using the program FASTQC. Low-quality reads and contaminating adapter sequences were removed using the program Cutadapt52. I obtained an average of 5 million reads for each hippocampal tissue sample and quantified the expression representing 22,485 genes in the mouse reference transcriptome M11. I used Kallisto50 for pseudo-alignment of reads and transcript counting using the-the Gencode M11 mouse transcriptome51. Transcript counts from Kallisto were imported into R44 and aggregated to yield gene counts using the ‘gene’ identifier from the Gencode transcriptome. I used DESeq253 to normalize and quantify gene counts with a false discovery corrected (FDR) p-value < 0.1. The DESeq2 model included treatment group and hippocampal subfield. Principal component analyses (PCA) and hierarchical clustering by correlation were used to describe broad patterns of variation. I used GO\_MWU59 using a –log(p-value) as a continuous measure of significance to identify gene ontology categories that are significantly enriched with either up- or down-regulated genes for a given two-way contrast. I used ggplot258, cowplot57, pheatmap60, viridis61, and colorblindr62 to make figures that are color-blind friendly.

#### Meta-analysis

I downloaded the gene counts from the Cembrowski et al. (2016) dataset (NCBI GEO: GSE74985) and the Harris et al. (2017) dataset (NCBI GEO: GSE99765). Briefly, the Cembrowski dataset contains hippocampal gene expression data for pools of 112 ± six cells for each of 5 cell types (CA1, CA2, and CA3 pyramidal neurons and DG mossy and granule cells) from behaviorally naive, transgenic mice that express a fluorescent protein label in the specific cell types. Fluorescently-labeled DG, CA3, and CA1 neurons from dorsal and ventral hippocampi were manually sorted (DG granule cells from Rbp4-Cre KL100, CA3 pyramidal cells from Mpp3-Cre KG118, and CA1 pyramidal cells from Vipr2-Cre KE2 mice.) The Harris et al. dataset contains DG, CA3, and CA1 transcriptomes from homecage mice, cognitively trained mice, and yoked control mice. The homecage mice were not examine in previous chapters thus a novel comparison between homecage and yoked controls can be made. The same statistical analyzes described in the previous section were run on these Cembrowski and Harris datasets to identity differentially expressed genes (FRD p-value < 0.1) in subfields of the hippocampus. I created a binary (0 or 1) list of genes that were significant or not for a GO enrichment analysis using Fisher's exact test to determine if GO categories are overrepresented among the significantly expressed genes.

#### Archival of data, code, and figures

I archived the raw sequence data and intermediate data files in NCBI's Gene Expression Omnibus Database (accession numbers GSE99765 and GSE100225). The data, code, and results are openly available on GitHub (<https://github.com/raynamharris/DissociationTest)>, with an archived version at the time of publication available at Zenodo93.

### Results

#### The effects of cellular dissociation

I identified 162 genes that were differentially expressed between the control and dissociated samples, 331 genes that were differentially expressed genes (DEGs) between any of the three hippocampus subfields, and 30 genes were shared between both sets of differentially expressed genes at p-value < 0.05 (**Fig. 3.1A, B**). A hierarchical clustering analysis of all differentially expressed genes does not give rise to distinct clusters that are separated by subfield or method; however, when examining the control, homogenized samples alone (identified with light grey boxes), the three subfields form distinct clusters, while the dissociated samples do not cluster by subfield (**Fig. 3.1C)**. A principal component analysis of normalized gene counts reveals that PC1 accounts for 40% of the variation and visually separates the DG samples from the CA1 and CA3 samples (**Fig. 3.1D**). To confirm statistical significance of this visual pattern, I conducted a two-way treatment x region ANOVA and confirmed a significant effect of region (F2,11= 17.69; p = 0.0004). P*ost hoc* Tukey tests confirmed CA1 = CA3 < DG. The effect of treatment and the treatment x region interaction were not significant. PC2 accounts for 22% of the variation in gene expression and varies significantly with treatment (F1,12=6.125; p = 0.03) but not by region or the treatment x region interaction. None of the other PCs showed significant variation according to either region or treatment.

#### The effects of stressful experience

I examined the effect of stressful experience, which is a common confound of behavioral manipulations because animals in different experimental groups often experience varying levels of stress, especially if the experimental procedure is not intentionally stressful. I identified 0 genes that were significantly expressed between samples from the home cage and shocked mouse samples; 1669 genes were significantly differentially expressed between any of the three brain regions at p-value < 0.05 (**Fig. 3.2A, B**). Hierarchical clustering of the differentially expressed genes gives rise to three distinct clusters corresponding to the three subfields, with CA1 (purple) and CA3 (green) being more similar to one another than to DG (orange), whereas the effects of the stress manipulation were not distinctive (**Fig. 3.2C**).

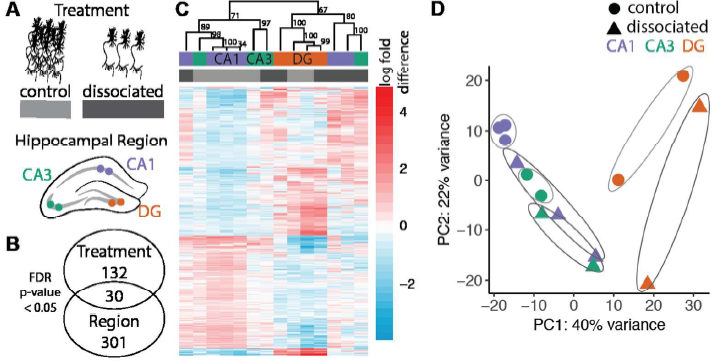


Fig. 3.1: The effect of cellular dissociation

**A)** From a single female mouse, I collected 2 CA1, CA3, and DG hippocampal tissue samples. One sample was subjected to a cellular dissociation treatment (dissociated) whereas the control samples (control) were standardly homogenized. **B)** I identified 162 dissociation-induced changes in gene expression, 331 genes with region-specific expression patterns, and 30 genes differentially expressed by both region and treatment (FDR p-value < 0.05). **C)** Hierarchical clustering separates the hippocampal subfields of the homogenized samples (light gray) but not the dissociated samples (dark gray). **D)** PC1 accounts for 40% of all gene expression variation and, by inspection, separates the DG samples from the CA1 and CA3 samples. PC2 accounts for 22% of the variation in gene expression and varies significantly with treatment. Ellipses are hand-drawn.

Next, I conducted a principal component analysis of all the genes that were measured. PC1 accounts for 31% of the variation and by inspection, separates the DG samples from the CA1 and CA3 samples (effect of subfield: F2,15= 42.89; p < 0.001; **Fig. 3.2D**). P*ost hoc* Tukey tests confirmed CA1 = CA3 ≠ DG. The effects of stress and the stress x subfield interaction were not significant. PC2 accounts for 18% of the variation and varies significantly between CA1 and CA3 and CA1 and DG (effect of region: F2,15= 11.41; p < 0.001; Tukey tests: CA1 ≠ DG = CA3). The effects of stress and the stress x region interaction were not significant. PC3 accounts for 15% of the variation and also explains some subfield-specific differences (effect of region: F2,15= 6.315; p < 0.01; Tukey tests: CA1 = DG ≠ CA3), whereas effects of stress and the stress x region interaction were not significant. PC4 is also influenced by region (F2,15= 6.315; p = 0.0102; Tukey tests: CA1 ≠ CA3. PC5 did not account for any significant differences according to region or treatment. PC6 significantly accounted for variance associated with the effect of a stressful experience (F1,16> 4.774; p’s < 0.04).

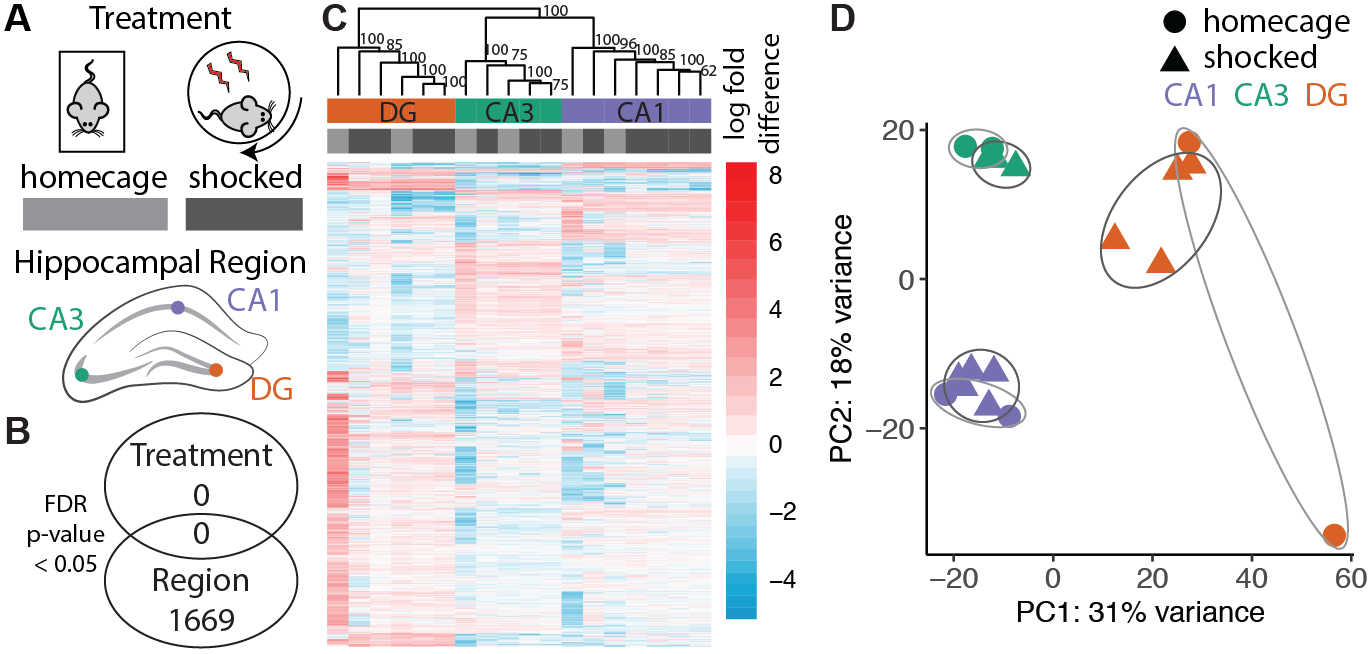


Fig. 3.2: The effects of a stressful experience

A) I compared CA1, CA3, and DG tissue samples from control mice taken directly from their home cage to mice that were subjected to a mild foot shock. B) I identified 0 genes that responded to treatment, and 1669 genes that were differentially regulated across regions of the hippocampus (FDR p-value < 0.05). C) Hierarchical clusters groups samples by brain region but distinct treatments clusters are not present. D) PC1 accounts for 31% of the variation and visually separates the DG samples from the CA1 and CA3 samples. PC2 accounts for 18% of the variation and distinguish the three subfields. Ellipses were hand-drawn.

#### The effects of cognitive training

I identified that 423 genes were differentially expressed between the yoked control and cognitively trained animals, 3485 genes that were differentially expressed across subfields, and 324 showed an interaction at FDR p < 0.05 (**Fig. 3.3A, B**). I see a large effect of brain region on gene expression, with 20% of detectable genes being differentially expressed between one or more brain-region comparisons (3485 differentially expressed genes /17320 measured genes). This is an order of magnitude greater than the 2% of the transcriptome that changed in response to learning (423 DEGs /17320 genes measured). Hierarchical clustering of the differentially expressed genes separates samples by both subfield and treatment (**Fig. 3.3C**). A principal component analysis of all gene expression data revealed that brain region explains 75% of the variance in the data (**Fig. 3D**). PC1 accounts for 56% of the variance and distinguishes DG from the Ammon’s horn samples (effect of subfield: F2,19= 226.1; p < 0.001; Tukey tests: DG ≠ CA3 = CA1 ), but the effects of training and the training x region interaction were not significant. PC2 accounts for 19% of the variance and distinguishes the three subfields (effect of region: F2,19= 255.3; p < 0.001; Tukey tests: DG ≠ CA3 ≠ CA1). PC3 and PC4 indicate a significant influence of cognitive training (PC3: F1,20=7.451; p = 0.01,) and (PC4: F1,20=10.11; p = 0.005), but no significant effects of region and the region x treatment interaction.

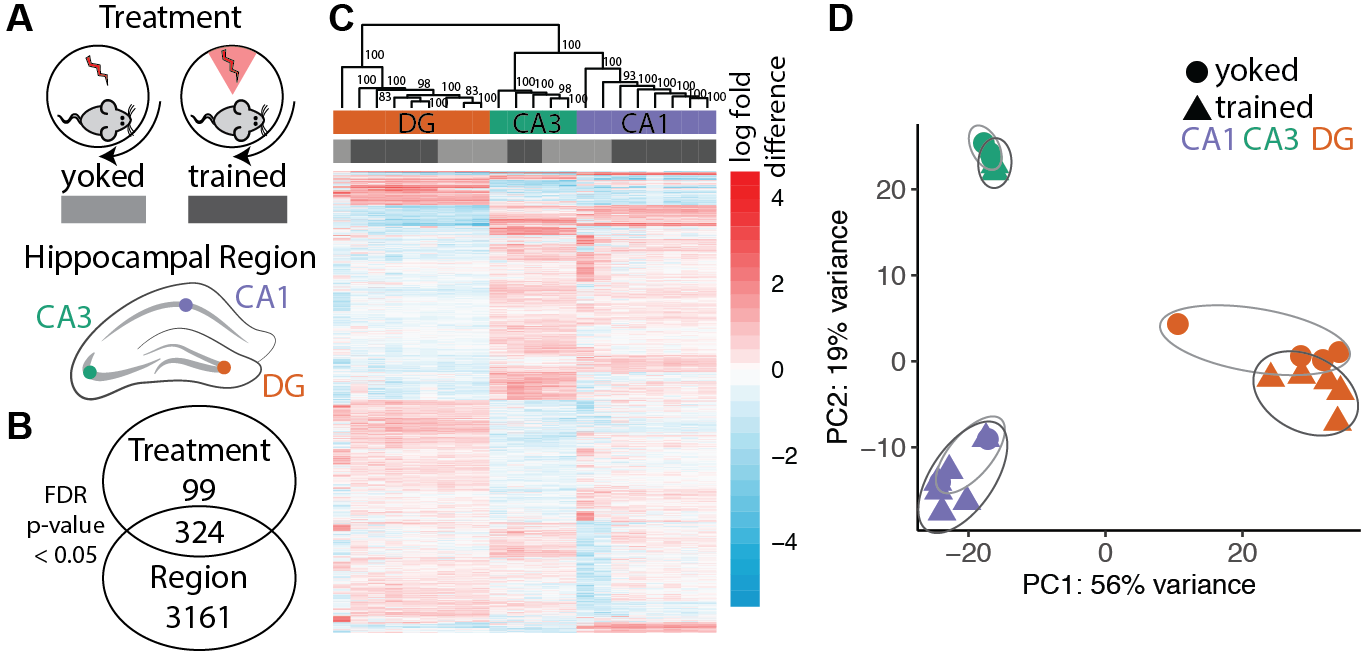


Fig. 3.3: The effects of learned avoidance behavior

A) Mice used in this study were either subjected to random but mild foot shocks (control) or subjected to mild foot shocks conditioned with spatial cues (trained). Tissue samples were collected from CA1, CA3, and DG. B) I identified only 423 genes that were significantly expressed according to cognitive training and identified 3485 genes that were differentially expressed between any of the three brain regions (FDR p-value <0.05). C) Hierarchical clustering of the differentially expressed genes gives rise to three distinct clusters corresponding to the three brain regions, with CA1 and CA3 being more similar to one another than to DG. D) A principal component analysis of all genes in the analysis (regardless of the level of significance) shows that PC1 accounts for 56% of the variation and visually distinguishes the DG samples and the CA1 and CA3. PC2 accounts for 19% of the variation and separates all three subfields. Ellipses were hand-drawn.

#### Identifying unique and general patterns of hippocampal genomic plasticity

Next, I examined the overlap in genomic response to the technical and biological perturbations. I identified three specific genes that responded to both cellular dissociation and cognitive training: *Grin2a,* *Epha6,* and *Ltbp3 (***Fig. 3.4A**).There was no overlap in differentially expressed genes compared to the cellular dissociation treatment *(***Fig. 3.4A**). Then, analyzed gene ontology at 5% FDR significant in each of the data sets to identify the molecular function of genes that changed in response to cellular dissociation (**Fig. 3.4B**) or cognitive training (**Fig. 3.4C**). The process of cellular dissociation results in a significant upregulation of molecular processes related to ribosomal activity, rRNA binding, oxidoreductase activity, and proton transport, while it caused a down-regulation of ligase and helicase activity (**Fig. 3.4B**).

Fig. 3.4: Overlapping responses to technical and biological perturbations

A) The number of genes that responded to chemical dissociation (163 genes), a stressful experience (0 genes), and cognitive training (423 genes). The three genes that respond to both technical and biological perturbation are *Epha6*, *Grin2a*, and *Itbp3*. B, C) The molecular function of gene ontology (GO) categories that are significantly enriched with either up- or down-regulated genes in response to cellular dissociation (B) or cognitive training (C). The top 10 most significant GO terms are visualized, each with a p-value < 0.001. The fraction next to GO category name indicates the fraction of genes in that category that survived a 10% FDR threshold for significance. Zero terms survived a 10% FDR threshold in response to a stressful experience.

The GO analysis detected no Molecular Function GO terms in the significantly overrepresented genes in response to the stressful experience. Cognitive training resulted in a significant upregulation of molecular processes related to glutamate receptors, signal transduction, calcium binding, and membrane transport, and it resulted in a significant downregulation of ribosomal activity, oxidoreductase activity, mRNA binding, and proton transport (**Fig. 3.4B**).

Fig. 3.5: Meta-analysis of primary and public data

A) This Venn diagram shows the overlap in brain-region specific gene expression across all four experiments (cellular dissociation, stressor habitation, cognitive training, and a public dataset examining subfield comparisons). Using this approach, I identified 146 genes that were differentially expressed between any two subfields of the hippocampus in all four experiments. B) Those 146 provide robust brain-region specific markers of gene expression belong to molecular function and cellular compartment GO.The top 10 most significant GO terms are visualized, each with a p-value < 0.05. The fraction next to GO category name indicates the fraction of genes in that category that survived a 10% FDR threshold for significance.

The gene ontology analysis identified 91 significant GO terms in response to cognitive training. Among the top 10 are glutamate signaling and membrane transport systems and a downregulation of oxidoreductase and ribosomal activity (**Fig. 3.4C**). Notably, learning induces a downregulation of a structural constituent of ribosomes and oxidoreductase, which were both up-regulated in response to cellular dissociation (**Fig. 3.4B, C**).  Using the public Cembrowski et al. 2016 dataset, I identified 10,751 genes that were differentially expressed between hippocampal sub-regions **(Fig. 3.5A)**. Using a meta-analysis of public the primary data, I identified 146 genes that showed robust subfield-specific gene expression patterns **(Fig. 3.5A)**. Those 146 genes are enriched in cellular compartments related to synapses and molecular functions related to calcium signaling, GTP exchange, and proteoglycan binding **(Fig. 3.5B)**.

### Discussion

The primary purposes of this study were 1) to test whether analysis of gene expression in hippocampus subfields is changed by tissue preparation procedures (cellular dissociation versus homogenization) and 2) to evaluate the effects of a stressful experience relative to cognitive training on analysis of gene expression. The work was designed to evaluate the extent to which technical (i.e., cellular dissociation) and biological confounds (i.e., stressful experience) can impact efforts to assess the transcriptomic response to cognitive processes. This is potentially important because it is increasingly necessary to conduct molecular analyses of single-cells.

Hippocampal subfield differences are well known 29–32. Across the three experiments with different treatments, the identity of the hippocampal subfield explained between 40 and 75% of all the variation in gene expression across samples **(Fig 3.1D, 2D, 3D)**. The samples that were subjected to cellular dissociation show the least amount of region-specific variation, suggesting that this process might alter the genes that typically distinguish the hippocampal subfields from one another. On the other hand, the Cembrowski et al., (2015) study identified a larger number of genes with subfield specificity; this is likely due to the cell sorting process that generates a relatively homogenous rather than a heterogeneous population of neurons. These results indicate that cell-type specific differences may be recovered by sorting cells from very heterogenous tissues into more homogeneous populations of cells.

Across the datasets, many genes consistently or robustly demonstrate hippocampal subfield specificity in their expression (**Fig. 3.5B**). The meta-analysis of the primary data and the public Cembrowski et al., (2016) data identified 146 genes that could potentially serve as robust subfield specific markers. The molecular functions of these potential marker genes are diverse, related to calcium channel regulation, proteoglycan binding, and guanyl-nucleotide (GTP) exchange, as well as cellular compartment categories related to the synapse and the postsynaptic density. This suggests that the phenotypic and functional differences amongst DG, CA3, and CA1 neurons may be driven or influenced by subfield differences in gene expression.

Concerning the effects of cellular dissociation of hippocampal gene expression, I found that 0.9% (162/16,709) of the genes measured changed in response to cellular dissociation (**Fig. 3.1B**). This is smaller than the 2.4% (423 /17,320) change I detected in response to cognitive training (**Fig. 3.2B**). The stressful experience produces a negligible response (i.e., no significant genes or GO terms were detected), indicating that the mild stress that likely accompanies most behavioral tasks does not have a lasting influence on hippocampal gene expression (**Fig. 3.2B**).

The extent to which cellular dissociation and unintended stress impact the expression of particular genes and signaling pathways limits the feasibility of investigating how genes contribute to behavior and other responses to organismal manipulations. I found that *Grin2a* responded to both cellular dissociation and cognitive training (**Fig. 3.4A**). *Grin2a* encodes subunits of N-methyl-D-aspartate (NMDA) type ionotropic glutamate receptors that are crucial for numerous cellular functions throughout the brain, including hippocampus-dependent synaptic plasticity and learning 94,95. Accordingly, care should be taken when studying the role of glutamate and *MAPK* signaling in combination with cellular dissociation techniques. *Epha6* and *Ltbp3* also responded to both cellular dissociation and cognitive training (**Fig. 3.4A)**. *Epha6* is involved with the MAPK-Erk signaling pathway. *Ltbp3* participates in binding calcium ions and shows altered gene expression in a mouse model of Alzheimer’s Disease 96.

I can look beyond the specific genes and examine which pathway responses are concordant or discordant to multiple treatments. In this case, I saw upregulation of ribosomal activity and rRNA biding in response to cellular dissociation, but I saw an opposing downregulation in ribosomal activity and mRNA binding in response to cognitive training (**Fig. 3.4B, 4C**). This suggests that cellular dissociation activates a general transcriptional response whereas cognitive training reduces the transcription of particular protein-coding genes. This demonstrates the possibility that such an interaction, in this case a downregulation in response to cognitive training could be overshadowed by technical artifacts if hippocampus tissue is first subjected to the cellular dissociation required for single-cell or single cell population investigations.

I found no detectable transcriptional response in the CA1, CA3, or DG following the stressful experience (**Fig. 3.2B**). The shock experience I used causes a substantial increase in plasma corticosterone levels, comparable to exposure to predator threat, that is observed after the initial shock exposure session but is absent 24-h later after the second training session 64. Our findings support the use of either home cage or shock-yoked animals as controls for active place avoidance training experiments. In the case of the homecage controls that do not experience shock, their stress response is indistinguishable from the trained mice, but their sensory experience is very different. In contrast, the shock-yoked mice have identical sensory experience as the experimental mice, but they experience stress that the experimental animals do not 64. It may be that untrained mice are ideal controls because they would have the identical experience of the environment as experimental mice, except at the brief 500 ms moments of shock that account for ~3% of the task assuming 20 shocks in 600 s. Depending on the question one control may be preferable over the others, but as demonstrated here, when assessed 24 h after the training experience, they appear to be equivalent regarding their gene expression profiles (**Fig. 3.2**).

### Conclusions

Many factors contribute to variation in gene expression. I set out to identify the extent to which the process of cellular dissociation – which allows for single-cell analysis of neurons – has an appreciable effect on our ability to detect biologically meaningful variation in hippocampal gene expression. I conclude that there are specific dissociation-induced and cognitive training-induced changes in gene expression that are largely non-overlapping. It is encouraging that the overlap between cellular dissociation and cognitive training is small, indicating that these technical and biological processes affect different transcriptional processes. It is also gratifying to know that the stressful experience had no substantial effect on hippocampal gene expression because its generalization to other tasks will allow for using behavioral control groups and behavioral manipulations that also induce modest, potentially confounding stress. These findings provide insight into how cellular and biological manipulations influence gene expression. Through meta-analysis with open-access data, I have identified a subset of robust sub-region specific markers of gene expression profiling. Further research is needed to uncover the influence of other variables on variation in hippocampal gene expression.

### Acknowledgments

This work was carried out as part of the Neural Systems and Behavior Course Directors Research Program run by HAH and AAF at the Marine Biological Laboratory. I thank The Jackson Laboratory for generously donating mice; Promega Corporation for generously donating molecular supplies and use of the MaxwellⓇ; and other vendors for donating materials to perform the electrophysiology. I thank members of the Hofmann lab, Fenton Lab, Boris Zemelman, Laura Colgin, and Misha Matz for helpful discussions. I am grateful to Dennis Wylie for insightful comments on earlier versions of this manuscript. The bioinformatic workflow was inspired heavily by Center for Computational Biology’s Bioinformatics Curriculum and Software Carpentry Curriculum on the Unix Shell, Git for Version Control, and R for Reproducible Research84–86. This work is supported by a Society for Integrative Biology (SICB) grant and a UT Austin Graduate School Continuing Fellowship to RMH; a generous gift from Michael Vasinkevich to AAF; NIH-NS091830 to JMA, IOS-1501704 to HAH; NIMH-5R25MH059472-18, the Grass Foundation, and the Helmsley Charitable Trust. The authors declare no competing interests.

## Concluding Thoughts and Future Directions

Outline

1. reflect on the overall knowledge gained from this body of work on transcriptional plasticity in the hippocampus and its role in avoidance learning.
2. described the outcomes of my efforts to improve interdisciplinary graduate training programs and increase the prevalence of reproducible research and open science practices.
3. conclude with an outlook to the future.

## Appendix 1: Candidate gene expression and synaptic plasticity do not explain differences in WT and FMR1-KO mice in avoidance behavior[[4]](#footnote-4)

### Abstract

In chapter 2, I ask if avoidance behaviors are regulated by molecular changes that alter the synaptic activity in a hippocampal circuit. My null hypothesis is that a conditioned place avoidance response is not due to long-term molecular changes at pre- and post-synaptic levels of a neuronal circuits. One alternative hypothesis is that changes in gene expression of candidate proteins whose activity underline functional changes in synapses subservient of memory that have long been studies will underlie changes in behavior. A type II statistical error would be that something other than knowledge acquisition and memory alter cell and molecular biology. I describe our first collaborative and integrative studying analyzing behavior, electrophysiology, and candidate gene expression in mice in a pop-up laboratory at the Marine Biological Laboratories. In this study, I fail to reject the null hypothesis that synaptic plasticity at the CA3-CA1 synapse and molecular activity in CA3 and CA1 do not regulate avoidance learning. However, I concluded that this outcome was the result of technical limitations (setting up a new lab takes time) and mishaps (using unfamiliar protocols) and not of biological significance. Therefore, I set out to repeat the experiment on a larger scale.

### Introduction

Many molecular events are considered relevant for the acquisition and storage of memory including pre-synaptic neurotransmitter release and post-synaptic expression of glutamate receptors, kinases, transcription factors, and protein synthesis regulators. Signaling through the glutamate receptors has been implicated in long-term synaptic modulation77–79. Activated mGluRs stimulate the synthesis of synaptic proteins postsynaptic density 95 (PSD-95) and fragile X mental retardation protein (FMR1)80. FMR1 has been shown to contribute the control to at least three attributes of brain functionality: neuronal network excitability, epileptogenesis, and place learning 78. mGluR dysfunction has been linked to neurodegenerative disorders like Alzheimer’s disease (AD) and Fragile X syndrome97. PKMζ promotes the opening of chromatin by interfering with a form of epigenetic regulation, possible through histone acetylation or methylation. Protein kinase M zeta (PKM zeta), has been found to maintain late LTP in hippocampal slices66.

Is avoidance behavior generated by molecular changes that alter the activity of synapses in neural circuits? Our null hypothesis is that the acquisition of knowledge and subsequent avoidance behavior is not based on long-term molecular changes at pre- and post-synaptic levels of neural circuits. The alternative hypothesis is that changes in gene expression of candidate proteins whose activity underline functional changes in synapses subservient of memory that have long been studies will underlie changes in behavior.

To answer this question, I collaborated to examine the neuromolecular and physiological correlates of memory formation. I measured avoidance behavior as a proxy for learning, CA3-CA1 synaptic plasticity, and candidate gene expression in the CA1 to estimate transcriptional plasticity (**Fig. 1**). Then, I analyzed the data from each level of analysis separately to determine large effects of a behavioral manipulation that induces avoidance behavior. Then, I repeated the study to look at the effect of a genetic manipulation of FMRP. Due to some unforeseen circumstances, I were not able to process the WT and FMR1 samples at the same time. Therefore, many factors such as laboratory location and tissue collector changed, so the genetic manipulation is confounded. Within each genotype, I find no differences in gene expression or synaptic plasticity, nor were there any correlations between levels. Thus I do not reject the null hypothesis. Therefore, a behavioral manipulation can result in a behavioral plasticity in the absence of both synaptic plasticity and transcriptional plasticity.

### Methods

All experimental animal procedures complied with NIH and institutional guidelines and were approved by the New York University Animal Welfare Committee and the Marine Biological Laboratory Institutional Animal Care and Use Committee. For this research, I set up a collaborative laboratory and office space at the Marine Biological Laboratories where all four co-authors worked together during June and July 2013 (**Fig. A1**).

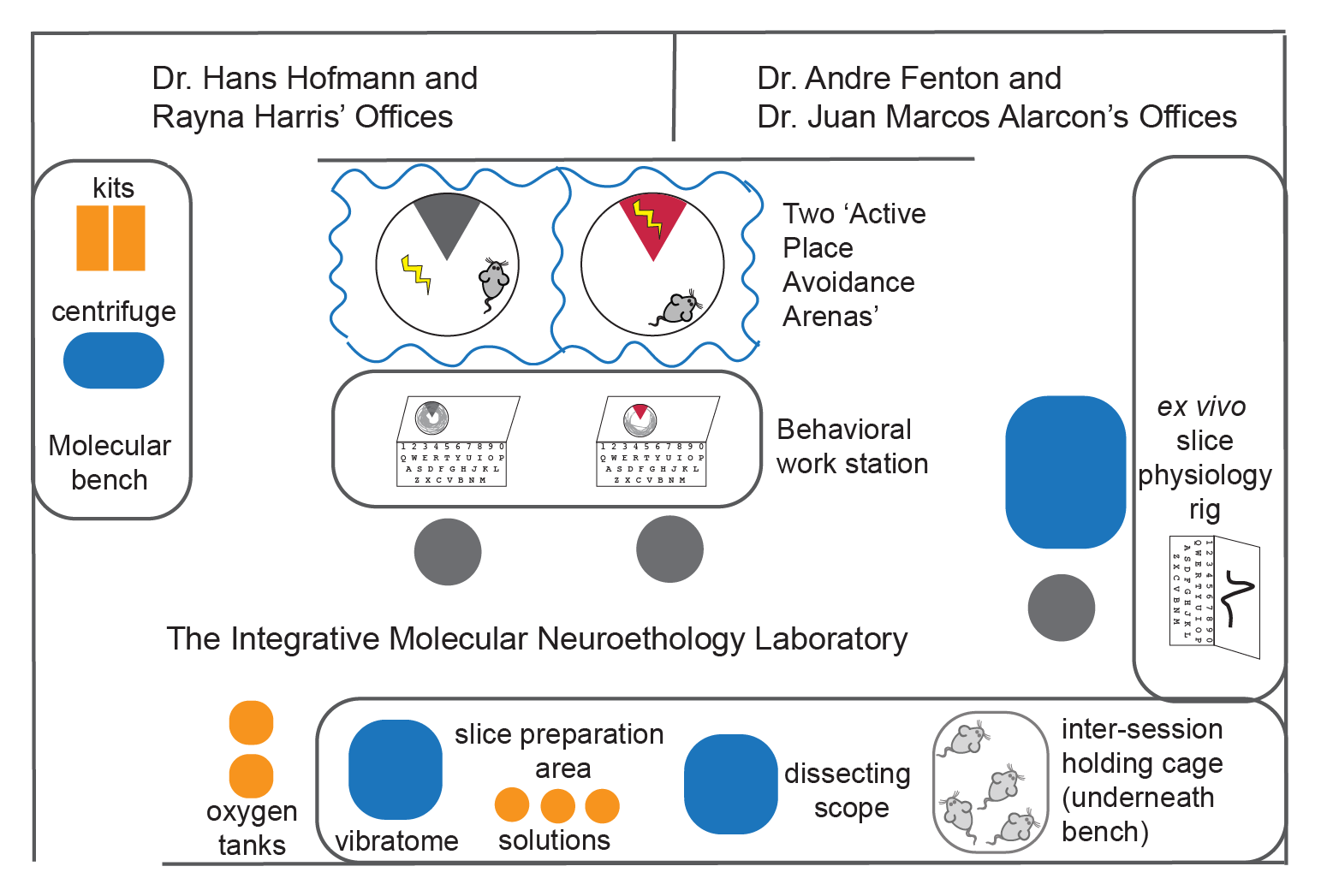
****

Fig. A1: The collaborative ‘Integrative Molecular Neuroethology Laboratory’ at the Marine Biological Laboratories.

#### Experimental design

Male C57BL/6J wild-type and Fragile X mental retardation protein knockout mice (FMR1-KO) were housed on a 12:12 (light: dark) cycle with continuous access to food and water.To examine spatial learning and memory,I used an active place avoidance (APA) paradigm (**Fig. 1A**). Mice were APA trained three times for 10 min each with 2 hours between training sessions. 24 hours later the mice were tested for memory of the shock zone. Behavior was quantified using digital video and automated tracking. After the memory retention test, hippocampal slices were collected for molecular and physiological analysis. Behavior was quantified using digital video and automated tracking.

#### Tissue collection

Each mouse was taken from its cage, anesthetized with 2% (vol/vol) isoflurane for 2 minutes and decapitated. The brain was rapidly dissected, and the hippocampus was removed. 400 um hippocampal slices were incubated at 36°C for 30 min and then at room temperature for 90 min in oxygenated artificial cerebrospinal fluid (aCSF in mM: 125 NaCl, 2.5 KCl, 1 MgSO4, 2 CaCl2, 25 NaHCO3, 1.25 NaH2PO4 and 25 Glucose)46,47. Alternate slices were used for electrophysiology and real-time qPCR. We recorded field extracellular responses to stimulation of the CA3 Schaffer collaterals to characterize CA1 synaptic circuit function 46.

#### RNA isolation, cDNA synthesis, and real-time PCR

CA1 subfields were isolated Using a 0.5 mm punch (Harris UniCore) and a Zeiss dissecting scope. Tissue punches were homogenized in lysis buffer (Promega) and stored at -80C. RNA was isolated using the SV Total RNA Isolation System (Promega) and eluted in 30 or 50 uL nuclease-free water. RNA concentration was quantified using the Quantas Fluorometer (Promega). I estimate that 2-8 ng RNA was isolated from 3 punches (1 per slice, 3 slices), of which 0.5-2 ng of RNA was used for cDNA synthesis. cDNA was synthesized using the GoScript Reverse Transcription System (Promega). Control reactions were run without reverse transcriptase. Taqman assays for probe-base qPCR were purchased from Invitrogen (Table 1). 10 μl real-time PCR amplification reactions were carried out with GoTaq Probe GoTaq Probe qPCR Master Mix with CXR reference dye (Promega) in a StepOnePlus System (Applied Biosystems) with the following cycling parameters: 95°C for 3 sec and 60°C for 30 sec for 35 cycles. All reactions were performed in triplicate. Preliminary analyses (Ct calling, melting curve analysis, and determination of primer efficiencies) were performed using StepOne Software v2.3 (Applied Biosystems).

|  |  |  |  |
| --- | --- | --- | --- |
| **Gene** | **Protein** | **Protein Name** | **TaqMan Assay** |
| grim1 | mGluR1 | metabotropic glutamate receptor 1 | Mm00810231\_s1 |
| prkcz | PKMζ | protein kinase M zeta | Mm00776345\_g1 |
| rpl19 | RPL19 | ribosomal protein L19 | Mm02601633\_g1 |

Table A1: TaqMan assays used in the study

#### Statistical analyses

The Ct values were then converted to values proportional to absolute amounts using the gene-specific amplification efficiencies as the base of the exponent using the R package MCMC.qpcr98. These values were then log2-transformed to yield Ca values, suitable for linear model analysis. A single Poisson-lognormal model was fitted to the counts data using Monte Carlo Markov Chain (MCMC) algorithm to jointly estimate the effects of all experimental factors on the expression of every gene and compute their credible intervals. This method can estimate gene expression changes without making any assumptions concerning control genes98. I calculated Pearson correlations to compare across all three levels of analysis to gain detailed insight into the behavioral, electrophysiological, and neuromolecular mechanisms regulating memory formation.

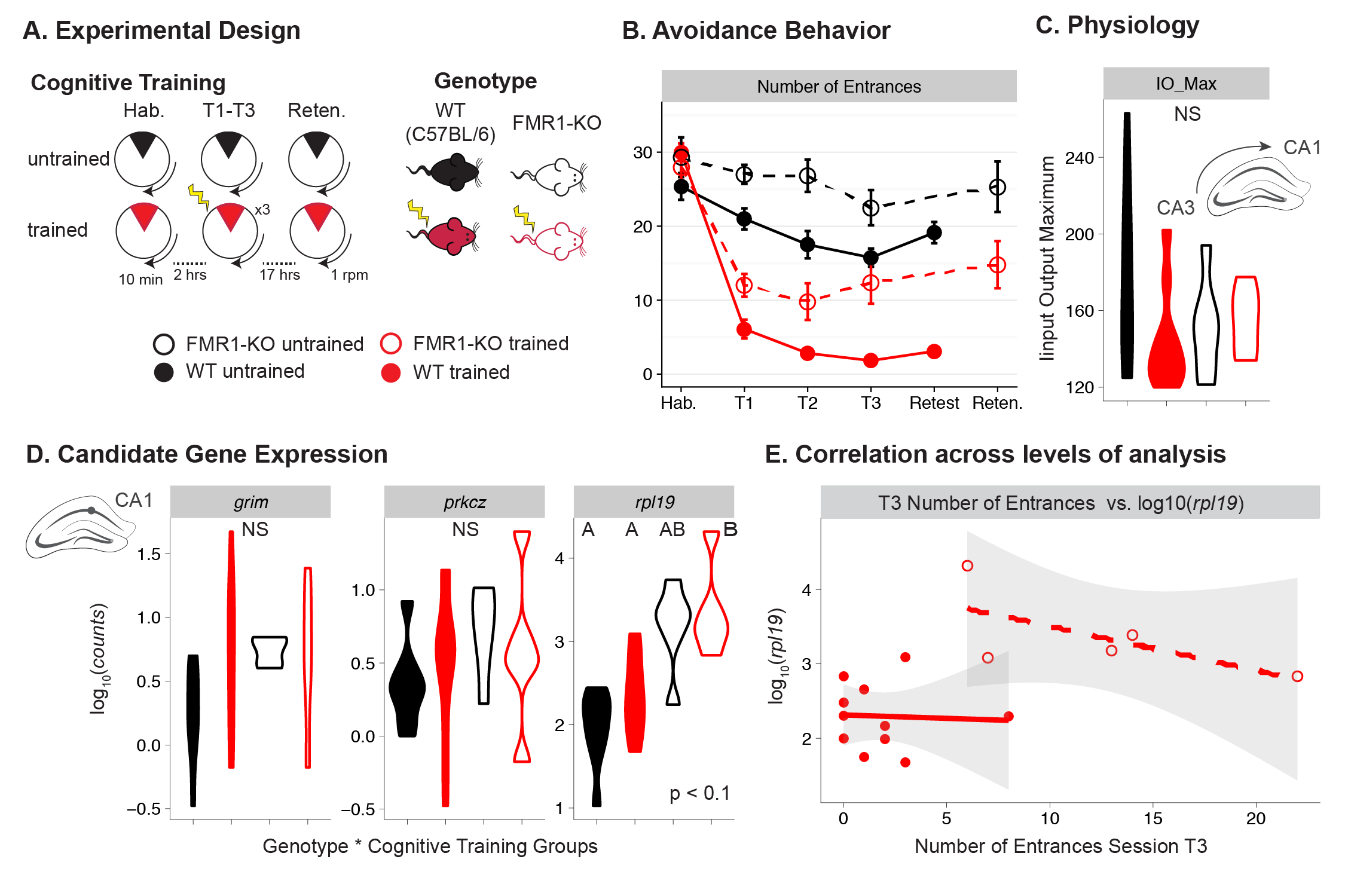


Fig A2: Experimental design, methods, and results

**(A)** I compared untrained and trained mice from WT and FMR1-KO mice strains. The black/red color scheme highlights the untrained to trained contrast. Closed and filled objects highlight the WT to FMR1-KO contrast. **(B)** WT mice (filled red cirles, solid line) enter the shock zone the least of all groups, followed by FMR-KO mice (filled red circles, dashed line), then untrained WT (filled black circles, solid line), then untrained FMR1-KO (open black circles, dashed line). **(C)** The maximum output response with increasing input stimulus intensity was not significantly different between groups. **(D)** *rpl19* was higher in FMR1-KO control mice relative to control and trained WT mice **(E)** Integrating across levels of biological organization, I found a negative correlation between rpl19 expression and the number of entrances into the shock zone in FMR1-KO trained mice.

#### Archival of data and code

All data, code, results, and figures are available at: <https://github.com/raynamharris/FMR1CA1qPCR>

### Results

The trained mice made significantly fewer entrances into the ‘shock zone’ than did untrained mice during all three training sessions and on the retest indicating that both B6 and FMR1-KO trained mice learned to avoid the shock zone and retained the memory for 24 hours (Fig 1B). The maximum output response (fESPS slope) with increasing input stimulus intensity was not significantly different between groups **(Fig 2C)**. I did not find any difference in *grim* or *pkrcz* gene expression between groups but r*pl19* was higher in FMR1-KO control mice relative to control and trained WT mice **(Fig 2D).** Integrating across biological, I found a negative correlation between rpl19 expression and the number of entrances into the shock zone in FMR1-KO trained mice **(Fig 1E).**

### Discussion

The experimental data provide evidence for learning and memory without overall changes in molecular or synaptic activity.A typical gap in knowledge is often linked to biomedical relevance or technological innovation. However, the broader impact of research is not limited to the p-value. Broader impact is a pre-requisite for some scientific funding opportunities. One of the gaps in knowledge I wanted to fill with this research was to promote cross-disciplinary research to neuroethology. To address this, I successfully built a new teaching module for the Neural Systems and Behavior Course. Students in different aspects of this project and determine that these protocols were sufficiently robust and reliable to work in a fast-paced course environment. Students in this new module were able to develop and test hypotheses using the skill set provided, exploring several other model systems and candidate molecular pathways. In the future, I will expand the opportunities that allow students to acquire the knowledge and skills to conduct integrative analyses of neural systems and behavior to understand thoroughly the neuromolecular and neurophysiological mechanisms regulating behavior in diverse model systems. Students will be in a position to potentially identify how forming a memory correlates with changed synaptic function and gene expression at defined regions within the hippocampal circuit.

### Acknowledgments

I am grateful to my lab members and colleagues from the Society for Neuroscience, International Society for Neuroethology, and Society for Integrative and Comparative Biology for fruitful discussion. I are grateful to Ain Chung for assistance with behavioral data collection. This work was carried out as part of the Neural Systems and Behavior Course Directors Research Program run by HAH and AAF at the Marine Biological Laboratory. I thank the Jackson Laboratory for generously donating mice; Promega Corporation for generously donating molecular supplies. This work is supported by NIH-NS091830 to JMA, IOS-1501704 to HAH; NIMH-5R25MH059472-18, the Grass Foundation, and the Helmsley Charitable Trust.

## Appendix 2: My publications with significance and contribution

I started this list in February 2017 as part of the application for a highly competitive Howard Hughes Medical Institute Young Investigator Award. I decided to keep updating the list annotated list of my publications for future personal or professional use. This appendix seems like the appropriate place to archive this record for public reference.

1. Chapter 1 in prep. The one with transcriptomics of avoidance. **Scientific significance:** The hippocampus is involved in spatial navigation, learning, and memory, but little is known about how dynamic activity of transcriptome supports these cognitive processes. I found that gene expression in CA1 and DG discriminates internal versus external variables because yoked and trained mice had the identical physical experience of the world but could interpret the experience differently. I observed that expression of nuclear signaling genes in DG are most responsive to memory formation; while changes in CA1 reflect membrane-level ion-channel regulation. Expression in CA1could be a molecular signature for a form of stress-related memory. **Contribution and career stage:** I helped to conceptualize and designed the experiments, collected experiments, analyzed the data, and wrote and revised the manuscript. This corresponds to chapter 1 of my research.
2. Chapter 2 in prep. The one with fragile X transcriptomics. **Scientific significance:** … **Contribution and career stage:** I helped to conceptualize and designed the experiments, collected experiments, analyzed the data, and wrote and revised the manuscript. This is chapter 2 of my thesis research.
3. Harris et al. 201733. The one with reproducible hippocampal transcriptomics. **Scientific significance:** The future will be filled withsingle-neuron gene expression studies designed to understanding nervous system structure and neural circuit function. Cellular dissociation is a prerequisite single-cell sequencing analysis, but the extent to which this process alters the molecular activity in neural tissues has not been determined. I found that chemical dissociation response is greater than the transcriptomic response to a stressful experience but weaker than cognitive training. A meta-analysis with publically available data identified concordant and discordant effects of technical and behavioral manipulations which will inform the design of future single-neuron transcriptome studies to facilitate a more comprehensive understanding of brain function. **Contribution and career stage:** I helped to conceptualize and designed the experiments, collected experiments, analyzed the data, and wrote and revised the manuscript. This is chapter 3 of my thesis research.
4. Devenyi et al99. The one with lesson development rules. **Scientific significance:** Lessons take significant effort to build and maintain. The collaborative development methods pioneered by the open source software community offer an effective, economical way to create and sustain lessons that can be used by large numbers of people in a wide variety of contexts. The ten simple rules outlined in this paper summarize the best practices that have been implemented by several successful open education projects aimed at researchers and research librarians in a wide range of disciplines. **Contribution and career stage:** I helped to conceptualize, create graphics, write, and edit the manuscript. During this time, I was supported by a fellowship form the Graduate School, and I was Software Carpentry Steering Committee Member.

HIDDEN TEXT: NOTE: If you have more than one appendix you can use the Heading 2,h2 style and label each appendix separately, e.g., Appendix A, Appendix B, etc. You could also title this section “Appendices” using the Heading 2, h2 style, and use Heading 3,h3 for each separate appendix

1. Renn et al. 2017100. The one with comparative transcriptomic of monogamy with cichlid microarrays. **Scientific significance:** The diversity of mating systems among animals is astounding; however, our understanding of the mechanisms underlying these convergently evolved phenotypes is limited. This research tooke a comparative transcriptomic approach to test the hypothesis that these independent transitions have recruited similar gene sets using field-collected males and females of four closely related species representing two independent evolutionary transitions from polygyny to monogamy from Ectodini cichlids from Lake Tanganyika. Our experiment identified 331 genes (∼6% of those assayed) associated with monogamous mating systems independent of species. This suggest deep molecular homologies underlying the convergent or parallel evolution of monogamy in different cichlid lineages of Ectodini. **Contribution and career stage:** I helped write the manuscript. I was a graduate student during the time of writing.
2. Dijkstra et al. 2017101.The one with body color and social behavior. **Scientific significance:** The melanocortin system regulates body color and social behavior. I found that a melanocortin receptor agonist increased aggressive behavior and body coloration while a melanocortin antagonist reduced aggression but had no affect on body color in *Astatotilapia burtoni*. Our comparison of two dominant male morphs with natural differences in yellow body color and aggression revealed that gene expression of melanocortin receptors and ligands was not differentially in the brain but is differentially regulated in the scales. These results suggest that the melanocortin system contributes to the polymorphism in behavior and coloration in A. burtoni but the mechanism is not completely understood. **Contribution and career stage:** I cloned the candidate genes and submitted their sequences to the National Center for Biotechnology Information (NCBI) database. I trained and supervised the first author in molecular and data analysis techniques for qPCR. I was a research a graduate student when these experiments were designed, conducted, and published. I contributed 15%.
3. Harris et al. 20171. The one with crab neurons and gene expression. **Scientific significance:** Gene expression analysis has become an increasingly prominent tool across biological disciplines. This publication presents discovery-driven research conducted during the Neural Systems and Behavior (NS&B) course at the Marine Biological Laboratories. I combined RNAi methods for gene knockdown with single cell electrophysiology and gene expression profiling to examine the electrophysiological and molecular consequences of gene manipulation. Our results provide insights into the neural mechanisms underlying generation and modulation of rhythmic motor patterns. This experiment launched an ongoing research and education program at NS&B that has expanded into single-neuron transcriptomics. **Contribution and career stage:** I led this project during the NS&B course. I worked with all authors to design the experiment and trained the 2nd and 3rd authors in gene expression methods and analysis. I wrote and revised the manuscript with approval of all co-authors. I was a graduate student when this project was conceived, executed, and published. I contributed 60%.
4. Harris, O’Connell, & Hofmann 2016102. The one with brain evolution and development. **Scientific significance:** Across animals there is astonishing diversity in the structure and function of nervous systems and the resulting behavior patterns. The question of how this diversity has evolved has long fascinated biologists, yet it was not until fairly recently that the mechanisms that make such variation possible have become a focus of study. In this book chapter, I review neural development, neural circuits and neurochemicals that modulate animal behavior, timescales of neural plasticity, and the evolution of mechanisms underlying neural plasticity. This is an important contribution to the field of neuroscience where researchers do not typically explore brain function in the context of evolution. **Contribution and career stage:** I co-wrote this review with another graduate student and our mentor. My portions were inspired by the readings from a graduate course on “Brain, Behavior, and Evolution”. I helped write the first draft, and I played a major role in the response to reviewer comments, revising the manuscript, and creating the illustration. I contributed 40%.
5. Ahmadia et al. 2016103. **Scientific significance:** The non-profit organization Software Carpentry teaches scientific computing workshops around the globe. This training manual contains materials for a two-day instructor certification program. The goals of the manual are to introduce: 1) general educational philosophy and evidence-based teaching and 2) specific teaching practices used in Software Carpentry’s scientific computing workshops. This training manual (or earlier editions) have been used to certify 850 instructors who have taught 20,000 learners worldwide. This manual provides researchers of all disciplines with tools and resources to be more effective teachers. **Contribution and career stage:** These open access educational material are collaborated developed on GitHub. I have contributed new background materials and exercises and helped with editing and organizing lessons. I was a PhD student during this time. I contributed 5%.
6. Goppert et al. 2016104.The one with aromatase and sex changing cichlids. **Scientific significance:** Estrogen and testosterone drive sexual development. Fish exhibit an amazing diversity of plasticity in sexual development. I found that inhibiting estrogen synthesis in *A. burtoni* females rapidly produces a male-like phenotype but does not transform ovaries into functional testis. I found that inhibiting estrogen synthesis has different effect son steroid receptor expression in the gonads and brain, which may contribute to the observed phenotypical changes. These results highlight the fact that hormonal impacts are species-specific and vary according to the level of biological organization. This is an important contribution to the literature on sexual development and plasticity. **Contribution and career stage:** I contributed to the hormone assays, experimental design, data collection, and data analysis. I taught the first and last author how to use R for data visualization and statistics. I contributed to writing and editing manuscripts. I was a research graduate student when these experiments were designed, conducted, and published. I contributed 15%.
7. Oldfield et al 2015105. The one with vasopressin and space use theory. **Scientific significance:** Understanding the causes and consequences of animal behavior is a fundamental goal in biology. This synthesis provided a conceptual integration of ultimate- and proximate-level perspectives on how neuroendocrine systems mediate variation in social behavior. Previously, researchers were confused why comparative studies of monogamous and polygamous mating systems did not reveal the same patterns of neuroendocrine activity. Our synthesis demonstrated that inclusion of information about an animal’s social organization resolved many conflicting results. This conceptual framework this is critical for understanding the evolutionary and environmental mechanisms that shape behavior across taxa. **Contribution and career stage:** My major contribution to this synthesis paper was restructuring the content to make the arguments more concise and accessible to a broad audience. I helped with the literature search, illustrations, writing, and revision. I was a graduate student at the time of writing and publication. I contributed 25%.
8. Simões et al 2015106. The one with neurogenomic responses to fish pheromones. **Scientific significance:** Animals adjust their behavior when interacting with conspecifics of different sex and social status. This study explored how behavioral plasticity is achieved through altered neural activity and gene expression in brain regions that respond to olfactory cues about their social environment. The results demonstrate that rapid transcriptional responses in specific brain regions are specific to male social status and female reproductive state. These findings underscore the extensive transcriptional plasticity that accompanies behavioral plasticity in rich social environments. **Contribution and career stage:** I helped designed the microarray experiment, trained the first author in the molecular techniques for the microarray experiments, and helped write and edit the manuscript. I was a research technician when the experiments were designed and conducted, and a graduate student when the manuscript was written. I contributed 10%.
9. Stiver et al 2015107. The one with wrasse reproductive tactics. **Scientific significance:** Plasticity in reproductive traits can lead to the evolution of discrete variation in behavior, morphology, and physiology. In the ocellated wrasse *Symphodus ocellatus*, males exhibit different reproductive tactics. I found that androgens were highest in “nesting males” that defended territories and provide parental care. Microarray data showed that “sneaker males” that use force copulation tactic had the highest number of differentially expressed genes compared females, nesting males, and “satellite males” that help nesting males defend territories. This research provides better understanding of how plasticity in reproductive traits arises from and is maintained by a common genome. **Contribution and career stage:** I designed, conducted, and anlalyzed the hormone assays, trained the first author in histological and microscopy techniques, and helped write and edit the manuscript. I was a research technician when the experiments were designed and conducted, and a graduate student when the manuscript was written. I contributed 15%.
10. Harris & Hofmann 2015108. The one with opsin gene evolution. **Scientific significance:** Gene duplication and deletion are important evolutionary mechanisms that facilitate phenotypic diversity. This publication is a commentary on Cortesi et al. 2016, which investigated the effects of opsin gene duplication and deletion on the adaptation to diverse light environments and visual displays in cichlid fishes. My commentary zooms out and looks at gene family evolution and phenotypic diversification across the vertebrate kingdom. Both the commentary and the original article underscore the power of the comparative approach for understanding gene family evolution and the origins of animal diversity. **Contribution and career stage:** An editor invited this commentary in response to a peer-review I co-wrote with my advisor. I conceived the narrative, performed a literature search, wrote the manuscript, created the figure, and made revisions following editorial review. I was a PhD graduate student at the time of publication. I contributed 70%.
11. Brawand et al 2014109. The one with the five cichlid genomes. **Scientific significance:** The rapid evolution of African cichlid fish represents one of the most recent and spectacular evolutionary events on the planet. The creation of five reference genomes and transcriptomes has been an incredibly valuable resource for the cichlid community, as demonstrated by over 200 citations since its publication. The major finding of this study was that cichlids have an excess of gene duplications and accelerated coding sequence evolution compared to other teleosts, which likely contributed to their radiation. This research has shaped our understanding of evolutionary processes and created a wealth of new resources and tools for studying genome evolution across the tree of life. **Contribution and career stage:** I isolated DNA and RNA from *Astatotialpia burtoni*, analyzed expression and evolution of paralogous genes, and helped write the manuscript. I was a technician when the experiments were designed and conducted, and a graduate student when the manuscript was written. My contributions were critical, but they only total 1% of this 10-year endeavor.
12. Harris & Hofmann 20142. **Scientific significance:** The Neural Systems and Behavior (NS&B) Course is the premier discovery-driven training opportunity for the next generation of neuroethologists and systems neuroscientists. This publication is an open-access training manual for a two-week research-driven educational course. This manual equips students with a comprehensive understanding of the relationships between neuromolecular and complex behavior. I used this manual to train a dozen graduate students. One pair of student co-authors a paper based on the research the conducted after learning the technique described herein. Forthcoming materials will include lessons for reproducible analysis of next-generation sequencing data. **Contribution and career stage:** I developed this curriculum based on techniques that I learned in the Hofmann Lab and adapted to utilize generous donations from Promega and Invitrogen. I was a graduate student when this was conceived, written, and implemented. My contribution was 70%.
13. Smith et al. 2015110. The one with swordtails, hormones, and body size. **Scientific significance:** Alternative reproductive tactics (ARTs) are alternative behavioral strategies that animals use to optimize reproductive fitness. ARTs with strong genetic components are often associated with variation in morphology that predicts the tactic. The study confirmed a genetic link between body size and ART, showing that male body size in the fish *Xiphophorus nigrensis* increases with increasing copy number in the melanocortin 4 receptor. However, there was no significant association between mc4r copy number and behavior or circulating testosterone levels. This is the first study to identify a genetic underpinning for variation in morphology but not behavior in an alternative reproductive tactic. **Contribution and career stage:** I validated hormone processing methods for this non-model organisms and trained the first author and two undergraduate students in new techniques. I contributed to writing, editing, and revision of the manuscript. I was a research technician when the experiments were conducted and a graduate student at the time of publication. I contributed 20%.
14. Harris et al. 2014111. The one with pro-opiomelanocortin gene evolution. **Scientific significance:** The melanocortin system is associated with phenotypic diversification across vertebrates. Pro-opiomelanocortin (*pomc*) genes are central to the highly pleiotropic melanocortin system. When researching *pomc*, I discovered a novel melanocortin peptide that resulted from of a tandem gene duplication in the *pomc B* gene of some fish. A comprehensive analysis of *pomc* promoters identified potential regulatory mechanisms that govern tissue-specific expression. Our results suggested a novel mechanism for melanocortin receptor regulation of neuroendocrine function. The discovery of the novel gene product sparked my interest in gene manipulation and inspired a new line of research for my thesis. **Contribution and career stage:** I spearheaded this project, conceived and designed the experiments, collected and analyzed the data, and wrote and revised the manuscript. I was a research technician when the experiments were designed and conducted, and a graduate student when the manuscript was written. I contributed 70%.
15. Fischer et al 2014112. The one with stressed-out guppies. **Scientific significance:** Trinidadian guppies from high-predation environments have repeatedly colonized and adapted to low-predation environments, resulting in parallel changes morphology and behavior. I investigated how the glucocorticoid system mediates stress responses in the Trinidadian guppy system. After validating methods for non-invasive hormone collection, I found that guppies from high predation environments had lower basal levels of glucocorticoids, and guppies reared in high predation environments had reduced glucocorticoid release. These results demonstrated that timescales of exposure to stress (evolutionary and short-term) reduce glucocorticoids but depend on distinct physiological mechanisms. **Contribution and career stage:** I validated hormone collection and processing methods for this non-model organisms and trained the first author in new techniques. I contributed to writing, editing, and revising the manuscript. I was a research technician when the experiments were conducted, and a graduate student at the time of publication. I contributed 30%.
16. Harris & Hofmann 201425. The one with social behavior and reverse genomics. **Scientific significance:** Biological diversity emerges from the interaction between genomes and their environment. In this book chapter, I review genomic studies investigating the neural processes the govern animal behavior. Then, I briefly describe forward and reverse genetics and argue that “forward genomic” studies should be combined with “reverse genomics”. I conclude that functional genomic studies can move beyond statistical associations between genes and behavior by manipulating candidate genes. This would reveal novel insights into the neuromolecular control of social behavior and its evolution. This review has been cited 15 times and has been incorporated into multiple graduate courses across the country. **Contribution and career stage:** I wrote this book chapter during my first semester as a PhD graduate student while taking a graduate-level genetics course. I coined the term reverse genomics to describe the process of measuring the transcriptional response to gene manipulation. My contribution was 70% of the total effort.
17. Sessa et al 2013113. The one with monogamy and steroid hormones. **Scientific significance:** Steroid hormones modulate behavioral responses to social stimuli. The results of this research demonstrate that aggressive and reproductive behaviors are regulated by androgen and estrogen receptors in a context-specific manner. This research filled a gap in the literature by disentangling the relationships between sex steroid hormones and context-specific social behaviors by measuring and manipulating behavioral and hormonal responses to multiple stimuli (including reproductive females, aggressive males, and non-social stimuli). These results extended our understanding of sex steroid regulation of behavioral responses to rich social environments. **Contribution and career stage:** I performed the hormone assay and wrote the corresponding methods. After the first author graduated I responded to all reviewer comments and revised the manuscript accordingly. I was a research technician when the experiments were conducted and a graduate student at the time of publication. I contributed 30%.
18. Oldfield et al 2013114. The one with monogamy and neuro-peptides. **Scientific significance:** Arginine vasotocin and its receptor 1a (V1ar) regulate variability in reproductive behavior across vertebrates. Our results demonstrated that a polygamous cichlid fish, Herichthys minckleyi, had higher levels of V1ar in the brain but lower levels of circulating testosterone compared to the monogamous Herichthys cyanoguttatus. I found that in this study, and in many other published studies, the difference between monogamous and polygamous mating systems was confounded by other ecologically relevant factors including territory size, space use, parental care and social dominance hierarchies. This gap in knowledge inspired the synthesis paper I co-authored (Olfield, Harris, & Hofmann 2015). **Contribution and career stage:** I designed and validated the qPCR and hormone assays for these two non-model organisms, analyzed the data, and helped write the manuscript. I was a research technician when the experiments were conducted and a graduate student at the time of publication. I contributed 25%.

## Bibliography

1. Harris, R. M. *et al.* Single Neuron Gene Expression Analysis Using the Maxwell 16 LEV System in the Neural Systems and Behavior Course. *bioRxiv* (2017). doi:10.1101/107342

2. Harris, R. M. & Hofmann, H. A. Laboratory Experiments for the 2014 Integrative Molecular Neuroscience Module. (2014). doi:10.5281/ZENODO.60725

3. Domjan, M., Grau, J. W. & Krause, M. A. *Principles of learning and behavior*. (Wadsworth, 2010).

4. Alivisatos, A. P. *et al.* The brain activity map project and the challenge of functional connectomics. *Neuron* **74,** 970–4 (2012).

5. Kandel, E. R., Markram, H., Matthews, P. M., Yuste, R. & Koch, C. Neuroscience thinks big (and collaboratively). *Nat. Rev. Neurosci.* **14,** 659–64 (2013).

6. Jonas, E. & Kording, K. P. Could a Neuroscientist Understand a Microprocessor? *PLOS Comput. Biol.* **13,** e1005268 (2017).

7. McDonald, R. J., Hong, N. S. & Devan, B. D. The challenges of understanding mammalian cognition and memory-based behaviours: an interactive learning and memory systems approach. *Neurosci. Biobehav. Rev.* **28,** 719–745 (2004).

8. Howard, M., Stern, C. & Hasselmo, M. Howard Eichenbaum 1947–2017. *Nat. Neurosci.* **20,** 1432–1433 (2017).

9. O’Keefe, J. & Nadel, L. *The hippocampus as a cognitive map*. (Clarendon Press, 1978).

10. Ramon y Cajal, S. Histologie du système nerveux de l’homme et des vertébrés. *Vol. 2. Paris Maloine* 891–942 (1911). doi:http://dx.doi.org/10.5962/bhl.title.48637

11. O’Keefe, J. & Dostrovsky, J. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res.* **34,** 171–5 (1971).

12. Bures, J., Fenton, A. A., Kaminsky, Y. & Zinyuk, L. Place cells and place navigation. *Proc. Natl. Acad. Sci.* **94,** 343–350 (1997).

13. Moser, E. I., Kropff, E. & Moser, M.-B. Place Cells, Grid Cells, and the Brain’s Spatial Representation System. *Annu. Rev. Neurosci.* **31,** 69–89 (2008).

14. Burghardt, N. S., Park, E. H., Hen, R. & Fenton, A. A. Adult-born hippocampal neurons promote cognitive flexibility in mice. *Hippocampus* **22,** 1795–808 (2012).

15. Denny, C. A., Burghardt, N. S., Schachter, D. M., Hen, R. & Drew, M. R. 4- to 6-week-old adult-born hippocampal neurons influence novelty-evoked exploration and contextual fear conditioning. *Hippocampus* **22,** 1188–1201 (2012).

16. Drew, M. R., Denny, C. A. & Hen, R. Arrest of adult hippocampal neurogenesis in mice impairs single- but not multiple-trial contextual fear conditioning. *Behav. Neurosci.* **124,** 446–454 (2010).

17. Denny, C. A. *et al.* Hippocampal memory traces are differentially modulated by experience, time, and adult neurogenesis. *Neuron* **83,** 189–201 (2014).

18. Lisman, J. Long-term potentiation: outstanding questions and attempted synthesis. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **358,** 829–42 (2003).

19. Sacktor, T. C. How does PKMζ maintain long-term memory? *Nat. Rev. Neurosci.* **12,** 9–15 (2011).

20. Lichtman, J. W., Livet, J. & Sanes, J. R. A technicolour approach to the connectome. *Nat. Rev. Neurosci.* **9,** 417–422 (2008).

21. Livet, J. *et al.* Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* **450,** 56–62 (2007).

22. Weissman, T. A., Sanes, J. R., Lichtman, J. W. & Livet, J. Generating and imaging multicolor Brainbow mice. *Cold Spring Harb. Protoc.* **6,** 763–769 (2011).

23. Hitzemann, R. *et al.* Genes, behavior and next-generation RNA sequencing. *Genes, Brain Behav.* **12,** 1–12 (2013).

24. O’Connell, L. A. & Hofmann, H. A. Genes, hormones, and circuits: an integrative approach to study the evolution of social behavior. *Front. Neuroendocrinol.* **32,** 320–335 (2011).

25. Harris, R. M. & Hofmann, H. A. Neurogenomics of Behavioral Plasticity. *Adv. Exp. Med. Biol.* **781,** 149–168 (2014).

26. Robinson, G. E., Fernald, R. D. & Clayton, D. F. Genes and social behavior. *Science* **322,** 896–900 (2008).

27. Aebersold, R., Hood, L. & Watts, J. Equipping scientists for the new biology. *Nat. Biotechnol.* **18,** 359–359 (2000).

28. Lein, E. S. *et al.* Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445,** 168–76 (2007).

29. Hawrylycz, M. J. *et al.* An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489,** 391–399 (2012).

30. Lein, E. S., Zhao, X. & Gage, F. H. Defining a molecular atlas of the hippocampus using DNA microarrays and high-throughput in situ hybridization. *J. Neurosci.* **24,** 3879–89 (2004).

31. Cembrowski, M. S. *et al.* Spatial Gene-Expression Gradients Underlie Prominent Heterogeneity of CA1 Pyramidal Neurons. *Neuron* **89,** 351–368 (2016).

32. Cembrowski, M. S., Wang, L., Sugino, K., Shields, B. C. & Spruston, N. Hipposeq: A comprehensive RNA-seq database of gene expression in hippocampal principal neurons. *Elife* **5,** e14997 (2016).

33. Harris, R. M., Kao, H.-Y., Alarcon, J. M., Hofmann, H. A. & Fenton, A. Analysis of hippocampal transcriptomic responses to technical and biological perturbations. *bioRxiv* 153585 (2017). doi:10.1101/153585

34. Lesburguères, E., Sparks, F. T., O’Reilly, K. C. & Fenton, A. A. Active place avoidance training is minimally stressful and does not elicit fear responses.

35. Radwan, B., Dvorak, D. & Fenton, A. A. Impaired cognitive discrimination and discoordination of coupled theta-gamma oscillations in Fmr1 knockout mice. *Neurobiol. Dis.* **88,** 125–38 (2016).

36. Cimadevilla, J. M., Wesierska, M., Fenton, A. A. & Bures, J. Inactivating one hippocampus impairs avoidance of a stable room-defined place during dissociation of arena cues from room cues by rotation of the arena. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 3531–6 (2001).

37. Cimadevilla, J. M., Fenton, A. A. & Bures, J. Functional inactivation of dorsal hippocampus impairs active place avoidance in rats. *Neurosci. Lett.* **285,** 53–56 (2000).

38. McKiernan, E. C. *et al.* *How open science helps researchers succeed*. *eLife* **5,** e16800 (eLife Sciences Publications Limited, 2016).

39. Blischak, J. D., Davenport, E. R. & Wilson, G. A Quick Introduction to Version Control with Git and GitHub. *PLoS Comput. Biol.* **12,** e1004668 (2016).

40. Wilson, G. *et al.* Best practices for scientific computing. *PLoS Biol.* **12,** e1001745 (2014).

41. Wilson, G. *et al.* Good enough practices in scientific computing. *arXiv Prepr.* **13,** arXiv:1609.00037 (2016).

42. Wilson, G. *et al.* Good enough practices in scientific computing. *PLOS Comput. Biol.* **13,** e1005510 (2017).

43. Cimadevilla, J. M., Kaminsky, Y., Fenton, A. & Bures, J. Passive and active place avoidance as a tool of spatial memory research in rats. *J. Neurosci. Methods* **102,** 155–164 (2000).

44. R Development Core Team. R: a language and environment for statistical computing | GBIF.ORG. (2013). at <http://www.r-project.org/>

45. Bakeman, R. Recommended effect size statistics for repeated measures designs. *Behav. Res. Methods* **37,** 379–84 (2005).

46. Pavlowsky, A. & Alarcon, J. M. Interaction between Long-Term Potentiation and Depression in CA1 Synapses: Temporal Constrains, Functional Compartmentalization and Protein Synthesis. *PLoS One* **7,** e29865 (2012).

47. Pavlowsky, A., Wallace, E., Fenton, A. A. & Alarcon, J. M. Persistent modifications of hippocampal synaptic function during remote spatial memory. *Neurobiol. Learn. Mem.* **138,** 182–197 (2017).

48. FastQC. at <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

49. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32,** 3047–3048 (2016).

50. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34,** 525–527 (2016).

51. Waterston, R. H. *et al.* Initial sequencing and comparative analysis of the mouse genome. *Nature* **420,** 520–62 (2002).

52. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17,** 10–12 (2011).

53. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15,** 550 (2014).

54. Neuwirth, E. RColorBrewer: ColorBrewer Palettes. (2014).

55. Suzuki, R. & Shimodaira, H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* **22,** 1540–1542 (2006).

56. Gentleman R, Carey V, H. W. and H. F., Gentleman, R., Carey, V., Huber, W. & Hahne, F. genefilter: genefilter: methods for filtering genes from high-throughput experiments. (2017).

57. Wilke, C. O. cowplot: Streamlined Plot Theme and Plot Annotations for ‘ggplot2’. (2016).

58. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag New York, 2009). doi:10.1007/978-0-387-98141-3

59. Wright, R. M., Aglyamova, G. V, Meyer, E. & Matz, M. V. Gene expression associated with white syndromes in a reef building coral, Acropora hyacinthus. *BMC Genomics* **16,** 371 (2015).

60. Kolde, R. pheatmap: Pretty Heatmaps. (2015).

61. Garnier, S. viridis: Default Color Maps from ‘matplotlib’. (2016).

62. McWhite, C. D. & Wilke, C. O. colorblindr: Simulate colorblindness in R figures. at <https://github.com/clauswilke/colorblindr>

63. Kheirbek, M. A. *et al.* Differential control of learning and anxiety along the dorsoventral axis of the dentate gyrus. *Neuron* **77,** 955–68 (2013).

64. Lesburguères, E., Sparks, F. T., O’Reilly, K. C. & Fenton, A. A. Active place avoidance is no more stressful than unreinforced exploration of a familiar environment. *Hippocampus* **26,** 1481–1485 (2016).

65. Ježek, K. *et al.* Stress-Induced Out-of-Context Activation of Memory. *PLoS Biol.* **8,** e1000570 (2010).

66. Pastalkova, E. *et al.* Storage of spatial information by the maintenance mechanism of LTP. *Science* **313,** 1141–1144 (2006).

67. Garber, K. B., Visootsak, J. & Warren, S. T. Fragile X syndrome. *Eur. J. Hum. Genet.* **16,** 666–672 (2008).

68. Cotney, J. *et al.* The autism-associated chromatin modifier CHD8 regulates other autism risk genes during human neurodevelopment. *Nat. Commun.* **6,** 6404 (2015).

69. Bassell, G. J. & Warren, S. T. Fragile X syndrome: loss of local mRNA regulation alters synaptic development and function. *Neuron* **60,** 201–14 (2008).

70. Voineagu, I. *et al.* Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* **474,** 380–384 (2011).

71. The Dutch-Belgian Fragile X Consorthium *et al.* Fmr1 knockout mice: A model to study fragile X mental retardation. *Cell* **78,** 23–33 (1994).

72. Pilaz, L. J., Lennox, A. L., Rouanet, J. P. & Silver, D. L. Dynamic mRNA Transport and Local Translation in Radial Glial Progenitors of the Developing Brain. *Curr. Biol.* **26,** 3383–3392 (2016).

73. Heyes, S. *et al.* Genetic disruption of voltage-gated calcium channels in psychiatric and neurological disorders. *Prog. Neurobiol.* **134,** 36–54 (2015).

74. Strom, S. P. *et al.* High-density SNP association study of the 17q21 chromosomal region linked to autism identifies CACNA1G as a novel candidate gene. *Mol. Psychiatry* **15,** 996–1005 (2010).

75. Lu, A. T.-H., Dai, X., Martinez-Agosto, J. A. & Cantor, R. M. Support for calcium channel gene defects in autism spectrum disorders. *Mol. Autism* **3,** 18 (2012).

76. Brown, V. *et al.* Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome. *Cell* **107,** 477–87 (2001).

77. Huber, K. M., Kayser, M. S. & Bear, M. F. Role for rapid dendritic protein synthesis in hippocampal mGluR-dependent long-term depression. *Science* **288,** 1254–1256 (2000).

78. Zhong, J. *et al.* BC1 regulation of metabotropic glutamate receptor-mediated neuronal excitability. *J. Neurosci.* **29,** 9977–9986 (2009).

79. Weiler, I. J. *et al.* Fragile X mental retardation protein is necessary for neurotransmitter-activated protein translation at synapses. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 17504–17509 (2004).

80. Zhong, J. *et al.* Regulatory BC1 RNA and the Fragile X Mental Retardation Protein: Convergent Functionality in Brain. *PLoS One* **5,** 7 (2010).

81. Lisman, J., Schulman, H. & Cline, H. The molecular basis of CaMKII function in synaptic and behavioural memory. *Nat. Rev. Neurosci.* **3,** 175–190 (2002).

82. Zhang, P. & Lisman, J. E. Activity-dependent regulation of synaptic strength by PSD-95 in CA1 neurons. *J. Neurophysiol.* **107,** 1058–66 (2011).

83. Soderling, T. R. CaM-kinases: modulators of synaptic plasticity. *Curr. Opin. Neurobiol.* **10,** 375–80 (2000).

84. Alexander, H. *et al.* *Software Carpentry: The Unix Shell*. (2017). doi:10.5281/ZENODO.278226

85. Allen, J. *et al.* *Software Carpentry: R For Reproducible Scientific Analysis*. (2017). doi:10.5281/ZENODO.278224

86. Ahmadia, A. *et al.* *Software Carpentry: Version Control With Git*. (2017). doi:10.5281/ZENODO.278219

87. Mizuseki, K., Diba, K., Pastalkova, E. & Buzsáki, G. Hippocampal CA1 pyramidal cells form functionally distinct sublayers. *Nat. Neurosci.* **14,** 1174–1181 (2011).

88. Danielson, N. B. *et al.* Sublayer-Specific Coding Dynamics during Spatial Navigation and Learning in Hippocampal Area CA1. *Neuron* **91,** 652–665 (2016).

89. Guzowski, J. F., McNaughton, B. L., Barnes, C. A. & Worley, P. F. Environment-specific expression of the immediate-early gene Arc in hippocampal neuronal ensembles. *Nat. Neurosci.* **2,** 1120–4 (1999).

90. Guzowski, J. F. *et al.* Recent behavioral history modifies coupling between cell activity and Arc gene transcription in hippocampal CA1 neurons. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 1077–82 (2006).

91. Chalancon, G. *et al.* Interplay between gene expression noise and regulatory network architecture. *Trends Genet.* **28,** 221–232 (2012).

92. Mo, A. *et al.* Epigenomic Signatures of Neuronal Diversity in the Mammalian Brain. *Neuron* **86,** 1369–1384 (2015).

93. Harris, R. M., Kao, H.-Y., Alarcon, J. M., Hofmann, H. A. & Fenton, A. A. The Github Repository For Analyses Of Hippocampal Transcriptomic Responses To Technical And Biological Perturbations. (2017). doi:10.5281/ZENODO.815081

94. Collingridge, G. L., Kehl, S. J. & McLennan, H. The antagonism of amino acid-induced excitations of rat hippocampal CA1 neurones in vitro. *J. Physiol.* **334,** 19–31 (1983).

95. Morris, R. G. M. NMDA receptors and memory encoding. *Neuropharmacology* **74,** 32–40 (2013).

96. Neuner, S. M., Wilmott, L. A., Hoffmann, B. R., Mozhui, K. & Kaczorowski, C. C. Hippocampal proteomics defines pathways associated with memory decline and resilience in normal aging and Alzheimer’s disease mouse models. *Behav. Brain Res.* **322,** 288–298 (2017).

97. Ménard, C. & Quirion, R. Group 1 metabotropic glutamate receptor function and its regulation of learning and memory in the aging brain. *Front. Pharmacol.* **3,** 182 (2012).

98. Matz, M. V., Wright, R. M. & Scott, J. G. No Control Genes Required: Bayesian Analysis of qRT-PCR Data. *PLoS One* **8,** e71448 (2013).

99. Devenyi, G. A. *et al.* Ten simple rules for collaborative lesson development. (2017). at <http://arxiv.org/abs/1707.02662>

100. Renn, S. C. P. *et al.* Gene Expression Signatures of Mating System Evolution. *Genome* (2017). doi:10.1139/gen-2017-0075

101. Dijkstra, P. D. *et al.* The melanocortin system regulates body pigmentation and social behaviour in a colour polymorphic cichlid fish. *Proc. R. Soc. B Biol. Sci.* **284,** 20162838 (2017).

102. Harris, R. M. R. M. *et al.* in *The Wiley Handbook of Evolutionary Neuroscience* (ed. Shepherd, S. V.) 422–443 (John Wiley & Sons, Ltd, 2016). doi:10.1002/9781118316757.ch15

103. Ahmadia, A. *et al.* *Software Carpentry: Instructor Training*. (2016). doi:10.5281/ZENODO.57571

104. Göppert, C. *et al.* Inhibition of Aromatase Induces Partial Sex Change in a Cichlid Fish: Distinct Functions for Sex Steroids in Brains and Gonads. *Sex. Dev.* **10,** 97–110 (2016).

105. Oldfield, R. G., Harris, R. M. & Hofmann, H. A. Integrating resource defence theory with a neural nonapeptide pathway to explain territory-based mating systems. *Front Zool* **12,** S16 (2015).

106. Simões, J. M. *et al.* Social odors conveying dominance and reproductive information induce rapid physiological and neuromolecular changes in a cichlid fish. *BMC Genomics* **16,** 1–13 (2015).

107. Stiver, K. A., Harris, R. M., Townsend, J. P., Hofmann, H. A. & Alonzo, S. H. Neural Gene Expression Profiles and Androgen Levels Underlie Alternative Reproductive Tactics in the Ocellated Wrasse, Symphodus ocellatus. *Ethology* **121,** 152–167 (2015).

108. Harris, R. M. & Hofmann, H. A. Seeing is believing: Dynamic evolution of gene families. *Proc. Natl. Acad. Sci. U. S. A.* **112,** 1252–1253 (2015).

109. Brawand, D. *et al.* The genomic substrate for adaptive radiation in African cichlid fish. *Nature* **513,** 375–81 (2014).

110. Smith, C. C. *et al.* Copy number variation in the melanocortin 4 receptor gene and alternative reproductive tactics the swordtail Xiphophorus multilineatus. *Environ. Biol. Fishes* **98,** 23–33 (2014).

111. Harris, R. M., Dijkstra, P. D. & Hofmann, H. A. Complex structural and regulatory evolution of the pro-opiomelanocortin gene family. *Gen. Comp. Endocrinol.* **195,** 107–115 (2014).

112. Fischer, E. K., Harris, R. M., Hofmann, H. A. & Hoke, K. L. Predator exposure alters stress physiology in guppies across timescales. *Horm. Behav.* **65,** 165–172 (2014).

113. Sessa, A. K., Harris, R. M. & Hofmann, H. A. Sex steroid hormones modulate responses to social challenge and opportunity in males of the monogamous convict cichlid, Amatitliana nigrofasciata. *Gen. Comp. Endocrinol.* **189,** 59–65 (2013).

114. Oldfield, R. G., Harris, R. M., Hendrickson, D. A. & Hofmann, H. A. Arginine vasotocin and androgen pathways are associated with mating system variation in North American cichlid fishes. *Horm. Behav.* **64,** 44–52 (2013).

## Vita …

Rayna Michelle Harris was born and raised in Hallsville, Texas. Rayna received the Bachelor of Science degree in Biochemistry from The University of Texas. After graduation, she was a Teaching Specialists in the Department of Chemistry and Biochemistry at The University of Texas, a bench scientist and SCUBA diver in Costa Rica at the Instituto Nacional de Biodiversidad, and a Spanish teacher at Futura Language Professionals in Austin, Texas. She joined the Hofmann Lab in 2008 as a research technician and lab manager. In 2012, she began her Ph.D. studies, remaining in the Hofmann Lab. While working on her Ph.D., Rayna served as the Course Developer for the Neural Systems and Behavior course at the Marine Biological Laboratory in Woods Hole, MA. During this time, she was also served as the Training and Outreach Coordinator for the Center for Computational Biology and Bioinformatics. In 2015, she joined the Software Carpentry community and served on the Steering Committee from January 2016 - December 2017. In 2016, Rayna was an Instructor Trainer for software Carpentry, which involved with designing lessons and teaching new instructors how to teach. In fall 2017, she was a teaching assistant in the undergraduate Biostatistics course at the University of Texas. At the time of writing, she was weighing a few post-doctoral options.

Permanent email: [rayna.harris@gmail.com](mailto:rayna.harris@gmail.com)

Website: <http://raynamharris.github.io>

ORCID: <https://orcid.org/0000-0002-7943-5650>

This dissertation was typed by Rayna Michelle Harris.

1. There are four co-authors on the corresponding manuscript: Hsin-Yi Kao, Juan Marcos Alarcón, André Fenton, and Hans Hofmann. [↑](#footnote-ref-1)
2. There are four co-authors on the corresponding manuscript: Hsin-Yi Kao, Eric Klann, André Fenton, and Hans Hofmann. [↑](#footnote-ref-2)
3. There are four co-authors on the corresponding manuscript: Hsin-Yi Kao, Juan Marcos Alarcón, André Fenton, and Hans Hofmann. [↑](#footnote-ref-3)
4. There are thee co-authors on the corresponding manuscript: Juan Marcos Alarcón, André Fenton, and Hans Hofmann. [↑](#footnote-ref-4)