Distribution Agreement

In presenting this thesis or dissertation as partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the outline submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as article or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Byron A. Gardner Date

Examination of Brain Chromatin Accessibility in the Prairie Vole (*Microtus ochrogaster)*

By

Byron A. Gardner
Doctor of Philosophy

Graduate Division of Biological and Biomedical Sciences
Neuroscience

_____
Larry Young
Advisor

_____
Brian Dias
Committee Member

_____
Shannon Gourley
Committee Member

_____
Randy Hall
Committee Member

_____
David Katz
Committee Member

Accepted:

_____
Kimberly Jacob Arriola
Dean of the James T. Laney School of Graduate Studies

_____
Date

Examination of Brain Chromatin Accessibility in the Prairie Vole (*Microtus ochrogaster)*

By

Byron A. Gardner
B.S., Marist College, 2012

Advisor: Larry J. Young, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
In Graduate Division of Biological and Biomedical Science, Neuroscience
2022

Abstract

Examination of Brain Chromatin Accessibility in the Prairie Vole (*Microtus ochrogaster)*
By Byron A. Gardner

Oxytocin (OXT) is a small and highly conserved neuropeptide that is known to modulate the neural circuitry that is associated with the processing of social information. Dysfunctions in OXT signaling are considered relevant to a handful of psychiatric conditions, including schizophrenia and autism spectrum disorder. Many scientific efforts to understand the biological mechanisms of the OXT system have involved research of the OXT receptor, OXTR, and the gene that encodes the receptor, *OXTR*. In humans, variation in the *OXTR* sequence has been linked to a wide variety of traits that relate to social cognition and behavior. In animals, behavioral paradigms combined with DNA sequencing and imaging of OXTR in the brain have revealed strong correlations between *Oxtr* single nucleotide polymorphisms (SNPs), OXTR distribution in reward and sensory processing networks, and pro-social phenotypes. Although much has been studied regarding the function of OXT in the brain, and associations between OXTR/*Oxtr* SNPs and behavior have been identified, less is known about underlying molecular genetic factors that contribute to variation in gene expression. To gain insight into potential mechanisms of *OXTR*/*Oxtr* regulation, we established a workflow for examining chromatin accessibility in brain tissue of the socially monogamous and highly affiliative prairie vole, *Microtus ochrogaster*. At its core, the Assay for Transposase Accessible Chromatin and DNA sequencing (ATAC-Seq) allows detection of open regions of DNA throughout the genome. We applied this method in brain areas that are known to differentially express *Oxtr* from previous studies, with the striatum being of specific interest due to a critical role in reward processing. Here, we demonstrate validation of the ATAC-Seq method with prairie vole brain tissue samples, and we perform tests for differential enrichment between the striatum, insular cortex, and inferior colliculus (n=11). Our findings present genes and genetic loci that have the strongest signals of striatum-specific chromatin accessibility. We also provide evidence that sex differences are a predominant factor contributing to variation in chromatin accessibility. We then analyze the prairie vole ATAC-Seq data at *Oxtr*, where we perform variant calling to compare SNP accessibility frequencies between each brain region. Interestingly, we show different sets of SNPs with high accessibility for the striatum compared to insular cortex, each of which we refer to as differentially accessible alleles. These SNPs may serve as candidates for genetic modification of *Oxtr* to alter its transcription in a region-specific manner. Additionally, the ATAC-Seq data presented here provides hundreds of thousands of accessible sites in the prairie vole genome that can be explored in further research. This dissertation is a unique resource that provides some of the first functional characteristics of the prairie vole genome, as well as experimental and analytical tools for further development of chromatin-based assays in the brain.

Examination of Brain Chromatin Accessibility in the Prairie Vole

By

Byron A. Gardner
B.S., Marist College, 2012

Advisor: Larry J. Young, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
In Graduate Division of Biological and Biomedical Science, Neuroscience
2022

Acknowledgements

Emory Neuroscience, I appreciate the knowledge and perspectives on life that were offered to me. Thank you to my committee for their contributions of time, expertise, resources, and patience towards the completion of this work. Larry, thank you for providing me with opportunities and guidance every step of the way. Thank you to my family and friends for love and support. Adin and Cassandra, I love you! The work presented here is dedicated to the memories of Bryson Gardner and Anna Rasch.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

LIST OF ABBREVIATIONS

ASD: Autism Spectrum Disorder
ATAC-seq: Assay for Transposase Accessible Chromatin with DNA Sequencing
AVP: Vasopressin
AVT: Arginine Vasotocin
Ca^{2+}: Calcium
ChIP-Seq: Chromatin immunoprecipitation and sequencing
DNA: Deoxyribose Nucleic Acid
GWAS: Genome Wide Association Study
HPA: hypothalamic–pituitary–adrenal axis
IP3: Inositol Triphosphate
KO: Knock-out
LD: Linkage Disequilibrium
mPFC: medial Prefrontal Cortex
MT: mesotocin
NAcc: Nucleus Accumbens
OXT: Oxytocin – peptide form
Oxt: Oxytocin - gene
OXTR: Oxytocin Receptor – protein
Oxtr: Oxytocin receptor – gene
PCR: Polymerase Chain Reaction
PLC: Phospholipase C
PPT: Partner Preference Test
PVN: paraventricular nucleus
SNP: single nucleotide polymorphisms
SON: supraoptic nucleus
VT: Vasotocin

Chapter 1:

Introduction

GENERAL INTRODUCTION: FROM BRAIN EXTRACTS TO DNA ACCESSIBILITY

The history of neuropeptide research began over 120 years ago and includes two early Nobel prize awards. Early studies in canine pituitary extracts displayed properties of raising blood pressure and slowing heart rate (Oliver and Schafer, 1895). Experiments in sheep then demonstrated that the posterior lobe of the pituitary gland was necessary for healthy blood pressure, milk production, and milk ejection in pregnant females (Howell, 1898). Comparably, in pregnant cats, human posterior pituitary extracts stimulated uterine contractions. It was from these specific findings when term oxytocin (OXT), from the Greek meaning "quick birth," was first used in 1909 (Dale, 1909). Sir Henry Dale later received the Nobel Prize in 1936 "for discoveries relating to chemical transmission of nerve impulses." OXT also became the first peptide to be completely sequenced and synthesized, leading to another Nobel prize (Du Vigneaud, 1954; Du Vigneaud et al., 1953). For several decades, a majority of OXT research then centered around reproductive biology. The development of pharmacological agonists and antagonists in the 1990s provided a foundation to study mechanisms of action for OXT in the periphery and brain, and OXT subsequently became widely studied for potential influences on human and animal social behavior. Scientists across many research fields have since continued to gain more knowledge about OXT; from the DNA sequences that encode OXT and its receptor gene, *OXTR*, to the unique cell-types and physiological mechanisms underlying production and secretion of the peptide.

After years of studying the OXT system in animals, inter- and intraspecies diversity in brain *Oxtr* expression emerged as a focal point in OXT research. Animal models, such as the socially monogamous prairie vole (*Microtus Ochrogaster*), provide a unique opportunity to understand how genetic variation in *Oxtr* contributes to variation in receptor distribution throughout the brain (Young, 1996). Prairie voles, compared to other laboratory reared colonies,

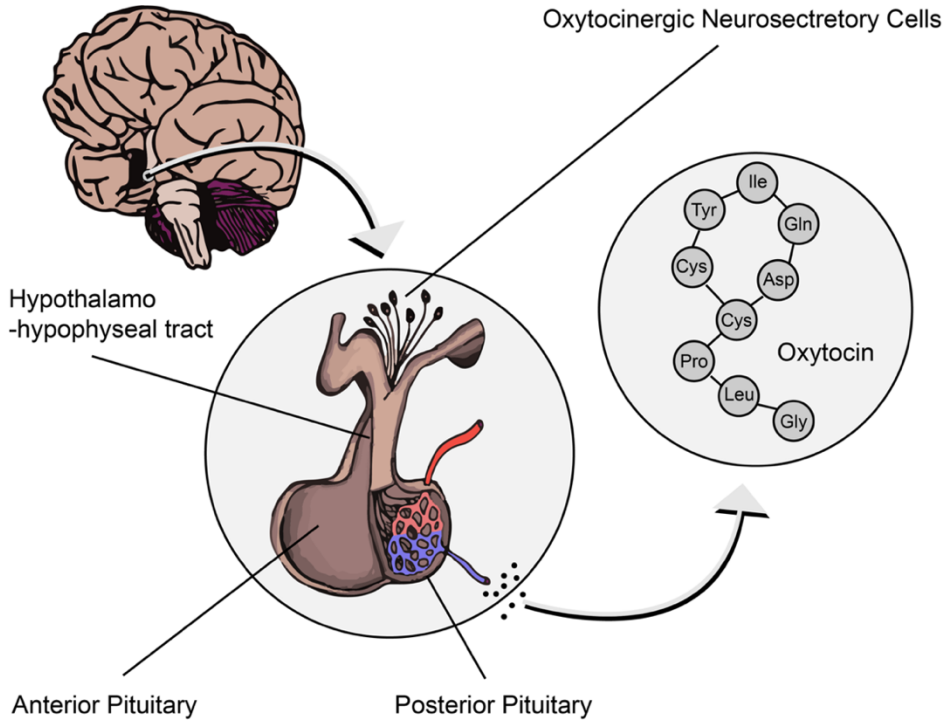
have remarkable variation in brain *Oxtr* expression within regions that are important for processing social information. Research in prairie voles has helped to show how OXT contributes to brain phenotypes and neural activity that mediates social behavior. Interestingly, in prairie voles, a group of single nucleotide polymorphisms (SNPs) spanning the *Oxtr* gene body were found to be associated with receptor distribution in the striatum, as well as affiliative behaviors (King, 2016). The molecular genetic mechanisms by which SNPs in *Oxtr* lead to phenotypic variability, however, remain widely unknown. A major goal of this dissertation is to provide functional information about chromatin regulation in the prairie vole brain. To this end, we have established and validated the Assay for Transposase Accessible Chromatin with DNA Sequencing (ATAC-Seq) for prairie vole brain tissues. ATAC-Seq is a method that involves isolation of nuclei, unique excision/ligation of nucleic acids, PCR amplification, and high throughput DNA Sequencing. The purpose of ATAC-Seq data is to detect accessible chromatin regions within the genome. To date, there are no other studies that have reported the use of ATAC-Seq in prairie voles. Our contribution will be to analyze and compare chromatin accessibility signals from three prairie vole brain areas: striatum, insular cortex, and inferior colliculus.

This dissertation seeks to demonstrate what can be gained from bringing chromatin-based experimental approaches to studying the biological mechanisms of *Oxtr* regulation. We review current knowledge of the OXT system. We explore some of the latest technical and informatics approaches available to study *Oxtr*, and we present findings that examine genome-wide chromatin accessibility in three sub-regions of the prairie vole brain. We then look more closely at the relationship between chromatin accessibility and genetic variation along the *Oxtr* gene. Finally, we summarize the current resources and approaches that may be important to future chromatin studies in the prairie vole model.

OXYTOCIN, VASOPRESSIN, AND SOCIAL BEHAVIOR

Oxytocin (OXT) is a neuropeptide that is known to regulate the salience and reinforcing value of social stimuli. OXT is historically known for its influence on peripheral functions that are related to reproductive physiology, such as uterine contraction and milk ejection (Fuchs and Saito, 1971; Gorewit et al., 1983; Nickerson et al., 1954). OXT is synthesized in the paraventricular nucleus (PVN) and supraoptic nucleus (SON) of the hypothalamus and released into peripheral circulation from terminals of magnocellular neurons at the posterior pituitary, also known as the neurohypophysis (Figure 1.1). The cellular mechanisms of OXT are believed to be primarily through G protein-coupled receptor signaling, and predominantly the G alpha-q pathway (Ku et al., 1995; Sanborn et al., 1998). G_q activation is associated with increase in intracellular calcium (Ca2+) concentration, and signal transduction mechanisms that increase phospholipase C (PLC) activity and inositol triphosphate (IP3) synthesis (Figure 1.2). Like many transmembrane receptors, autoregulation of OXT and its closely related neuropeptide vasopressin (AVP) has also been described (Dayanithi et al., 2000).

Vasopressin (AVP) is another commonly studied neuropeptide, and OXT and AVP differ by only two amino acid positions. Evidence has been found for the cross binding of OXT to AVP receptors, and vice versa (Barberis et al., 1998; Chini et al., 1996). Pharmological studies that have involved treating isolated PVN neurons with neuropeptide receptor antagonists have revealed the differential mechanisms by which OXT and AVP influence cellular Ca^{2+} concentrations (Dayanithi et al., 1996). Increases in Ca^{2+} due to OXT binding are demonstrated to be the result of intracellular release, whereas AVP has been shown to interact directly with voltage gated Ca2+ channels on the cell surface. In humans, the genes encoding OXT and AVP, are separated by just less than 11,000 base pairs, and gene homologs have been identified in virtually all animals, from



Animal	Peptide Name	Sequence
Annelida (worms)	Annetocin	CFVRNCPXG
Mollusca (snails)	Conopressin	CFIRNCPKG
Hymenoptera (ants)	Inotocin	CLIVNCPRG
Osteichthyes (bony fish)	Isotocin	CYISNCPIG
Aves (birds)	Mesotocin	CYIQNCPIG
Mammals (humans)	Oxytocin	CYIQNCPLG

Figure 1.1. Oxytocin Production and Secretion in the brain.
Oxytocin is an evolutionarily conserved neuropeptide with more than a dozen nonapeptide homologs described in vertebrates and invertebrates. Oxytocinergic neurosecretory cells are housed within the paraventricular and supraoptic nuclei of the hypothalamus. Magnocellular neurons descend along the hypothalamo-hypophyseal tract and terminate into circulation. Upon stimulation, OXT is released into the periphery and diffuses centrally into the brain.

birds and humans to insects and worms (Figure 1.1). Only a single nonapeptide is found in invertebrates and primitive vertebrates, and it is well established that a gene duplication of the common ancestor gene, vasotocin, occurred early in vertebrate evolution. Between cyclostomes (lampreys) and bony fishes, for example, two nonapeptide forms can be identified: oxytocin-like and vasopressin-like. Conservation of the OXT peptide sequence across taxa likely promotes the synonymous physiological effects that can be observed. In the earthworm, *Eisenia foetida*, the OXT homolog annetocin induces egg laying behavior (Oumi et al., 1996). In birds and most non-mammalian vertebrates, mesotocin (MT) and vasotocin (VT) function to regulate species-specific sex behaviors, as well as some endocrine functions. For example, in the zebra finch, *Taenioypygia guttata*, OXT-neurons of the PVN contribute to opposite sex aggression and flock size preference (Kelly and Goodson, 2014). In the frog, *Rana esculenta*, MT and VT have been found to stimulate neuro-steroid production (Do-Rego et al., 2006). From the observed similarities in neuropeptide gene and protein sequence between many animals, it has been estimated that prior to the divergence of the last common vertebrate ancestors, roughly 700 million years ago, a gene duplication event of vasotocin likely led to the emergence of the two distinct peptides OXT and AVP (Van Kesteren et al., 1995).

In the central nervous system, receptor distributions for binding OXT and AVP can be highly variable between different species, and even among individuals of the same species in some cases (Young, 1996). Recent human studies have demonstrated a role for the receptor genes in potentially mediating the processing of social information. For example, both cognitive and emotional empathy have been respectively associated with genetic variation in the AVPR1a and *OXTR* genes (Uzefovsky et al., 2015). In a foundational study, Young et al. proposed three potential factors that they believed contribute most to the diversity of neuropeptide receptor

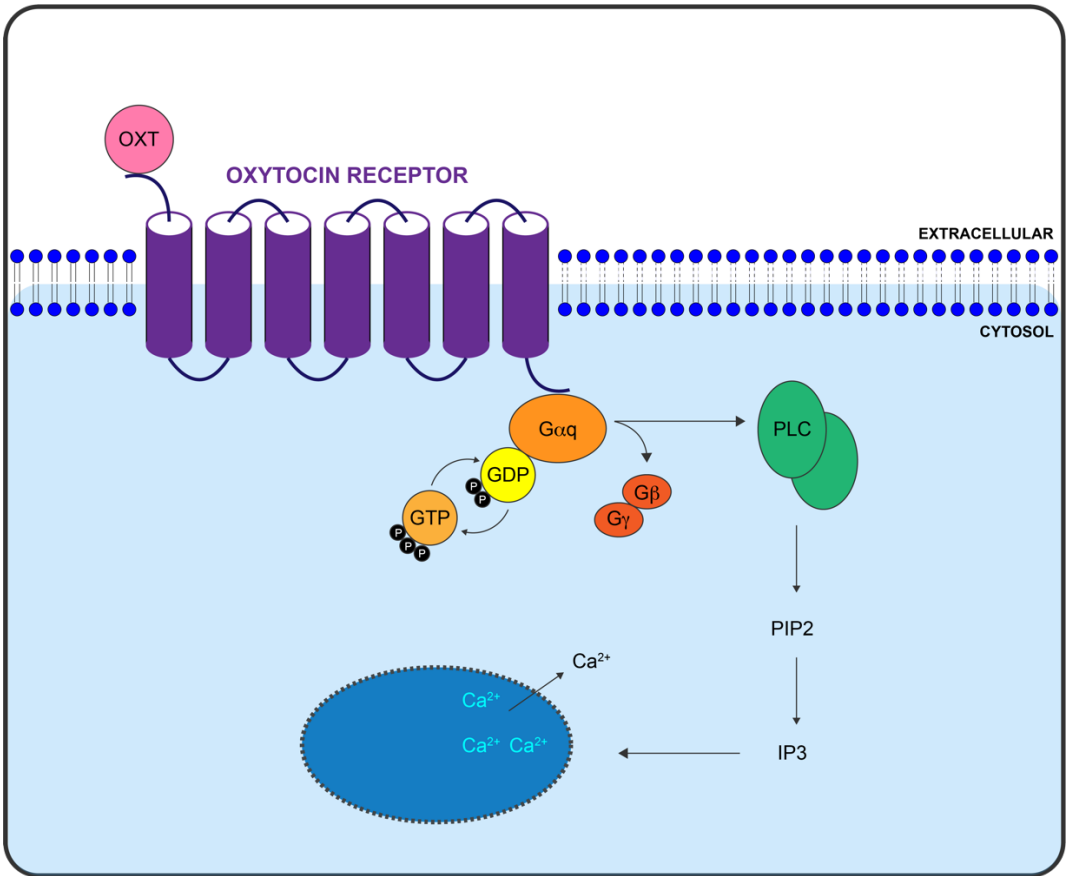


Figure 1.2. OXT Gq Signaling Pathway.
OXT binds to the 7 transmembrane domain G-protein-coupled OXT receptor. The Gq pathway is one of a handful of cellular mechanisms that involve a series signaling cascades that involve the de-coupling of cytosolic GTP binding proteins from the receptor and cell membrane. Gq activation signals increases in Phospholipase C (PLC) production, which directs the hydrolysis of Phosphatidylinositol Biphosphate (PIP2) into Inositol Triphosphate (IP3) and DAG. In neurons and muscle cells, signal transduction mediated by OXT binding are also associated with increases in intracellular Ca^{2+} due to stimulated release from the endoplasmic reticulum.

expression in the brain : Genetic diversity at DNA binding sites, the expression patterns of regulatory proteins in different brain regions, and early life experiences such as the pre- and perinatal environment (Young, 1996). Field research has come to support these early hypotheses, where we can observe balancing selection for the phenotypes that influence social behavior. In prairie voles, for example, AVPR1a density in the retrosplenial cortex interacts with space use and social interaction to predict male fidelity (Okhovat et al., 2015). Based on over a century of research, our understanding of the connection between brain neuropeptide function and variation in behavior now considers that additive phenotypes operate in specific networks of the brain and interact with one another and with the environment.

OXYTOCIN IN THE BRAIN AND SOCIAL COGNITION

Lessons from Rodent and Animal Models

Animal studies have demonstrated that OXT mediates many aspects of social behavior, including parental nurturing, social recognition, empathy, and pair bonding (Burkett et al., 2016; King et al., 2016; Klatt and Goodson, 2013; Ross et al., 2009). In gene knockout (KO) experiments of *Oxt*, KO mice can be more likely to fail at discriminating previously encountered mice from novel strangers (Ferguson et al., 2001; Pobbe et al., 2012). *Oxt* KO mice could, however, distinguish between nonsocial odor cues, which suggests that OXT may function to specifically promote social memory. Some research has also identified sexual dimorphisms in OXT phenotypes, one being plasma oxytocin levels. (Marazziti et al., 2019). In mice, OXT plays a role in modulating neural activity in the auditory cortex to recognize pup calls, therefore promoting maternal behavior in females (Marlin et al., 2015; Mitre et al., 2017). In CB57BL/6 mice, age and sex play a role in the expression level of *Oxtr* in brain areas such as the hypothalamus and olfactory

bulb (Olazabal and Alsina-Llanes, 2016). Evidence for interactions between OXT and other genes has grown in recent years, with some results being linked to specific aspects of information processing. In a KO study of *Cntnap2*, a gene found to be associated with ASD phenotypes, mice lacking the gene showed a reduction in OXT producing neurons (Penagarikano et al., 2015). Similarly, in a sensory deprivation model using neonatal mice, OXT secreting hypothalamic neurons of mice that were reared in the dark exhibited reduced firing rate and OXT secretion (Zheng et al., 2014). In this same study, OXT signaling was shown to regulate experience-dependent cortical development in the somatosensory, auditory, and visual cortices. Interestingly, treatment of OXT injection or environmental enrichment were able to elevate excitatory synaptic transmission and was able to rescue the effects of sensory deprivation. These studies suggest that the combination of OXT signaling and environmental cues likely work together to support social development.

OXTR Polymorphisms in Clinical Research

Several clinical studies have focused on the interaction between OXT and *OXTR* variation in social contexts. In one such study, exogenous OXT administration was combined with a socioemotional task in a sample of 203 men (Chen et al., 2015). Each of the men were genotyped for a collection of commonly carried genetic variants known as a haplotype block, which, for this experiment included six single nucleotide polymorphisms (SNPs) spanning the *OXTR* promoter and intron 3. The researchers found that a TTCGGG haplotype was positively associated with increased emotion recognition when given exogenous OXT compared to a placebo. These results suggest that an interaction exists between individual base pair variation at *OXTR* and behavioral response to OXT. Some researchers have, however, expressed open criticisms regarding the

mechanisms of action for intranasal OXT in reaching the brain (Leng and Ludwig, 2016). Other scientists have postulated that a more indirect role for the effect of exogenous OXT administration could exist, where trace amounts are able to stimulate local endogenous release of the peptide (Quintana et al., 2019). Despite conflicting perspectives regarding a mechanism of action for peripheral OXT, what cannot be ignored are the many independent lines of evidence for the role of OXT in mediating socially specific behaviors. For example, in an article that represented three studies of small student populations in Singapore, Haifa, and Amsterdam, creative cognition was found to be independently linked to plasma OXT levels, *OXTR* variation, and administration of intranasal OXT (De Dreu et al., 2014).

The positive effects of intranasal OXT on social cognition and behavior should be taken into consideration with the negative effects that have been demonstrated as well. Research has shown that chronic intranasal OXT inhibits partner preference in prairie voles (Bales et al., 2013). Another study examined a social stress paradigm in pigs, where researchers hypothesized that OXT administration would buffer the effects of early life stress. In contrast, the group found that repeated exposure of intranasal oxytocin produced dysregulation of the HPA axis and increased aggressive behavior (Rault et al., 2013). Taken together, these results indicate that OXT signaling may be subject to some degree of physiological tolerance, and that variation in response to OXT can influence behavioral outcomes. Experimentation for links between *OXTR* variation and social cognition has helped to identify dozens of candidate SNPs, which may play a functional role in regulating *OXTR* expression. For example, the *OXTR* polymorphism rs53576 is seen in higher frequency among sales professionals with high motivation to help customers, and also in individuals with stronger facial recognition skills (Verbeke et al., 2013). In a larger sample of 1081 European American and 214 African American males, the SNP rs1042778 was found to be

associated with size and diversity of social networks and support systems (Creswell et al., 2015). In a study cohort of 530 White males, the intron *OXTR* SNP rs2268493 was found to be associated with Asperger Syndrome (Di Napoli et al., 2014). Similarly, in the 5' noncoding region of the *OXTR* gene, rs2270465 was found in higher frequencies in individuals with high functioning autism (Wermter et al., 2010). Another study investigated the haplotype structure of *OXTR* and *AVPR1*a from 14 populations that are represented in the 1000 Genomes Project, and polymorphisms spanning intron 3 of *OXTR* were found to be under positive directional selection based on three independent statistical approaches (Schaschl et al., 2015). This study also highlighted rs2268496, which they suggest could be under balancing selection of both major and minor allele, across all populations. Taken together, there is a continuously growing line of evidence that suggests SNPs spanning *OXTR* may have perhaps evolved at different rates among different ancestral haplotypes.

Many SNPs that meet significance from genetic association studies are also likely in some level of correlation with other nearby SNPs, making it difficult to pinpoint the regions of DNA that could be bound or influenced by the activity of regulatory proteins. The OXTR SNP rs2268496 is, however, located within the binding sequence for the insulator protein CTCF. Also known as the CCCTC binding factor, CTCF is a member of the zinc-finger protein family and is well-known for defining chromatin boundaries to influence transcription signals. CTCF has also been found to act directly as a transcription factor, to activate or repress gene expression across the genome (Fang et al., 2015). Although the consensus DNA binding site for CTCF can be identified in the OXTR intron of multiple species, no causal link has been identified between polymorphisms and the CTCF protein. The paternally expressed gene 3, *Peg3*, is currently the only DNA-binding protein that has been demonstrated to directly regulate *Oxtr* (Frey et al., 2018).

Towards Molecular Mechanisms of OXTR Regulation

The lack of molecular data that can be focused on *Oxtr* regulation continues to complicate a scientific understanding of how variation in the gene underlies social aspects of behavior and cognition. Information that can bridge our knowledge of genetic variation and DNA interactions could push forward many current perspectives from correlations to more basal molecular factors that are involved in mediating OXT related behaviors. Biological variables that contribute to complex phenotypes such cognition and behavior also extend to modifications of the DNA and the protein complexes that form chromatin, such as nucleosome subunits known as histones. DNA is typically methylated between C-G dinucleotides, and histones have amino acid chains that can also be methylated and phosphorylated. Modifications such as methylation that can influence gene expression independent of DNA sequence are defined as epigenetic factors. Methylated DNA can be detected methodologically with sodium bisulfite treatment and DNA sequencing, or in the case of detecting histone modifications, with chromatin immunoprecipitation and sequencing (ChIP-Seq). With bisulfite treatment, the chemical converts unmethylated cysteine (C) residues into uracil (U) through a process known as deamination, and after PCR (polymerase chain reaction) the U's are copied as thymine (T); leaving the remaining C-G dinucleotides to be inferred as methylated. The ChIP-Seq method involves formaldehyde fixation of DNA-protein complexes with DNA shearing and immunohistochemistry. The end result of ChIP-Seq are DNA libraries (PCR amplicons) only believed to be associated with a protein of interest (Mardis, 2007). When used with proper controls, bisulfite sequencing and ChIP-Seq have provided additional layers of knowledge to research that seeks to look beyond genetic variation and into the more realistic regulatory environment of DNA. Early work in humans and nonhuman primates has demonstrated that assays inferring on regulation of *OXT* and *OXTR* may aid in our understanding of how

experiences may shape molecular phenotypes. For example, children who experience childhood abuse also exhibit higher *OXTR* intron methylation (Smearman et al., 2016). Tri-methylation of the 4th lysine residue of histone 3 (H3K4me3), a typical marker for active promoters of genes, is decreased at the *Oxtr* exon 2 in hippocampal samples of adult male rhesus macaques that experienced maternal deprivation compared to males that were reared normally (Baker et al., 2017). These studies suggest a potential connection between peripheral epigenetic modifications of *Oxtr* and neural circuitry underlying social information processing during development.

OXTR SNPs are often prime candidates for gene-of-interest studies that involve novel and expensive measurement techniques combined with small numbers of participants from similar ethnic backgrounds. The results of many behavioral genetic associations can therefore be heavily biased from the perspectives of measurement, construct, and correction for number of comparisons. Nevertheless, several associations have been found between *OXTR* SNPs and social outcomes, including decision making in economic games, optimism, self-esteem, support seeking, and reactions to betrayal (Israel et al., 2008; Kim et al., 2010; Saphire-Bernstein et al., 2011; Tabak et al., 2014). Further, it can be observed that several polymorphisms tied to variation in behavior among healthy subjects, such as rs53576 and rs2254298, have also been reported as statistically significant in autism studies. The opposing direction of behavioral effects of the same SNPs could be explained by varying degrees of relation to additive genotypes that are associated with general social functioning. This trend of intercorrelation of genetic variants is known as linkage disequilibrium (LD). LD is higher among closely related individuals, and higher as the distance between any two nucleotides is decreased within any heritable DNA sequence. Minimizing LD can help to identify functional variants, and can be achieved by studying more diverse groups, or by combining data from several independent studies, also known as a meta-analysis. One of the

best to date for *OXTR* was performed for 16 *OXTR* SNPs in 3941 individuals, and interestingly, 4 of the polymorphisms survived statistical significance for an association with autism (LoParo and Waldman, 2015). However, what remains unclear is the mechanism by which OXTR SNPs can alter DNA/chromatin state and gene transcription. Many SNPs that are identified in behavioral association studies are not further characterized for biological function. Below, we present research in one model organism that can provide promising insight into the neurobiology of OXT function and *OXT*R regulation: prairie voles.

PRAIRIE VOLES AS A MODEL FOR UNDERSTANDING OXYTOCIN FUNCTION

Microtus ochrogaster, also known as the socially monogamous prairie vole, is a microtine rodent species that is native to the North American plains, and one of the most valuable research models for building a systems level perspective on the neurobiological mechanisms that are associated with OXT functioning. Prairie voles are highly affiliative, and for many male and female voles, mating facilitates the development of a life-long pair bond, and biparental care of offspring (Carter et al., 1995; Carter and Getz, 1993). Comparatively, close evolutionary relatives of *M. Ochrogaster*, montane voles (*M. montanus*), predominantly display solitary and asocial mating strategies (Insel and Shapiro, 1992). The more popular monogamous mating strategy of prairie voles can be taken in contrast to field experiments which have found that some male prairie voles, in fact, overlap similar territories, and factors such as resources and predation may ultimately influence the ratio of single to biparental care of young within some vole populations (Cushing et al., 2001; Okhovat et al., 2015).

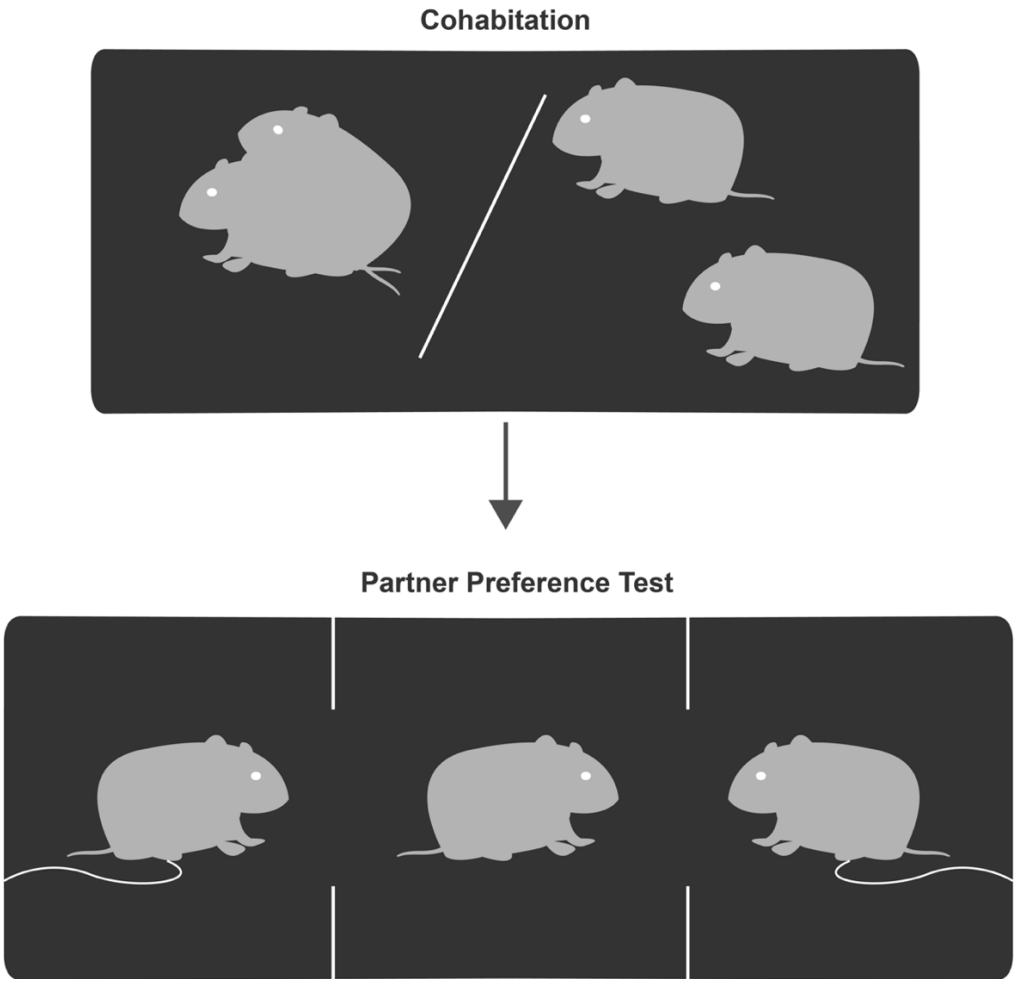


Figure 1.3. Prairie Vole Partner Preference Behavioral Test
An experimental subject is given time to freely interact with an opposite sex conspecific, also known as the cohabitation period. If mating occurs during this time, both animals form a social bond. A bond is usually not formed if the animals do not mate or spend a significant amount of time making physical contact. In the next phase, the partner preference test, the strength of the social bond that is formed is tested in a three chambered arena. The familiar individual and a novel opposite sex conspecific are tethered to opposite ends of the aera, and the experimental animal, unrestrained, is placed into the center. The amount of time the experimental animal spends with the other two animals is then recorded.

In the laboratory, a behavioral test known as Partner Preference (PPT) is used to estimate the presence of a bond by giving individuals the option of spending time with a previously cohabited mate or a stranger of the opposite sex (Figure 1.3). Under normal conditions, prairie voles will spend significantly more time with a previously cohabited partner (Williams et al., 1992). Researchers can then compare more specific behavioral results such as huddling, grooming, and mating behavior under varying conditions. OXT receptor agonists, antagonists, and genetic manipulations of *Oxt* and *Oxtr* have all been reported using the PPT paradigm (Cho et al., 1999; Lim et al., 2004; Williams et al., 1994). In one prairie vole study, PPT helped to demonstrate that OXT function may interact with dopamine neurons in reward areas of the brain to function as a social reinforcement mechanism (Aragona et al., 2003). Similarly, partner preference formation is inhibited in female prairie voles with repeated exposure to amphetamine stimulants, where neural circuitry within the striatum may be altered (Young et al., 2014). Indeed, the effect of amphetamines is mostly explained by reductions in *Oxtr* and dopamine receptor 2 (*Drd2*) expression in the medial-prefrontal cortex (mPFC) and nucleus accumbens (NAcc). In some animals, researchers were ultimately able to restore partner preference behavior following central infusion of an OXTR agonist into the mPFC.

Alloparental behavior in female prairie voles is also known to be mediated by NAcc *Oxtr* expression. Virgin females that spontaneously display maternal behavior such as licking and grooming towards unrelated pups also display higher average density of NAcc *Oxtr* levels (Olazabal and Young, 2006). Site-specific infusion of an OXTR antagonist into the NAcc blocks alloparental behavior, whereas upregulation of *Oxtr* using viral-mediated gene transfer enhances alloparental behavior (Keebaugh and Young, 2011). Later experiments that employed the

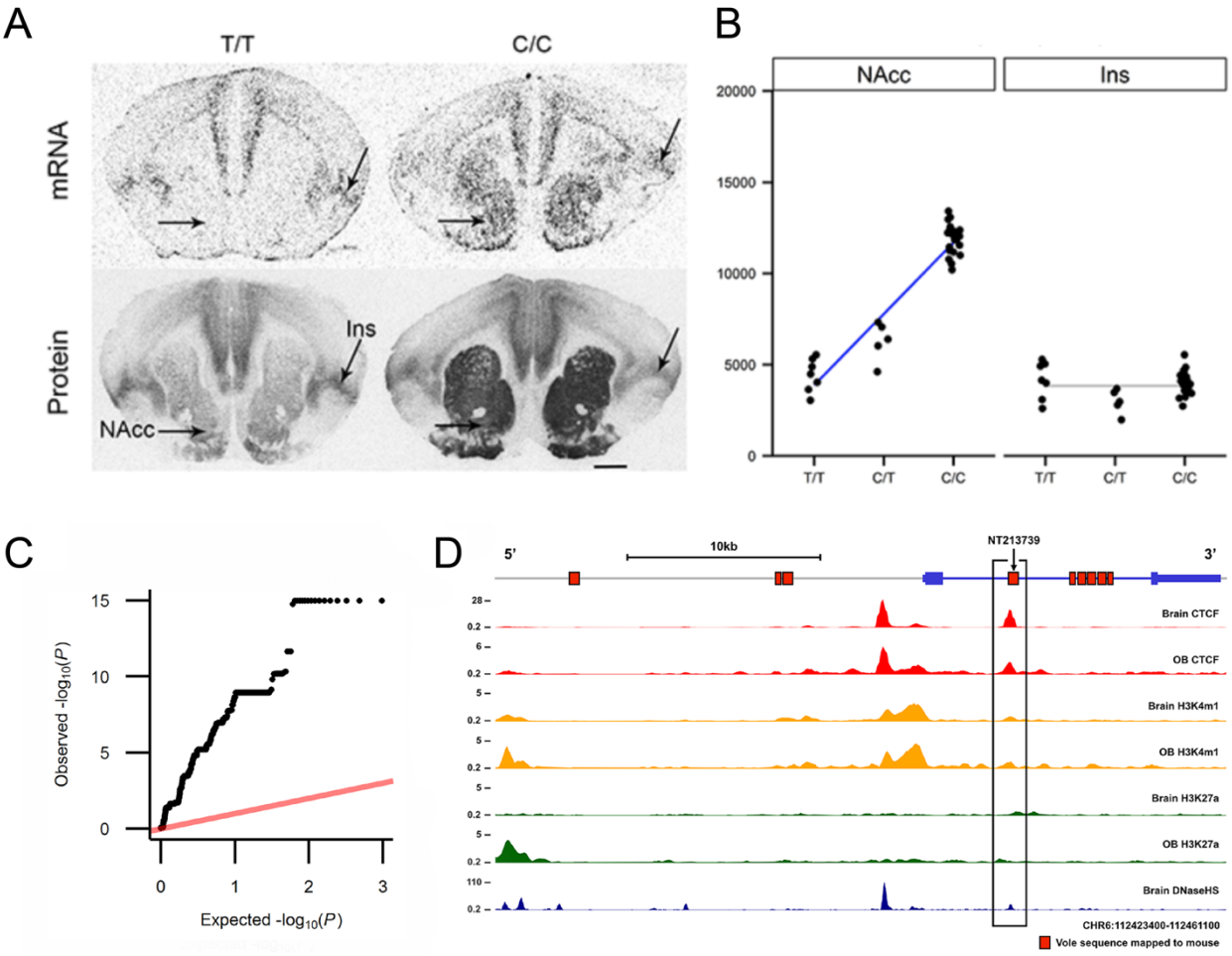


Figure 1.4. Prairie Vole *Oxtr* SNPs Predict NAcc Expression
(A) Representative brain sections from prairie vole imaged for Oxtr mRNA (in situ hybridization) and OXT receptor distribution (autoradiography). Right column C/C genotype individual can be seen with elevated levels of mRNA and protein in the Nucleus Accumbens (NAcc) compared to the T/T genotype. The Insular Cortex (Ins) expression levels are indistinguishable between brain sections and not predicted by Oxtr genotype. (B) Receptor autoradiography measurements for NAcc and Ins compared by genotype (n=84). A strong correlation is identified between NT 213739 genotype and NAcc OXT receptor density. (C) Plot of Observed/Expected P value ratio for all genotyped prairie vole SNPS. The SNPs explaining the most variation ($2x10^{-14}$) are also in 100% linkage disequilibrium, indicated also by the distribution of identical P values (horizontal) near the top of the graph. (D) Prairie Vole *Oxtr* gene diagram, coding sequences (blue bars), with top significant SNPs annotated (red bars). Below are aligned ENCODE data from Mouse brain and olfactory bulb for CTCF transcription factor binding and histone methylation/acetylation. Peaks indicating transcription factor occupancy can be seen spanning the gene, and particularly near one of the significant SNPs, NT213739 (King et al., 2016).

combination of mating and post-mortem analysis of Fos expression, an immediate early gene used as a proxy for neuronal activation, found that an OXTR antagonist was able to reduce the correlation of neuronal activation that occurs between reward areas during mating (Johnson et al., 2016). These findings provide direct evidence that OXT signaling may function as an integrator of communication between an entire brain network that is specifically tuned for the processing of socially relevant cues. Indeed, other work has demonstrated that modulation of OXT signaling in the mPFC-NAcc network using optogenetics can facilitate partner preference behavior between strangers. (Amadei et al., 2017). Decades of research with prairie voles has led to the establishment of methodologies that continue to advance our understanding of how OXT signaling within reward circuitry is involved the neurobiology of social cognition and behavior.

Many efforts have been made to bring the latest tools from the field of genetics to the prairie vole research model. In 2012, 76 genes, including 21 known to be relevant to behavior, were sequenced in prairie voles for the first time with the use of bacterial plasmid cloning (McGraw et al., 2012) The first prairie vole genome, MicOch1.0, was also sequenced in 2012, with the first annotation completed in 2017. Targeted sequencing approaches have been performed in large samples of prairie voles to define variation spanning *Oxtr*. Sequencing data from these animals combined with autoradiogram data for OXT receptor binding affinity has led to the discovery of SNPs in the prairie vole *Oxtr* that are strongly associated with receptor density in specific regions of the brain, and with pair bonding behavior (King et al., 2016). King et al. demonstrated that up to 80% of variation in OXTR density in the NAcc may be explained by an LD block containing less than 12 SNPs along the gene body of Oxtr. When aligning these predictive genetic variants from the prairie vole *Oxtr* to the mouse *Oxtr* sequence in the ENCODE database, a subset of these polymorphisms could be found near transcription factor binding sites

and positive signals for histone modifications known to be present at gene enhancers (Figure 1.5). Work from this study has highlighted that in other brain areas of the prairie vole, such as the insular cortex, there is much less individual variation of *Oxtr* expression. Our laboratory has also demonstrated that that in brain areas such as the inferior colliculus, there is little to no expression of *Oxtr*. A follow up goal for the SNP associations in prairie voles has been to provide information that can help to infer on the complex regulatory processes that may lead to differences in *Oxtr* expression across different areas of the brain. With the increased availability of new experimental tools in DNA library preparation and Next Generation Sequencing, we now can begin to generate datasets that will inform previously unknown characteristics of the prairie vole genome, such as the regulatory landscape of chromatin in the brain.

CONCLUSIONS

Years of work has come to show that OXT plays a significant role in regulating the salience and reinforcing value of social stimuli, individual variation in OXTR can influence social phenotypes, and brain areas such as the striatum and encapsulated NAcc are important in processing social reward. What remains largely unexplored is the mechanistic relationship between brain-region specific effects of genetic polymorphisms and actual physical chromatin availability or transcription factor binding. To date, no data has been generated that provides a picture of chromatin accessibility in the prairie vole brain. A major focus of this dissertation will be to provide the first deep sequenced brain-region specific dataset for the prairie vole species.

In Chapter 1, we provided an overview of OXT function, signaling mechanisms, and behavioral research that suggests OXT plays a critical role in regulating social behaviors in animals, as well as humans. We have also pointed to evidence that variation in expression of *Oxtr*

in the brain contributes to diversity in behavior across species, and between individuals of the same species. Moreover, information on specific SNPs in *Oxtr* that are statistically significant in genome wide association studies (GWAS) are reported here. The prairie vole is then presented as a model for studying OXT function and *Oxtr* regulation.

In Chapter 2, chromatin accessibility data in the form of 2.1 billion Next Generation Sequencing DNA reads that represent 11 samples from 3 prairie vole brain areas is provided, with both sexes included. Next, the ATAC-Seq approach is validated, and features of whole genome chromatin accessibility are identified. Some of the latest software packages in genomics analysis were then used to quantify differentially enriched accessible regions between each brain area. Further, we demonstrate a more predominant relationship between sex and chromatin accessibility at a genome-wide level.

In Chapter 3, we focus on the *Oxtr* gene to align accessible regions of chromatin with previously identified SNPs that have been associated with *Oxtr* expression. Variant calling along the *Oxtr* gene locus was performed with ATAC-Seq reads that represent different brain tissues from the same pooled animals, for the purpose of identifying variation in allele accessibility across brain tissues. Overall, we observe high individual variability in DNA accessibility at *Oxtr*. Some evidence is provided for differentially accessible SNPs between the striatum and insular cortex, which could be important for identifying tissue specific transcriptional activity. Finally, in Chapter 4, we summarize the current achievements of prairie vole genomics and their limitations, and we briefly overview the next experimental approaches that may help to identify how individual *Oxtr* base pair positions can functionally contribute to variation brain region specific gene expression.

Chapter 2:

Chromatin Accessibility in the Prairie Vole Brain

ABSTRACT

Prairie voles (*M. Ochrogaster*) are a unique mammalian species that are frequently studied for affiliative behavior, socially monogamous mating, and biparental care. Recent targeted DNA sequencing projects in voles have identified a strong correlation between single nucleotide polymorphisms (SNPs) in the prairie vole oxytocin receptor gene (*Oxtr*) and variation in oxytocin receptor protein levels. Moreover, early prairie vole genetic associations have been localized to specific areas of the brain that are important for social information processing, such as the striatum. Previous work has demonstrated that increased *Oxtr* expression in the nucleus accumbens, a sub-region of the striatum, is related to increased pair bonding and alloparental behavior. Several Genome Wide Association Studies (GWAS) in humans have also outlined a relationship between *OXTR* SNPs and various aspects of social behavior and cognition. The mechanisms by which *OXTR*/*Oxtr* can be differentially regulated across various neuronal populations and tissues of the brain is, however, largely unknown. In the present study, we report the development of a workflow for the Assay for Transposase Accessible Chromatin with DNA Sequencing (ATAC-Seq) in three regions of the prairie vole brain. A genome-wide comparison of the ATAC-Seq signal between the striatum, insular cortex and inferior colliculus reveals that in the striatum, the predominant differentially enriched regions include genes known to function in DNA binding and transcriptional regulation. We also find that sex explains up to 69% of variation among 65,000 peaks tested for differential enrichment. Sex specific chromatin accessibility can be identified globally on the X chromosome, and at autosomal genes that are known to function in g-protein signaling and neuropeptide transport. These findings serve as the first documented source of genome-wide chromatin accessibility in various prairie vole brain tissues and add support to a growing line of evidence for sex differences in gene regulation within the brain.

INTRODUCTION

Chromatin organization is a critical factor in the development of phenotypic complexity, especially in the mammalian brain (Dulac, 2010; Harabula and Pombo, 2021; Su et al., 2017). Understanding precise regulatory interactions that influence the unique expression patterns of genes in specific cell and tissue types are a logical step towards unraveling mechanisms of disease risk and resilience. Genome-wide association studies (GWAS), which involve statistical comparisons between human traits and millions of individual positions that are variable in the human genome, have helped to identify specific variants that predict risk for many diseases. The ability to genotype previously identified individual base pair variants known as single nucleotide polymorphisms (SNPs) have subsequently promoted further discoveries of genetic associations with social cognition and behavior in humans. The oxytocin receptor gene, *OXTR*, is one very popular gene for testing SNPs with aspects of social functioning (Israel et al., 2009; Kim et al., 2010; Lerer et al., 2008; Saphire-Bernstein et al., 2011; Wermter et al., 2010). In the more recent post-GWAS era, physical characterization of DNA at regions identified from association studies are of great interest to biomedical researchers and clinicians (Freedman et al., 2011; Gallagher and Chen-Plotkin, 2018; Wijmenga and Zhernakova, 2018). To date, few opportunities have been presented to explore aspects of gene regulation in the brain in vivo for *OXTR*/*Oxtr*.

Prairie voles are an important research model for studying the neurogenetics of social cognition and behavior. They are highly affiliative and form long term pair bonds, which often results in biparental care of offspring and partner preference for one another over strangers (Williams et al., 1992). Prairie vole research has demonstrated that the oxytocin system is involved in mediating partner preference behavior, and specifically the processing of socially relevant information (Burkett et al., 2016; Olazabal and Young, 2006; Ross et al., 2009). In prairie voles,

specific brain areas such as the striatum and pre-frontal cortex have been shown to coordinate via oxytocin signaling to promote affiliative behavior (Amadei et al., 2017). Single nucleotide polymorphisms (SNPs) have also been identified to correlate with region specific expression of the oxytocin receptor gene, *Oxtr,* in the striatum (King et al., 2016). With the publishing of the first prairie vole genome, MicOch1.0, and the availability of some of the first annotations for this species, research efforts within this model organism can now begin to characterize the regulatory landscape of behaviorally relevant genes such as *Oxtr*. To this end, a major consideration for the development of functional genetic information in the prairie vole model is the inclusion of both sexes, as numerous studies have highlighted substantial differences in neural gene regulation between males and females (Gegenhuber and Tollkuhn, 2020; Trabzuni et al., 2013; Xu et al., 2002; Xu et al., 2008).

The Assay for Transposase Accessible Chromatin (ATAC-Seq) is a more recently developed DNA sequencing method that allows for the detection of accessible regions of chromatin (Buenrostro et al., 2013; Buenrostro et al., 2015). ATAC-Seq involves careful quantification of starting material in the form of viable cell nuclei. The method then combines novel forms of DNA library preparation and analysis of sequence data to measure the availability of DNA at any given position in the genome. Evidence has demonstrated that binding sites for regulatory factors can, in fact, also be confidently inferred from ATAC-Seq data (Bentsen et al., 2020; Li et al., 2019). Reporting on the use of ATAC-Seq has grown substantially in recent years, and many laboratories are performing experiments and analyzing chromatin data for the first time in several model species (Fullard et al., 2018; Jauregui-Lozano et al., 2021; Liu et al., 2019). Early work in the brain has been particularly focused on mechanisms of disease and development (de la Torre-Ubieta et al., 2018; Fullard et al., 2017; Wang et al., 2018). In humans, there is already some

evidence that chromatin accessibility is altered by neural activity, and perturbations in chromatin structure may underlie disruptions at the neuronal and circuit level that increase risk for psychiatric conditions (Rizzardi et al., 2019; Su et al., 2017). The importance of understanding changes in chromatin accessibility was further demonstrated in a recent report on the effects of prenatal e-cigarette exposure to the developing rat brain, where markers for excitatory vs. inhibitory neurons could be established from single cell ATAC-Seq data (Chen et al., 2022). As the method becomes more commonly used, chromatin accessibility data is proving to be essential to classifying unique tissues and cell-types at a level that precedes gene expression and physiology.

To move towards a mechanistic understanding of the functional roles for DNA variation in behaviorally relevant gene regulation, we find it vital to establish reproducible methodologies that can combine *in vivo* brain dissection and whole genome sequencing of multiple samples. ATAC-Seq stands out as one of few chromatin-based experiments that can be performed same day from neural tissue to DNA libraries. Here, were employed ATAC-Seq to examine chromatin accessibility in three areas of the prairie vole brain: the striatum, insula cortex, and inferior colliculus. Our experimental design specifically seeks to capitalize on previous findings related to striatum-specific expression of *Oxtr*, which showed that the striatum expresses high levels of *Oxtr* that robustly varies with SNPs, the insular cortex expresses high levels of *Oxtr* that does not vary with genotype, and inferior colliculus does not express *Oxtr* (King et al., 2016). This resource can serve as a foundational genomic dataset for prairie vole research, while contributing towards the discovery of tissue and cell-type specific genetic regulatory mechanisms within the context of studying behavior.

METHODS

Animals

Animals were housed at the Emory National Primate Research Center vivarium with Enviro-Dri enrichment and cotton squares for nesting. Light cycles used for prairie voles were 7AM-9PM, and a standard diet consisted of 3-4 large milled rabbit pellets that were replenished 2-3 times per week. Same sex littermates were pooled in groups of 3 for each ATAC-Seq sample (n=11). An important consideration for this study was to minimize relatedness between the duplicates for each brain region and sex. Our vole colony represents field captured voles from Illinois, and we maintain regular outbreeding cycles to promote overall genetic diversity. Breeding is tracked within the colony to produce vole kinship coefficients that can used to approximate the degree of relatedness between any two litters. We used this information to ensure that the samples collected for this study (striatum, insular cortex, and inferior colliculus) were from distinctly unrelated prairie vole litters, with a degree of relatedness calculated to be 0.02-0.08.

Dissection and Pooling

The striatum, insular cortex, and inferior colliculus were individually dissected from adult prairie voles aged 6-9 months. Voles were CO_2 euthanized in cages of 3, following immediate decapitation and brain dissection on a cold, flat 4C aluminum surface with sharp razor blades. A 1mM coronal section approximately 1.8-2mm anterior to bregma was taken, containing the striatum and insular cortex, of which approximately 1-2mg of brain tissue was collected from each hemisphere and immediately stored in a 15mL centrifugation tube containing cold 1X HBSS. Inferior colliculi were then removed from the brain and processed as described for striatum and insula. After the pooling of brain regions into individual tubes that represented each of the six

conditions (female striatum, male striatum, female insular cortex, male insular cortex, female inferior colliculus, and male inferior colliculus), samples were cleaned three times with 1X cold PBS and moved immediately to dissociation.

Cell Lysis and Nuclei Quantification

Cell lysis was performed using MACS adult brain dissociation kit as described by the manufacturer (Miltenyi Biotec, Gladbach, Germany). The end-product resulted in cell lysate volumes of approximately 500uL per sample. 5-10uL of each sample were then separately stained with trypan blue (Sigma-Aldrich, St Louis, MO) and visualized using a standard hematocytometer. Viability estimates were calculated to be approximately 80-85% for insular cortex and the inferior colliculus samples, and 50-60% for striatum samples. The viability estimates were used to quantify nuclei per μL for each lysate. Individually dispensed aliquots that contained approximately 100K freshly isolated nuclei per sample were then immediately used as input material for the transposition reaction.

Transposition and Library Amplification

Transposition, adapter ligation, and library amplification were carried out using the Nextera Flex Prep Library Kit (Juno Beach, FL). During transposition the regions of chromatin that are more frequently accessible to the integrase enzyme Tn5 will be cut and ligated to sequencing adapters. Provided in the kit, unique dual indexes for each reaction were used as primers to allow for the pooling of samples and multiplex sequencing. A PCR program for 12 cycles; the recommendation for 100K nuclei samples, was then created to amplify ATAC-Seq DNA libraries.

Quality Control, Sequencing and Alignment

Samples were first sent to the Emory Primate Research Center Genomics Core for quality assurance and library quantification, and then to The Hudson Alpha-Discovery Life Sciences department of genomics services (Huntsville, Al) for further validation and whole genome sequencing. Bioanalysis traces of DNA libraries were used to determine fragment size ranges of 150-1000 base pairs, a 350 bp target insert size, and library concentrations of 1-8ng/μL. Next generation sequencing was performed on 11 final samples (5 Female = 2 striatum, 2 insular cortex, 1 inferior colliculus; 6 Male = 2 striatum, 2 insular cortex, 2 inferior colliculus). Samples were sequenced on a single Novaseq S4 flow cell and 150 bp paired end (PE) strands were generated for 2.1 billion reads, yielding a range of 150M to 250M reads per sample. Raw data was provided in two .fastq files per sample, representing forward (R1) and reverse reads (R2). Files were downloaded from the Hudson Alpha server, concatenated depending on unique molecular index, and quality assessed for read quality and sequence duplications using FastQC. Sequencing reads with Phred Q>30 and no duplications consisted of 93-97% of the 11 libraries' raw read quantity and were then aligned to the MicOch1.0 assembly using the Bowtie2 software package. Parameters for alignment include: --very-sensitive -p <number of cores> -x <bowtieindex> -1 <R1.fastq.gz > -2 <R2.fastq.gz >.

Data Analysis

Peak Calling

All sequencing analysis were performed within the Mac OS UNIX command line. PCR duplicates and reads from mitochondrial DNA were removed using Samtools (github.com/samtools). Peak calling was performed using the version 2 software of the Model-

based Analysis for Chip-Seq (MACS2) software package (github.com/MACS). A corrected p value cutoff of 0.05 was used as a threshold for defining peak regions, and the -f BAMPE flag was used to indicate peak calling for only properly paired reads. Total peak quantities for the 20 largest chromosomes of the MicOch1.0 assembly were output using Samtools and visualized with R.

Differential Enrichment

Differential enrichment of all shared peaks was quantified with DeSeq2 and executed with the GUAVA-1 MacOS application. R Outputs from DeSeq2 code include "gained open" and "gained closed" regions from each brain region and a principal component analysis (PCA). P-value sorted output containing top striatum differentially enriched regions was generated in R and written as .csv file. Gene interaction networks for representative striatum enriched genes was performed within the UCSC genome browser. MichOch1.0 chromosomes are not named with respect to karyotypes that have been established in the prairie vole model, and there is a lack of annotation for many of the chromosomal regions identified within the top differentially enriched peaks. To overcome this, all 20 positions were cross-referenced with Genescan predictions using the Ensemble database (https://m.ensembl.org/Microtus_ochrogaster/Info/Annotation).

Visualization

Visualizations were produced from quality filtered reads that are normalized by ATAC-Seq signal based on the entire group of samples, rather than from variation within individual samples. .bam files were converted into a compressed .tdf format with IGV tools and visualized with IGV. Heatmaps were produced with IGV tools bar viewer. Volcano and PCA plots were

produced as outputs from the DESeq2 package. Bar plots and tables were produced using R version 3.0 and/or Microsoft Excel.

RESULTS

Prairie Vole ATAC Workflow

Our methods predominantly followed the established ATAC-Seq protocol as previously described (Buenrostro et al., 2013; Buenrostro et al., 2015). We included modifications to accommodate the technical limitations of our laboratory and provide some the latest available materials and methods for PCR amplification and data analysis. Newly developed brain dissociation methods, DNA library prep, and unique indexing primers, were integrated into our workflow to allow us to maintain a high quality first assessment of chromatin accessibility in the prairie vole brain. As such, data is presented from 11 individually transposed ATAC-Seq DNA libraries, including 4 striatum samples (two males and two female), 4 insular cortex samples (two male and two female), and 3 inferior colliculus samples (one female and two male). Each individual sample comprises a pooled triplicate of same sex littermates. Brain samples from striatum, insular cortex, and inferior colliculus were used to represent brain regions known to have different ranges of individual variation in *Oxtr* expression (King, 2016), with the striatum possessing the highest receptor levels, the insular cortex showing relatively static *Oxtr* expression, and the inferior colliculus displaying near zero expression. Nuclei were isolated from freshly dissected brain tissue and quantified into 100,000 aliquot samples for chromatin extraction and library prep. Sequenced DNA libraries ranged from 2-6ng of pure nucleic acids which were quantified via Kapa qPCR. In preliminary experiments, we performed ATAC-Seq under different

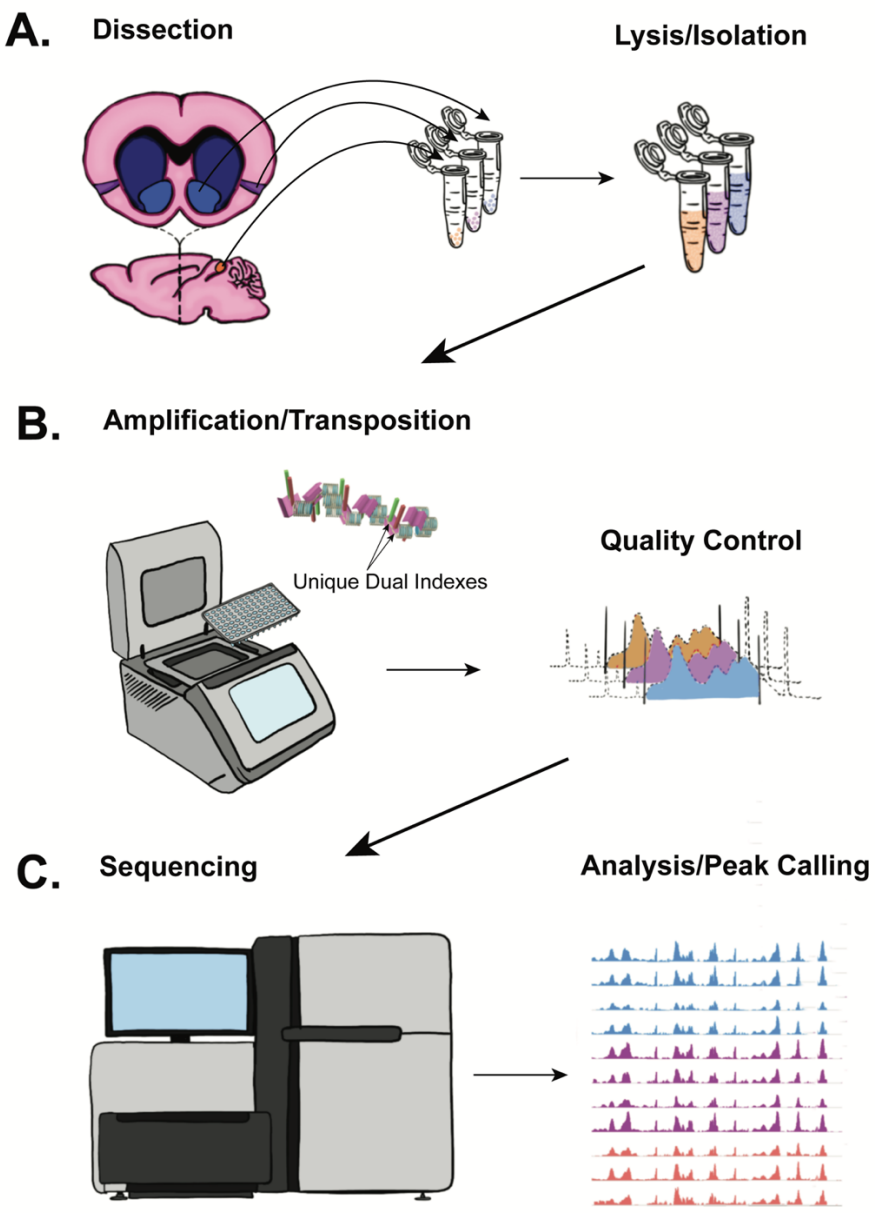


Figure 2.1. Prairie Vole ATAC-Seq Workflow.
(A) Freshly dissected brain tissue from striatum (blue) insular cortex (purple) and inferior colliculus (orange) are used to create samples that are each a pool from three animals. Cells are lysed and nuclei are isolated and quantified with a hematocytometer. (B) Lysates serve as template DNA for PCR amplification, which simultaneously involves tagmentation and fragmentation of chromatin with the Tn5 enzyme. Unique index primers are ligated to individually label both strands of DNA. Quality control analysis is performed to determine specific fragment size distribution and nucleic acid concentration. (C) Whole genome sequencing of validated samples is conducted, and sequencing reads are aligned to the genome. Further analysis allows for the inference of chromatin accessibility through defining peak regions at areas of high coverage.

storage conditions and protocol lengths. We found that flash frozen tissue produced roughly 50% of DNA library yields compared to freshly dissected brain tissue. We also compared preparations of 50,000 nuclei to 100,000 but found high concentrations of large fragments (500-2000bp) and thus inconsistent target fragment lengths for DNA libraries generated from 50k. A Phred quality score, q>30, is commonly used to assess the quality of sequencing reads as a function of errors per 1000 sequenced bases. The 11 final libraries used for this study all resulted q>30 over 90%, which can provide high confidence in reliable data to be carried into further analysis. From nuclei isolation and quantification to DNA sequencing, it is possible to adapt this workflow to examine comparisons between many other regions of the brain or specific cell types (Figure 1).

Coverage, Validation, and Peak Calling

Our sequencing approach allowed us to obtain likely the most deeply sequenced whole genome dataset for *M. Ochrogaster*, with over 2.1 billion total raw sequencing reads (Figure 2a), and up to 10x coverage across the genome for 11 samples. Two distinct files were produced for each uniquely indexed sample, which contained sequencing reads that represent information from each side of the double stranded DNA libraries, also known as paired end data. Generally, paired-end reads present reduced error call rates and improved overall alignments to reference genomes compared to single-end reads (Corley et al., 2017; Maher et al., 2009). To validate the chromatin accessibility signal, a major expectation is the ability to detect open regions of DNA in the size ranges that are expected. For some methods of analysis, researchers may introduce bias during peak calling by allowing smaller aligned reads to be included. To mitigate this type of bias, we constrained the peak calling algorithm's extension size parameter to 150, which is also roughly the base pair amount of DNA that wraps around a single nucleosome. Nucleosome scaled fragment

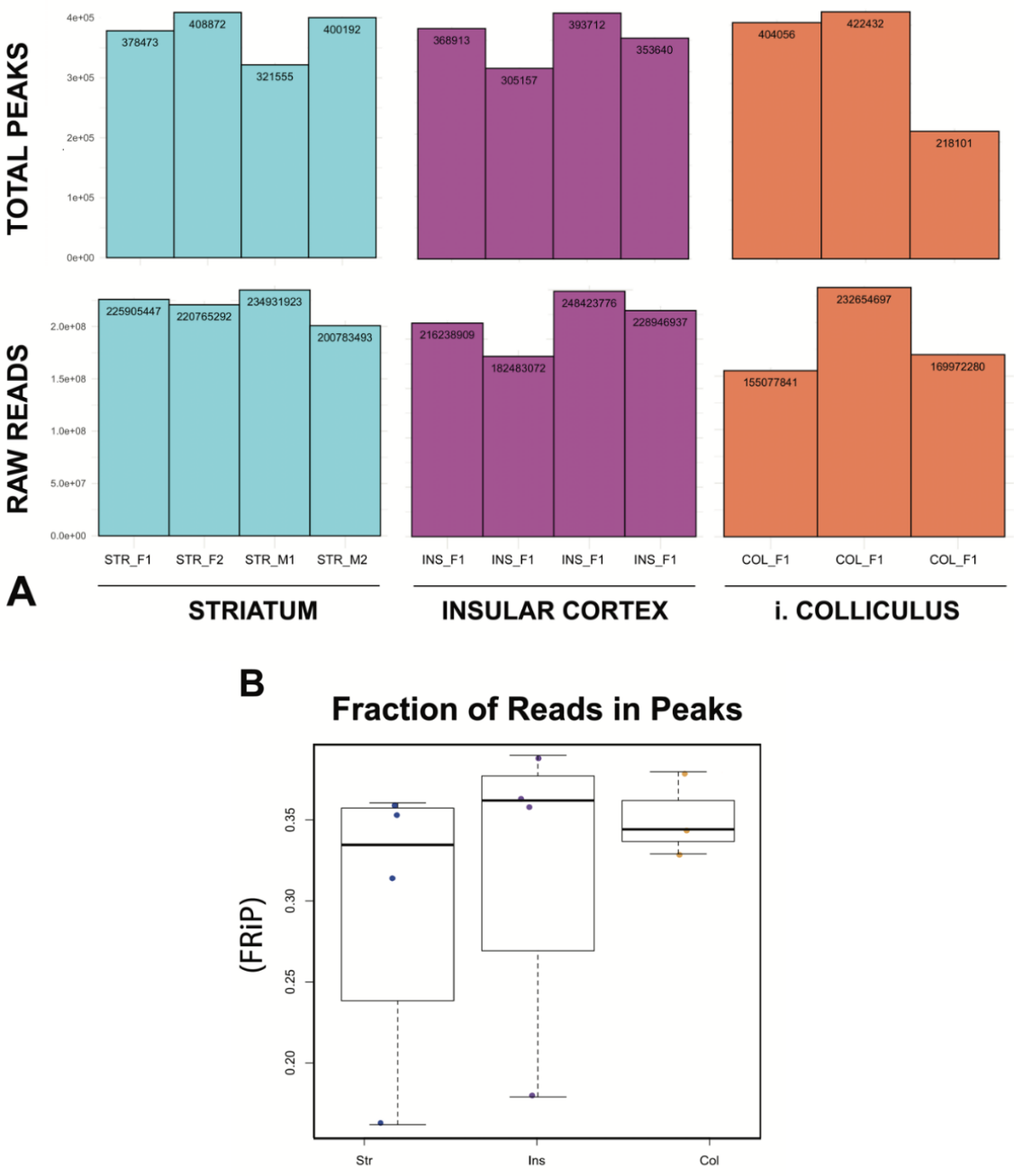


Figure 2.2. Sequencing Statistics and Peak Calling.
(A) Lane statistics for raw reads obtained from whole genome sequencing of 11 prairie vole brain ATAC-Seq libraries using the NovaSeq DNA sequencing platform. Approximately 2.5 billion total paired end reads were generated with a coverage range of 150 million to 250 million reads per sample that were aligned to the prairie vole reference genome. Peak calling with MACS2 identified 218,000-422,000 peaks ($p<0.01$, insert size 150 pairs). (B) Bar plot representing the fraction of genomic reads within peaks (FRiP) for each sample and grouped by brain area. Here, an average range of approximately 35% of reads can be found within identified peaks between the three brain tissue types.

lengths ultimately increase resolution for peak detection and reproducibility of peaks across samples. Here, we detected 218,101 to 422,432 peaks between 11 samples, with no specific influence of read depth on peak detection (Figure 2.2a – top row). On the contrary, samples with lower coverage appear to be slightly more likely to detect peaks due to a higher ratio of aligned reads to total reads. ATAC-Seq data from inferior colliculus tissue presented the highest and lowest number of detected peaks between all 11 samples. This is potentially due to higher individual variation among brain dissections performed for this study. The striatum and insular samples were taken routinely from precise brain slices, whereas the inferior colliculi were completely dissected and then cut into non-specific fractions prior to being weighed. For this study, the inferior colliculus generally served as a negative control. Another important metric for interpreting peak quantity with ATAC-Seq is the fraction of DNA reads that are found within peaks (FRiP) compared to all sequenced reads. The ENCODE data consortium recommends that ATAC-Seq peaks be validated at a minimum FRiP score of 0.2 (https://www.encodeproject.org/atac-seq/). The range of FRiP scores for this dataset was 0.30-0.40 (Figure 2.2b), and within typical range of those reported in recent ATAC-Seq publications (Jauregui-Lozano et al., 2021; Peng et al., 2021; Rocks et al., 2022).

Housekeeping Regions for Prairie Vole Brain Tissue

After validating the alignment and peak detection, we were interested in representing the overall coverage of DNA reads throughout the prairie vole genome because it is not well documented. This can be partially explained by the fact the MichOch1.0 assembly is heavily fragmented and exists as contiguous chromosomal sequences (contigs) that span in length from

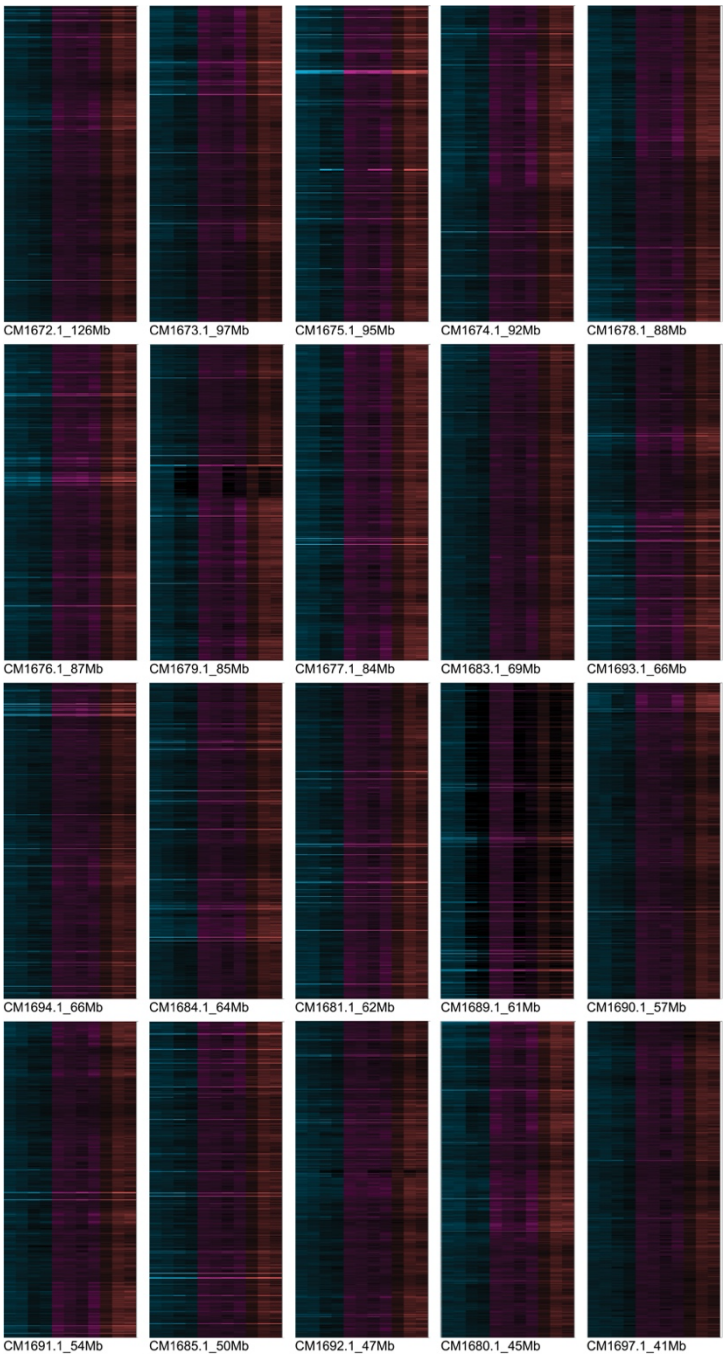


Figure 2.3. Chromosomal Heatmap of ATAC-Seq Coverage.
Heatmap representing alignment of ATAC reads to the 20 largest and most complete chromosomal segments of the prairie vole reference genome (MicOch1.0). Brighter regions indicate higher quantities of aligned reads, and shaded areas indicate low alignment. All 11 final samples (4 striatum - blue, 4 insular cortex-purple, 3 inferior colliculus-orange) are represented from left to right and genomic coordinates are sorted from top to bottom. The contigs above represent approximately 50% of the total MicOch1.0 assembly and range from 126Mb (CM1672.1) to 41Mb (CM1697.1).

1,000 bp to more complete chromosomes around 100Mb. We subset the 20 largest contigs to produce an alignment heatmap that includes approximately 90% of the genome, where brighter colors indicate higher frequency of aligned reads and darker areas represent lower frequency of alignment (Figure 2.3). The heatmap demonstrates that a majority of the genome is covered by ATAC-Seq, and that very specific positions of high coverage were generally shared between all samples. To provide confirmatory evidence for performing ATAC-Seq of brain tissues, we next wanted to identify consistent peaks within genes known to be expressed in the brain. Uniform coverage at cytoskeletal and DNA methylation genes such as *Gapdh* (GeneID:ENSMOCT00000009886) and *Dmn3* across all 11 final samples (STR-striatum, INS-insular cortex, COL-inferior colliculus) serves as an indicator for consistency of the ATAC-seq signal, as well as a resolution that can distinguish between promoter, coding, and intron regions of genes (Figure 2.4). Uniform coverage at genes such as *Astn1* and *Rgs8*, which respectively play a role in neuronal adhesion and regulation of g-protein signaling, provide some confirmatory evidence for the isolation of intact nuclei derived from neural tissues.

Sex Explains Most Variation in Brain Chromatin Accessibility

To examine the extent of individual variation in the prairie vole ATAC-Seq signal, we generated a matrix which resulted in a subset of approximately 65,000 shared peaks between every sample. This information was input into the Deseq2 analysis pipeline, with the 4 striatum samples loaded as the treatment condition and insula and colliculus samples as control. For these shared positions of DNA accessibility, sex resulted as the predominant factor that explains the most variation in the prairie vole ATAC-Seq peaks. A principal component analysis (PCA) plot from

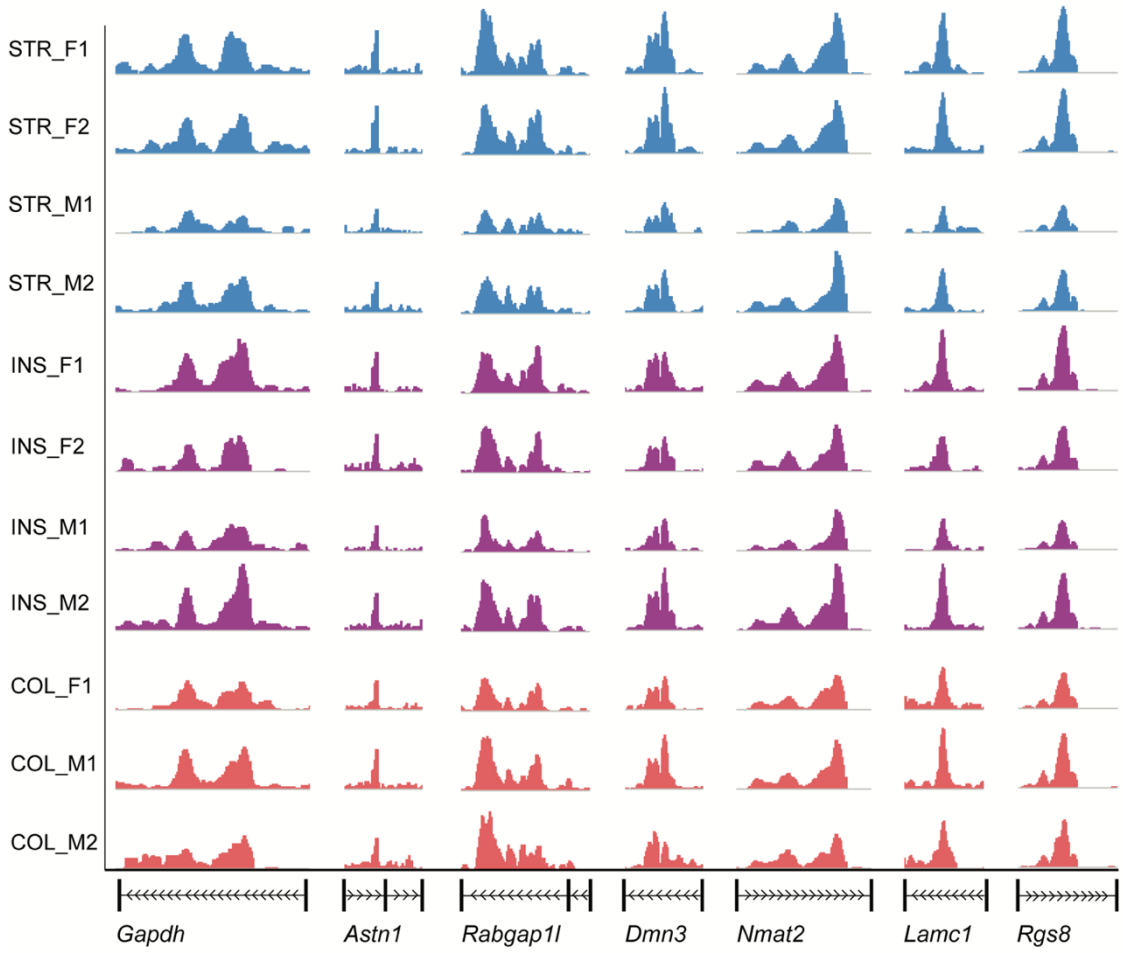


Figure 2.4. Prairie Vole ATAC Housekeeping Regions.
Integrative Genomics Viewer (IGV) display of the distribution of aligned ATAC-Seq reads to the prairie vole genome (MicOch1.0) at representative housekeeping regions. Reads mapped across all 11 samples (STR, INS, COL) at cytoskeletal (Gapdh), DNA methylation (Dmn3), neuronal adhesion (Astn1), and g-protein signaling (RGS8) indicate consistency of ATAC-seq signal with prairie vole brain tissue.

the DESeq2 program demonstrates that principal component 1 (Figure 2.6, x-axis) explains up to 69% of individual differences in magnitude of any given shared peak. From the position of samples in the plot, it can be inferred that males and females cluster together on opposite ends of this axis. Sex differences in gene regulation have been previously reported for brain tissues, and specifically striatum-based tissues (Hodes et al., 2015; Zou et al., 2020). Nevertheless, the most significant individual differences between males and females that we identified in prairie voles were found in noncoding, and often non-annotated regions of the prairie vole genome. Sex-bias in chromatin topology may therefore give rise to broad differences we observe at shared accessible sites.

Due to the strong influence of sex on variation in peak magnitude, we also wanted to search for prominent sex differences in chromatin accessibility outside of shared peaks. A qualitative assessment of the ATAC-signal considering sex yielded several positions that were within promoter and intron regions of known genes (Figure 2.7). Interestingly, the genes we identified with apparent male/female bias in accessibility also have known nervous system functions. The *Scaper* gene, for example, is much more widely accessible in females than in males. Deficiencies in the *Scaper* gene are associated with intellectual disability and retinitis pigmentosa, which interestingly are heavily male-biased. At *Slc15a5* and *Gria3*, the opposite can be observed, with male specific peaks at loci where females have little to no accessibility. We next quantified peaks per chromosome for all samples and observed strong similarities in the number of peaks called for autosomes. The X chromosome, however, results in an over 50% reduction in peaks for male samples compared to female samples across each of the three brain areas (Figure 2.5b). This is unexpected because of X chromosome inactivation in females and could potentially be a sign of dosage compensation or escape from X chromosome inactivation. Conversely, because the female

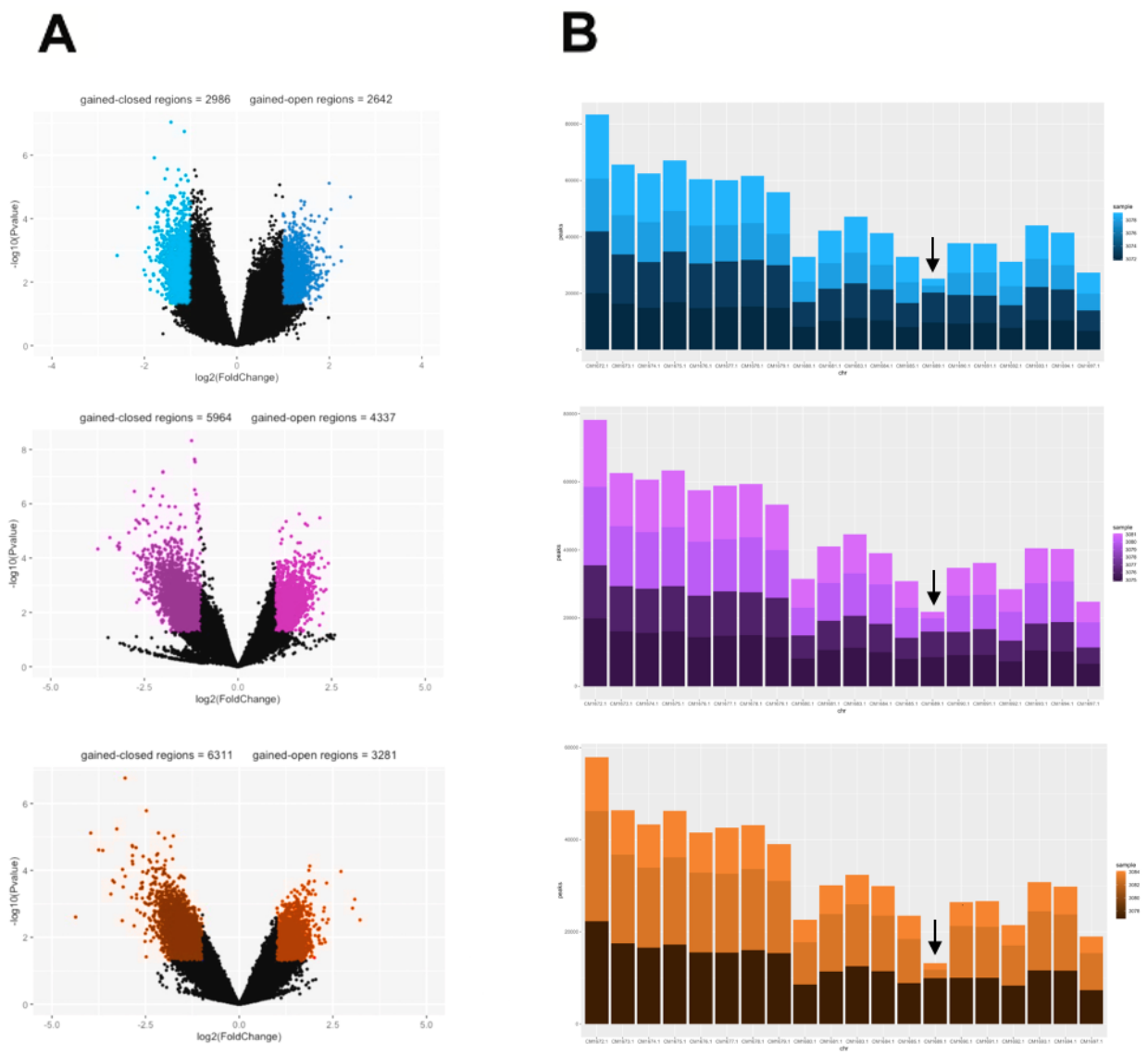


Figure 2.5. Differential Enrichment and Peak Calling per Chromosome.
(A) DeSeq2 volcano plot view of tests for brain-region specific differentially enriched peaks (STR-top, INS-middle, COL-bottom). Individual fold changes in accessibility are graphed against p value, where higher variation is represented in the top left and right corners and defined as gained open or closed. 5628 differentially accessible regions are identified in striatum, 10,301 for Insula, and 9,592 for colliculus. (B) Total peaks identified for each brain area for the 20 largest chromosomal regions (STR-top, INS-middle, COL-bottom). Bars are segmented to represent peak quantity for individual samples, with female samples being darker shaded, and male samples lighter. On average, 10,000-20,000 peaks are called for most large chromosomes. The arrows indicate the peak quantities for the X chromosome, where a high reduction in peak detection of >50% is present in male samples compared to female.

X chromosome peak quantities are more in line with the average peaks detected in autosomes, it could be more likely that male samples have reduced DNA accessibility on the X chromosome. ATAC-Seq experiments have confirmed the existence of higher chromatin accessibility in the female active X chromosome compared to inactive X (Balaton et al., 2021). However, we did not identify any reporting of sex differences in DNA accessibility between male and female X chromosomes. One recent study has shown that sexual differentiation in the mammalian brain is driven strongly by sex hormones and their influence on large gene expression networks (Gegenhuber et al., 2022). Considering this, future analysis with this data could allow us to see if peaks that are female specific contain binding sequences for transcription factors such as gonadal hormone receptors and glucocorticoids. Our initial examination of ATAC-Seq signal in the prairie vole brain elucidates features of substantial sex differences in overall chromatin accessibility a clear reduction in peaks on the male X chromosome, and in the magnitude of accessibility at shared peaks. These findings provide a strong foundational for considering sex an important variable in chromatin studies of the brain.

Differentially Enriched Genes in Striatum are Associated with DNA binding

Due to the dynamic expression patterns for *Oxtr* that have been previously observed in the prairie vole striatum, a primary interest for this study was to generate a list of the most differentially accessible sites in the striatum compared to the insular cortex and inferior colliculus. To obtain this information from our data, the peak magnitudes from male and female striatal samples were compared as the treatment condition against insular cortex and inferior colliculus samples as control conditions with Deseq2 analysis. We also performed this analysis independently with insular and colliculus samples each as treatment conditions and the other two brain areas as

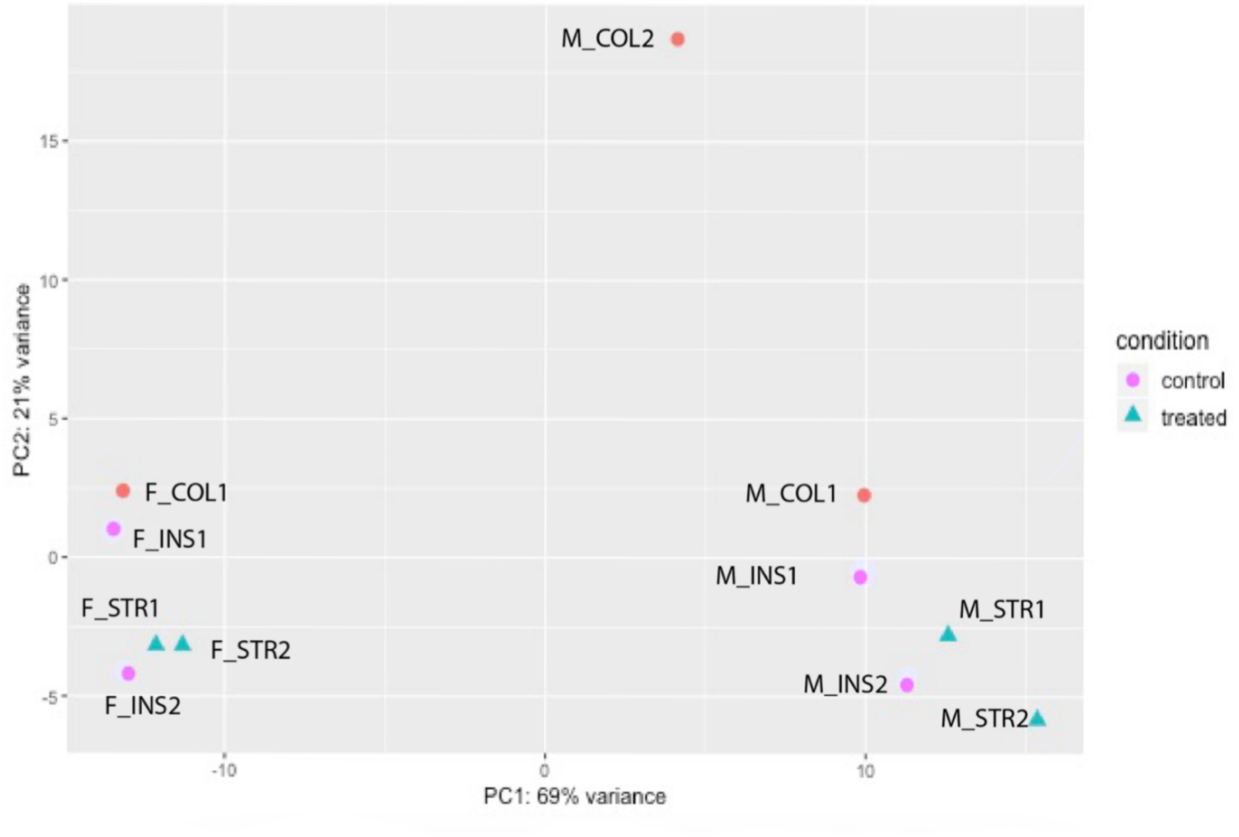


Figure 2.6. Sex Explains Most Variation in Differentially Enriched Regions.
Principle Component Analysis. DeSeq2 plot view of principle component analysis from 11 ATAC-Seq libraries representing 3 brain areas (STR=Striatum, INS=insular cortex, COL=inferior colliculus). All Striatum samples were loaded as treatment condition, and all insula and inferior colliculus samples were labeled as control. This plot shows that an assessment of variation in peak enrichment across all samples produces clustering based predominantly on sex, and that the first principal component (PC1) is a factor explaining most variation in the data (69%).

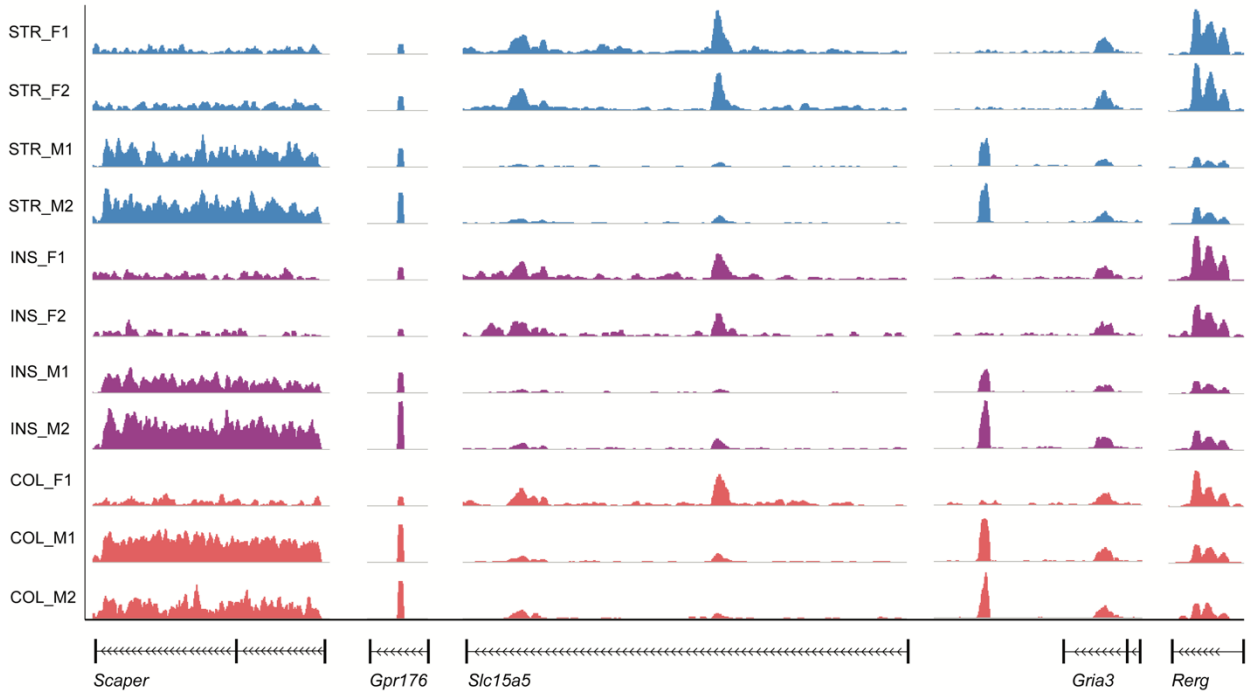


Figure 2.7. Sex Differences in Chromatin Accessibility.
Sex differences in ATAC-Seq signal identified in differentially enriched regions of known MicOch1.0 genes. Genes such as *Rerg*, *Scaper* and *Slc15a5* show variation in accessibility at multiple positions along their respective gene bodies in autosomes. In the Gria3 gene, found on the X chromosome (CM 1689:6,622,450-6,230,000), a male specific peak spanning the approximate width of two nucleosomes (450-500 bp) is likely within the 3' untranslated region. Gria3 encodes the glutamate receptor subunit 3.

controls (Figure 2.5a). This allowed us to capture brain region specific characteristics of the ATAC-Seq signal. The analysis identified approximately 5,000 differentially accessible peaks in striatum samples, and around 10,000 differentially accessible peaks for insula and colliculus samples. For the striatum, differentially enriched regions were then sorted based on p-value and the top 20 most significant genomic loci were extracted. Differentially enriched regions from striatal tissue contained modest fold enrichments of 1.5-2.0, and p-values from 0.006-1.23e^{-6}. The tests for differential accessibility were performed for all shared peaks across the genome, most of which were not within known genes. Most of the top 20 regions that were identified as either 'gained-open' or 'gained-closed' did include genes, with several encoding transcription factors such as *Arhgap5*, *Cux1*, *Zhx3*, and *Ctcf* (Figure 2.8). Although the chromatin state at these genes may be important in regulating rates of their own transcription, we hypothesize that the specific binding sites of these factors elsewhere in the genome are a more important focus for establishing any causative role in tissue specific regulation in the striatum. CTCF, for example, is a known insulator and transcription factor at millions of locations in the genome. In human cells, CTCF sites identified by ChIP-Seq have a high rate of detection within ATAC-Seq peaks (Karabacak Calviello et al., 2019). This work described here may be complemented in the future by analysis that can overlay the presence of factors such as CTCF in the prairie vole brain with regions of DNA accessibility.

DISCUSSION

The results from the present study confirm that the ATAC-Seq method can be used to examine genome-wide chromatin accessibility in the prairie vole brain. We provide evidence that most

name	Chr	Start	End	baseMean	log2FoldChange	pvalue	regulation	gene	chromosome
Peak_45855	CM001673.1	21196923	21197833	68.7979323	-1.780923	1.23E-06	gained-close	Cux1	2
Peak_456401	CM001697.1	480054	480466	31.9955019	2.00165537	7.96E-06	gained-open	Olfr1152	LG9
Peak_197250	CM001678.1	73548756	73549292	37.4180277	-1.9312953	1.55E-05	gained-close	ArMh3	8
Peak_36387	CM001672.1	117939140	117939450	18.8320404	2.46191712	2.15E-05	gained-open	noncoding	1
Peak_334628	CM001687.1	4828486	4829099	45.6991197	-2.1370848	4.51E-05	gained-close	Usp15	24
Peak_146864	CM001676.1	81003016	81003725	46.0042159	-1.7290815	4.69E-05	gained-close	intergenic: Cd34/Cd46	6
Peak_299467	CM001684.1	35127743	35128486	31.9928838	2.03782349	5.27E-05	gained-open	noncoding	19
Peak_20255	CM001672.1	60695331	60695840	39.2430407	-1.7123469	6.94E-05	gained-close	Arhgap5	1
Peak_187966	CM001678.1	41673751	41674069	37.8370766	-1.734471	7.22E-05	gained-close	Dnajc4	8
Peak_307250	CM001684.1	61822627	61823511	59.7126158	1.93993538	0.00010564	gained-open	noncoding	19
Peak_262966	CM001682.1	14475338	14476098	44.1942827	-1.7216292	0.00013454	gained-close	Pinx1	17
Peak_435844	CM001694.1	53087807	53088157	21.5892265	1.89606165	0.00013912	gained-open	noncoding	LG5
Peak_453466	CM001696.1	17546807	17547260	24.8158897	-2.5876335	0.00014598	gained-close	Zhx3	LG8
Peak_69207	CM001674.1	5266109	5266564	44.9250336	-1.7635596	0.00018545	gained-close	Ctcf	4
Peak_88701	CM001674.1	71348078	71348442	35.9234043	-1.8774996	0.0001977	gained-close	intergenic: Wdsub1/Baz2b	4
Peak_46544	CM001673.1	23174797	23175209	40.9994992	-1.7530634	0.00025376	gained-close	noncoding	2
Peak_224092	CM001679.1	81276086	81276512	38.2803373	-1.7810671	0.00026166	gained-close	Camta1	10
Peak_30821	CM001672.1	98684844	98685320	38.5498171	-1.7293256	0.00042683	gained-close	noncoding	1
Peak_40702	CM001673.1	5486950	5487410	57.5565141	-1.7332612	0.00047	gained-close	Pxn	2
Peak_536145	JH996436.1	11980017	11980440	36.4275571	-1.7208936	0.0006738	gained-close	Thoc5	NA

Figure 2.8. Top 20 differentially enriched genes from the striatum
Results from an assessment for striatum specific differentially accessible ATAC-Seq peaks. Results are sorted by p-value ($6e^{-3}$-$1.23e^{-6}$) and contain information for chromosome name, position, and the inferred gene and chromosome number as cross-referenced with the Ensemble genomics database (https://www.ensembl.org/Microtus_ochrogaster). Regulatory status is also included as "gained-close" and "gained-open" to indicate the direction of differential enrichment. Many genes in this list function in DNA binding and transcription regulation, including *Cux1*, *Zhx3*, *ArhGap5*, and *CTCF*.

variation in peak enrichment is due to sex, but also that most shared peaks are comparably equal in magnitude. Our aligned sequencing data demonstrates a signal-to-noise ratio that we believe closely elucidates variation in chromatin accessibility given the nuclei quantities that are commonly used. We show that when generating 150-250 Million reads per sample, coverage depths above peak threshold can range from magnitudes of 10^2-10^4. In addition, our analysis reveals that 10-20% of total peaks are differentially enriched across three brain areas, with most peaks found in non-coding regions of the genome. It is not currently known what role these sites may play as regulatory elements, but there is potential for further investigation that can be performed informatically to identify targets for DNA-protein interactions. Interestingly, the top genes for differentially enriched peaks of the striatum include transcription factors that have a known function in neural gene regulation. Characterizing the binding sites for CTCF or Cux1 within ATAC-Seq peaks we have identified could help to inform if the presence of these transcription factors overlays with differentially regulated regions of the genome specific to the striatum.

The MicOch 1.0 genome assembly presents some drawbacks regarding the mapping of the first gene list files for the species within the relevant software packages for analysis. Some, but not all genes can be confidently identified in the MicOch 1.0 assembly when integrated into workflows that are standard in mice and humans, such as IGV and DeSeq2. The Ensemble database, which has integrated karyotyping and gene predictions, is a valuable resource for confirming coding sequences and individual loci when working with the MichOch1.0 assembly. A newly annotated genome in prairie voles could justify a re-evaluation of the chromatin data generated from this study. Quantification of peaks near introns, transcription factor start sites and promoters of genes, for example, would be of interest.

Comparisons between brain chromatin data can be difficult due to variable experimental approaches, including sample preservation, depth of sequencing, and significance thresholds for peak calling. Biologically, the organization of chromatin is also dependent on phylogenetic and cell-lineages. In a study that compared cow, pig, and mouse samples, it was estimated that as few as 20% of ATAC-Seq peaks are shared between species, which is correlated with evolutionary distance (Halstead et al., 2020). In addition, within organism comparisons of 16 different tissue types demonstrated up to 70,000 tissue specific peaks exist. The translatability of inferring individual open chromatin regions from one sample to the next may therefore be heavily based on species and tissue type. Understanding tissue specific differences can help to uncover the changes in DNA organization that are related to unique regulatory mechanisms. Interestingly, the same study examined two different types of brain tissues and the found the most similarities of any two tissue types were between cortex and hypothalamus. This finding is also supported by our observation of strikingly similar ATAC-Seq signals between the striatum, insular cortex, and inferior colliculus, where over 30% of total peaks were shared between all three areas.

Taken together, our results and others suggest that neuronal populations throughout the brain may contain a more generalized chromatin conformation than previously assumed. The ability of the unique cell types of the brain to produce numerous gene expression profiles is perhaps highly dependent on the binding activity of specific tissue-specific transcription factors and epigenetic modifications. In addition, many of these regulatory processes may be engaged in response to specific environmental stimuli, which is a topic that was not included in the scope of this work but may be integral to understanding the true extent of variability in the chromatin landscape of behaviorally relevant genes. Indeed, a growing body of evidence has gained support for the role of histone modifications and DNA methylation as markers for gene x environment

interactions and assessing risk for neuropsychiatric conditions (Baker et al., 2017; Kraaijenvanger et al., 2019; Smearman et al., 2016).

Prairie voles are a unique model organism for studying behaviorally relevant genes due to their socially monogamous behavior in the wild that has been transferred to the lab as well-established behavioral paradigms. The neuroanatomical pathways related to social information processing are also well-documented in prairie voles, with a major focus on the neuropeptides OXT and AVP. Genomic resources in the prairie vole model have not been developed at the same pace. Our research here aims to advance knowledge within the prairie vole research model by investigating chromatin accessibility genome-wide, and in samples of brain tissue taken from both sexes. As demonstrated by our discovery of sex differences in several genes that are previously unreported in the species, we demonstrate the value of chromatin-based assays for understanding functional aspects of DNA regulation in the brain. The successful development and implementation of an ATAC-Seq workflow in the prairie vole brain is a critical step towards functional annotation of noncoding regulatory motifs. In the future, proper annotation of regulatory elements on a specifies specific level will contribute to a more in depth understanding of the role between genetic variation and phenotypic complexity between and within species.

Our results bring much attention to sex as a biological variable when investigating chromatin accessibility in the brain. The sexually dimorphic profiles observed from our study may have a substantial impact on our understanding of prairie vole gene regulation, particularly the reduction in peaks identified on the male X chromosome compared to females. There are two possibilities when considering the larger implications to other mammals and the human brain. 1 – that sex differences in chromatin accessibility are prominent across multiple species, and underlie many aspects of sex-biased gene expression and neural functions. 2 - characteristics of broad sex

differences at autosomes and sex chromosomes are unique to prairie voles. Evidence can be found for both cases, as sexually dimorphic gene expression has been localized to several areas of the brain, including the nucleus accumbens (Gegenhuber and Tollkuhn, 2020; Hodes et al., 2015). On the contrary, no other studies investigating chromatin accessibility have reported similar results to our observations, where a significant reduction in X chromosome peak quantities exist in males compared to females. Currently, it can be said that this phenomenon is unique to the prairie vole, although further work in voles and perhaps other species will be necessary to confirm the extent of variation in chromatin accessibility that is explained by sex.

Regarding sexual differentiation in the brain, evidence can be found that gene regulation is mediated by the activity of gonadal steroids, glucocorticoids, and more internally programmed by sex-linked genes (Gegenhuber et al., 2022; Gray et al., 2017; Manoli and Tollkuhn, 2018; Zou et al., 2020). To some extent, it must be considered that each of these biological systems operate in combination to drive phenotypic differences between sexes, as well as similar phenotypes via convergent molecular mechanisms. Many of the sex differences seen at the level of chromatin likely infer functional differences in gene expression programs within neuronal subpopulations which ultimately influence brain connectivity and neural activity. Historically, the inclusion of both sexes in neuroscience research has been underdeveloped for human and animal studies. What remains to be discovered are the unique interactions that occur at accessible regions of chromatin to directly impact the onset of sexually dimorphic phenotypes. In a broader context, monogamous species tend to physically be less sexually dimorphic, largely due to decreased competition for mating and reproduction. Where male and female voles show much more similar forms of parental behavior than other rodents, the mechanisms by which these behaviors are achieved by the brain may be entirely different. We postulate that functional differences in the regulation of the same

genes may therefore be necessary to produce the same behavior. This is supported in our data by a strong overlap in the detection of peaks between males and females, but substantial variation in peak magnitudes across the genome. Alternatively, we do show positions of the genome that have a stark contrast in accessibility between males and females, which suggests some sex differences in gene regulation may be more hardwired in biology of the sexes than tuned for environmental adaptations.

Major considerations for employing the use of ATAC-Seq in brain tissue should be variation in cell viability, nuclei quantity, and cell-type heterogeneity. Although we were able to control for nuclei quantity with the use of a traditional hematocytometer, the variation in accessibility due to cell type was not captured between the three brain areas studied here and could be one confound for the interpretation of region specificity. We do demonstrate a feasibility of generating consistent ATAC-Seq signals across pooled samples taken from freshly dissected brain tissue. Some studies have shown success in ATAC-Seq peak detection from frozen cell and tissue samples (Corces et al., 2017; Peng et al., 2021). In our preliminary experiments, we identified a significant reduction in quantities of viable nuclei from snap frozen tissue compared to freshly dissected. It is worth noting that many studies using frozen samples are prepared *in vitro*. Future work could also seek to compare chromatin landscapes from the tissues mentioned in a more cell-type specific manner, such as distinctions between neurons or glia. Although peaks of expected total quantities were detected from our experiments, many genes with hypothesized expression in the brain displayed relative low coverage compared to regions containing peaks. It is not clear from this data how open or closed regions may specifically be contributing to gene expression. Understanding the accessibility of DNA at low vs. high expressed genes could therefore be one of the strongest ways to compliment this dataset in future research. One possibility for explaining

low coverage at genes believed to be expressed in neural tissue may be due to occupation by regulatory proteins that could be, in fact, increasing gene transcription. Conversely, it may be possible that other factors at occupied regions of DNA are decreasing transcription. In addition to measuring gene expression, identifying transcription factor binding sites at known ATAC-seq peaks may help to identify further mechanisms of tissue-specific gene regulation.

The combination of more advanced cell and molecular techniques such as single cell sorting, RNA-Seq, ChIP-Seq, and further immunohistochemistry classification of dissected brain regions could strongly add a biological context to the measurement of nucleosome free DNA presented by ATAC-Seq. Nevertheless, we provide the first protocol for performing a genome-wide assessment of the ATAC method in prairie voles, and we hope our work can support further research that explores the relationship between genetic variation and brain phenotypes using this model.

Chapter 3:

Oxtr Chromatin Accessibility in the Prairie Vole Brain

ABSTRACT

The oxytocin (OXT) receptor gene (*OXTR*/*Oxtr*) is often reported in research that investigates the connections between DNA variation, brain phenotypes and social outcomes. Prairie voles (*Microtus Ochrogaster*) are a socially monogamous rodent species that has natural variation of *Oxtr* expression in reward processing areas of the brain. Single nucleotide polymorphisms (SNPs) identified in prairie voles have been previously associated with individual variation in *Oxtr* expression. At the molecular level, much less is known about how regulation of *OXTR*/*Oxtr* is mediated in different areas of the brain, and specifically which DNA variants in the gene have functional roles. The Assay for Transposase Accessible Chromatin and DNA Sequencing (ATAC-Seq) has recently become a popular technique for measuring DNA accessibility throughout the genome. ATAC-Seq may provide useful information towards understanding the molecular mechanisms that mediate regulated transcription of genes such as *Oxtr*. Here, we analyzed male and female ATAC-Seq data from three prairie vole brain regions: inferior colliculus, insular cortex, and striatum. Previous studies have shown that these tissues express *Oxtr* at respectively low, invariable, and highly variable levels, respectively. Our results indicate that the chromosomal region that houses *Oxtr* in prairie voles presents high degrees of individual variability, and no consistent peak detection at the promoter or intron. Variant calling at *Oxtr* identified 274 previously undefined polymorphisms in *M. Ochrogaster*, and a wide range of allele frequencies across the gene body. From this data, we report two distinctly unique groups of SNPs with the highest differential accessibility between the striatum and insular cortex. Minimal overlap in the positions of high frequency SNPs is suggestive that separate alleles can vary in overall accessibility across different brain tissues. These findings provide support for the complex dynamics between *Oxtr* variation and DNA accessibility in the prairie vole brain.

INTRODUCTION

Social deficits are often shared characteristics seen in psychiatric disorders such as autism, schizophrenia, and bipolar disorder. The overall formation of social cognition and behavior is highly influenced by the environment, but years of population-based genetics research has also helped to demonstrate strong heritable components to some of the most critical personality traits, including trust, empathy, and aggression (Cesarini et al., 2008; Porsch et al., 2016; Warrier et al., 2018). For many researchers, a critical barrier is still knowing which genes and types of molecular phenotypes to examine. Identifying the specific genetic risk factors for behavioral traits that are characterized in psychiatric conditions is therefore important for the future of understanding and promoting mental health.

Single nucleotide polymorphisms (SNPs) are single base pair DNA variants that are shared within a population. SNPs are among the most widely studied observations for Genome-wide association studies (GWAS) that examine individual genetic risk for specific diseases or traits. Much of this work has specifically been applied to the field of psychiatry, and as a result, thousands of potential SNPs have been statistically correlated with various forms of social behavior and cognition (Visscher et al., 2012; Yang et al., 2012). From these results, a small fraction of significant SNPs are assumed to each play minor additive roles in driving phenotypic differences in the brain. One of the more frequently reported genes from behavioral association studies has been the oxytocin receptor gene (OXTR) with over 150 unique polymorphisms that have been reported in humans (Baribeau et al., 2017). Research that can establish functionality between behaviorally relevant SNPs such as those at OXTR, and their influence on gene expression, may provide unique insights that underlie a molecular basis for the relationship between DNA variation and behavior.

Oxytocin (OXT) is an evolutionarily conserved neuropeptide that functions to mediate the reproductive physiology and social behavior of many species. Historically, OXT was discovered to regulate uterine contraction and milk ejection in mammals (Nickerson et al., 1954). Decades of pharmacological, histological, and behavioral experiments have also tied OXT function to neural mechanisms of socially reinforced behavior (Olazabal and Young, 2006; Ross et al., 2009; Skuse et al., 2014). One of the more valuable research models for understanding mechanisms of OXT in the brain is the socially monogamous prairie vole. In voles, it has been demonstrated that the neurons of the posterior pituitary, which function to release OXT into the periphery, also have collateral fibers projecting to areas of the striatum (Ross et al., 2009), which is involved in reward processing and is critical for pair bond formation (Aragona et al., 2003). Targeted DNA sequencing research combined with tests that measure affiliative behavior in prairie voles have found associations between *Oxtr* SNPs and increased pair bonding (King et al., 2016). Further, *Oxtr* SNPs in prairie voles have been linked to individual variation in gene expression in the nucleus accumbens (NAcc), a sub-region of the striatum. A follow up goal to this work has been to provide information that can help to infer on the complex regulatory processes that lead to differences in *Oxtr* expression across different brain regions. Newly developed methods in chromatin-based research may now provide an opportunity to explore physical characteristics of the prairie vole genome within the context of the brain and behavior.

The Assay for Transposase Accessible Chromatin (ATAC-Seq) is a method that allows for genome-wide detection of open chromatin (Buenrostro et al., 2013; Buenrostro et al., 2015). To date, no studies have been reported that have performed ATAC-Seq in prairie vole. In Chapter 2, we first validated the ATAC-Seq methodology in prairie vole brain tissue and performed a comprehensive analysis of open chromatin peaks and examined differential accessibility between

different areas of the brain. In the present study, we address a primary purpose for generating this ATAC-Seq dataset, and further analyze DNA accessibility of the striatum (STR), insular cortex (INS), and inferior colliculus (COL) specifically at the chromosomal regions spanning *Oxtr*. Additionally, we have applied a new framework for variant calling with ATAC-Seq data, and have defined novel polymorphisms at *Oxtr* as well as the respective allele frequencies for the three brain areas studied.

METHODS

Prairie Vole ATAC-Seq Data and Visualization

The data analyzed for this study was collected from the ATAC-Seq experiments as previously described in Chapter 2. Briefly, representing 11 freshly dissected brain samples that were transposed and amplified for library preparation same day. 5 Female (2 STR, 2 INS, 1 COL) and 6 Male (2 STR, 2 INS, 2 COL) were used for analyzing *Oxtr* and performing variant calling. Each sample represents a pooled triplicate of same-sex littermates. Paired-end DNA sequencing reads ranged from 150M to 250M total reads per sample, with Q>30 ratios between 0.88 and 0.92. Visualizations were produced from quality filtered reads that are normalized by an average of transcripts per million (TPM) bases for all samples. .bam files containing normalized sequencing reads were converted into a compressed .tdf format with the Integrative Genomics Viewer IGV tools and visualized with IGV. Gene track diagrams and Ven diagrams were created using the base R statistical programming software and Gviz package. Custom annotations for *Oxtr* SNPs were overlayed using SnapGene and Adobe Illustrator.

OXTR Variant Calling

All steps during variant calling were performed within the UNIX command line. Using the samtools (https://github.com/samtools) -view command, sequencing reads within a region of interest including the entire *Oxtr* (JH996431.1: 26350000-26352000) were extracted from the genomic alignment files for each of 11 ATAC-Seq samples. Alignment files were grouped based on brain tissue to maximize coverage depth for SNP detection, and input into the bcftools program (https://github.com/samtools/bcftools) -mpileup command as three separate files (STR, INS, COL) with an average of 200K reads per sample for the *Oxtr* locus. The MichOch1.0 reference genome was indexed and placed within the working directory of the DNA reads. Merged reads were output directly into a pipeline for the bcftools -call command for variant calling and calculating allele frequencies. Data containing single nucleotide polymorphisms were exported into three .bcf files. The vcftools software package (https://github.com/vcftools/vcftools) was then used to export position and allele frequency statistics for the prairie vole *Oxtr* into a single .csv format. (Figure 3.2A)

RESULTS

Individual Variability of Prairie Vole ATAC-Seq Oxtr Signal

To visualize the open chromatin profile for *Oxtr*, we cross-referenced the gene with Ensemble database coordinates (https://www.ensembl.org/Microtus_ochrogaster), resulting in the locus JH996431.1:26.35Kb. The alignment files from the ATAC-Seq data were then individually plotted for each sample, and a graphic was constructed to represent coverage along the corresponding MicOch1.0 chromosomal scaffold JH996431.1 (Figure 3.1, left). From the ATAC-Seq data, we expected to confirm peaks at *Oxtr* in tissues known to express the gene. Contrary to

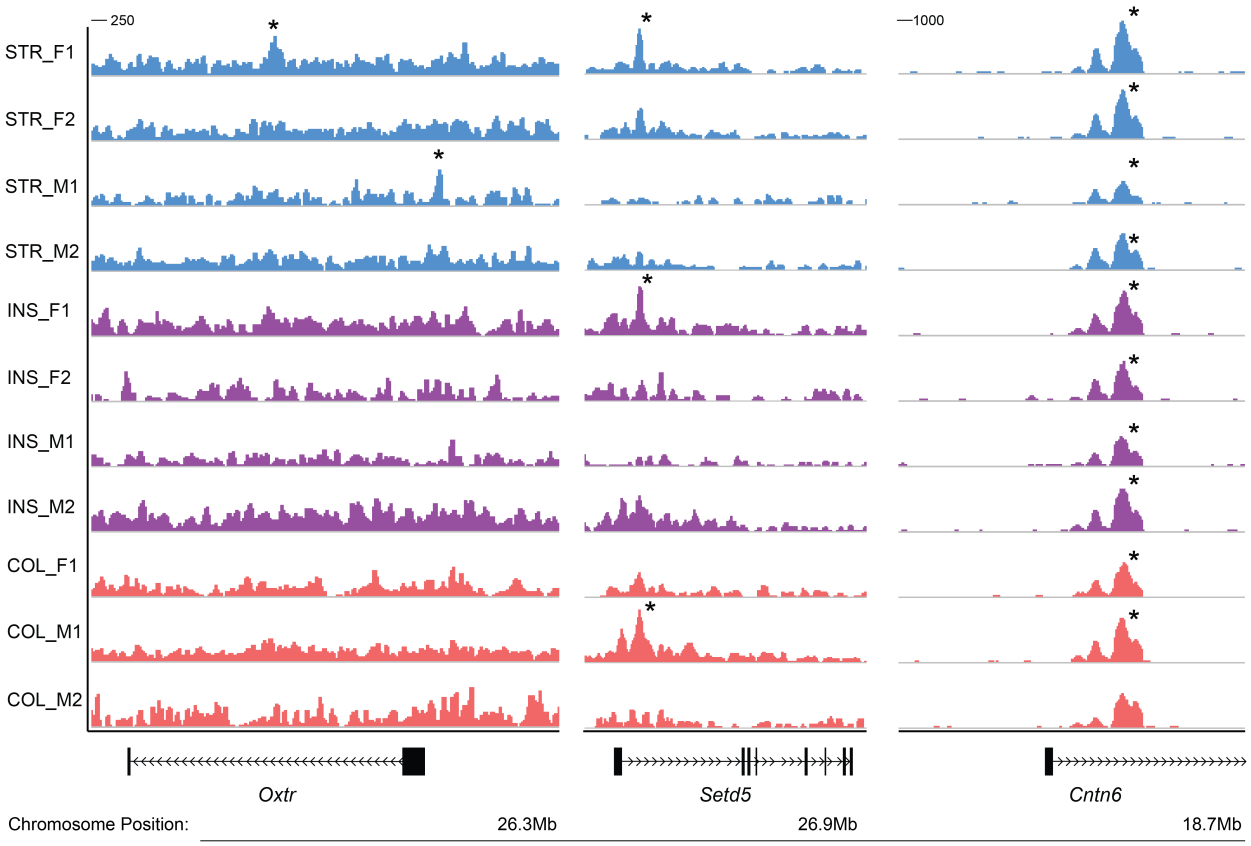


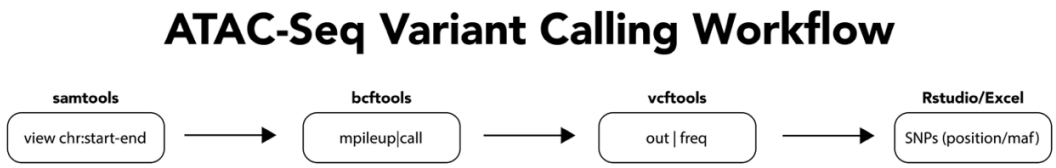
Figure 3.1. Individual Variation in *Oxtr* ATAC-Seq Results.
Representative ATAC-Seq signals for prairie vole chromosome 6 (JH996431.1), including *Oxtr,* for 11 individuals - Male/Female Striatum (STR), Insular Cortex (INS), and Inferior Colliculus (COL). *Setd5* (middle, 26.90Mb) and *Cntn6* (right,18.7Mb) are genes where ATAC-Seq peaks closest to the *Oxtr* locus (26.35Mb) can be identified. (* = ATAC-Seq peak, p=.01)

our hypothesis, a high signal to noise ratio was identified for much of the genomic coordinates spanning 24-26Mb, which includes *Oxtr*. Alignment of peaks identified in Chapter 2 demonstrates overall inconsistency of peak detection at *Oxtr* and other nearby genes for all three brain tissues. *Setd5*, for example, is only 600Kb from *Oxtr* and shows a similar profile for open chromatin (Figure 3.1, middle). We wanted to then explore the extent of high background alignments spanning chromosome 6, to confirm if our observations at *Oxtr* and *Setd5* were unique or a more general characteristic. We found that read depth elsewhere on the chromosome displayed very robust coverage that is coupled with low background. Peak detection in these areas appear to be more consistent as well, such as with the *Cntn6* gene (Figure 3.1, right). The observed individual variation at *Oxtr* can therefore be confidently assessed as a true feature of the ATAC-Seq signal at the *Oxtr* locus for all three tissue types. The *Oxtr* ATAC-Seq signal is suggestive that DNA accessibility at the gene is highly dynamic, with perhaps several different open chromatin profiles that cannot be deduced from bulk tissue samples. In these samples, the data are also clear that neighboring genomic positions to *Oxtr* are also lowly covered, which points to a potential larger cis regulatory complex that may have a larger influence on the overall chromatin state within this chromosomal region.

Establishing a Pipeline for Variant Calling of Prairie Vole ATAC-seq data

Previous work to define polymorphisms with prairie vole genetic data have involved targeted DNA sequencing experiments and a bacterial plasmid clone as a genomic reference. Early variant calling results may therefore be more prone to bias due to PCR duplicates and/or reference insertions and deletions. Our approach involved extracting reads from a whole genome ATAC-Seq alignment, and the SNPs defined in this experiment may be one of the first to be identified

A



B

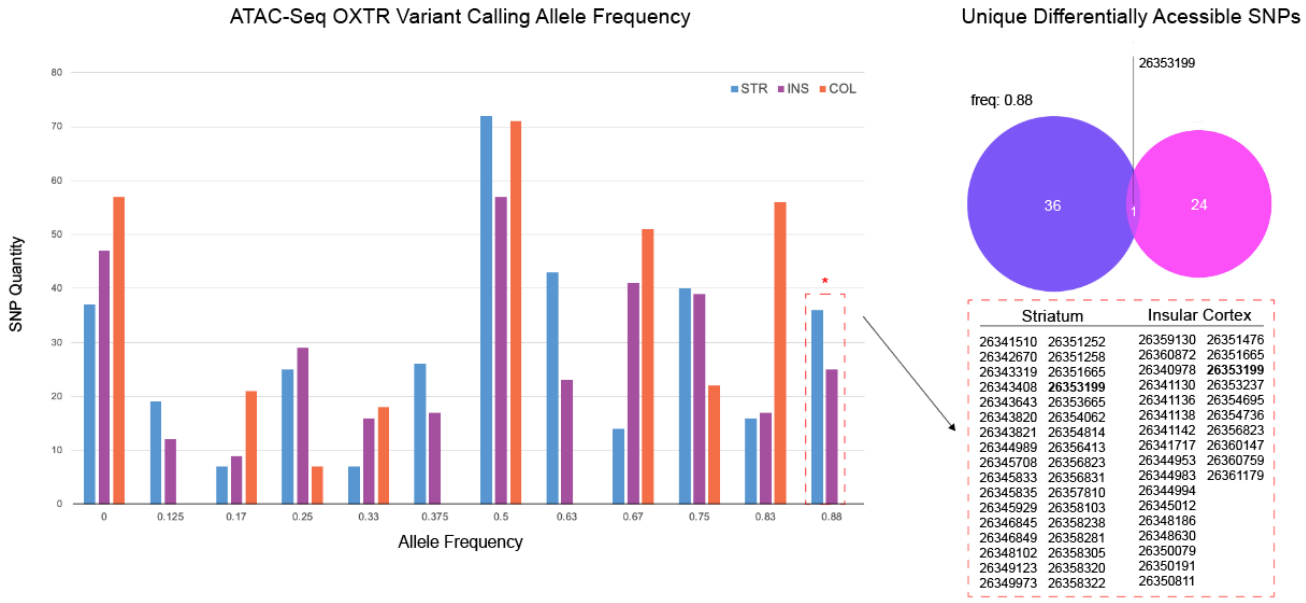


Figure 3.2. Variant Calling Workflow and Brain-region Specific Allele Frequencies.
(A) Graphic overview of variant calling of prairie vole brain ATAC-Seq data. Samtools view command is used to extract aligned reads from a 20Kb chromosomal region spanning *Oxtr;* JH99431.1:26,340,000-26,620,000. The bcftools and vcftools programs are used to perform variant calling and export frequency statistics, which can be .csv formatted and visualized between Microsoft Excel and RStudio. (B) Left - Bar graph of allele frequency intervals and quantity of SNPs defined at each frequency for three different brain areas (STR, INS, COL). SNPs with an allele frequency at the highest interval (0.88) are considered passing threshold for differential accessibility. Right – Ven diagram and list of differentially accessible SNPs between prairie vole Striatum and Insular Cortex tissues. Size of circles correlates with quantity of SNPs passing threshold, and overlap indicates the quantity of SNPs that are shared between the two brain tissues. A list of the individual SNP positions is shown below with the single overlapping SNP, 26353119 in bold.

from prairie vole next generation sequencing data. To capture maximum coverage for each base pair call, sequence data for each brain area was merged. Three aggregate files were fed into a .vcf/.bcf pipeline that contains a variant calling command for exporting the location, chromosome, base pair call, and major allele frequency of every defined SNP. In the Oxtr region spanning 20Kb, we identified 300-342 SNPs for each brain area, with 274 shared SNPs (85%) identified in all three areas (Figure 3.2a).

Differentially Accessible SNPs at Prairie Vole *Oxtr*

Approximately 25% of the prairie vole *Oxtr* SNPs we identified are in an allele frequency of 0.5, which is indicative an equal percentage of accessibility for both alleles. We were interested in classifying the extent of variation in allele frequencies in reads between each brain area, given that the samples are derived from separate brain tissues but from the sample animals. For the 274 shared SNPs, the output SNP frequencies were subset by brain region, and grouped into discrete measurements that reflect the total number of chromosomes used for variant calling, which for our study allowed a major allele frequency (MAF) range from 0.12-0.88. We defined a threshold for a differentially accessible allele as being the maximum frequency detected (0.88). This analysis was able to detect two distinct subsets of SNPs for the striatum and insular cortex with the highest observed frequency (Figure 3.2b). Interestingly, the location of these SNPs is almost entirely non-overlapping with only 1 SNP in high accessibility shared between two brain areas (Figure 3.3a). These results imply that preference for DNA accessibility at *Oxtr* can be driven by one allele over another and that this varies from one area of the brain to another.

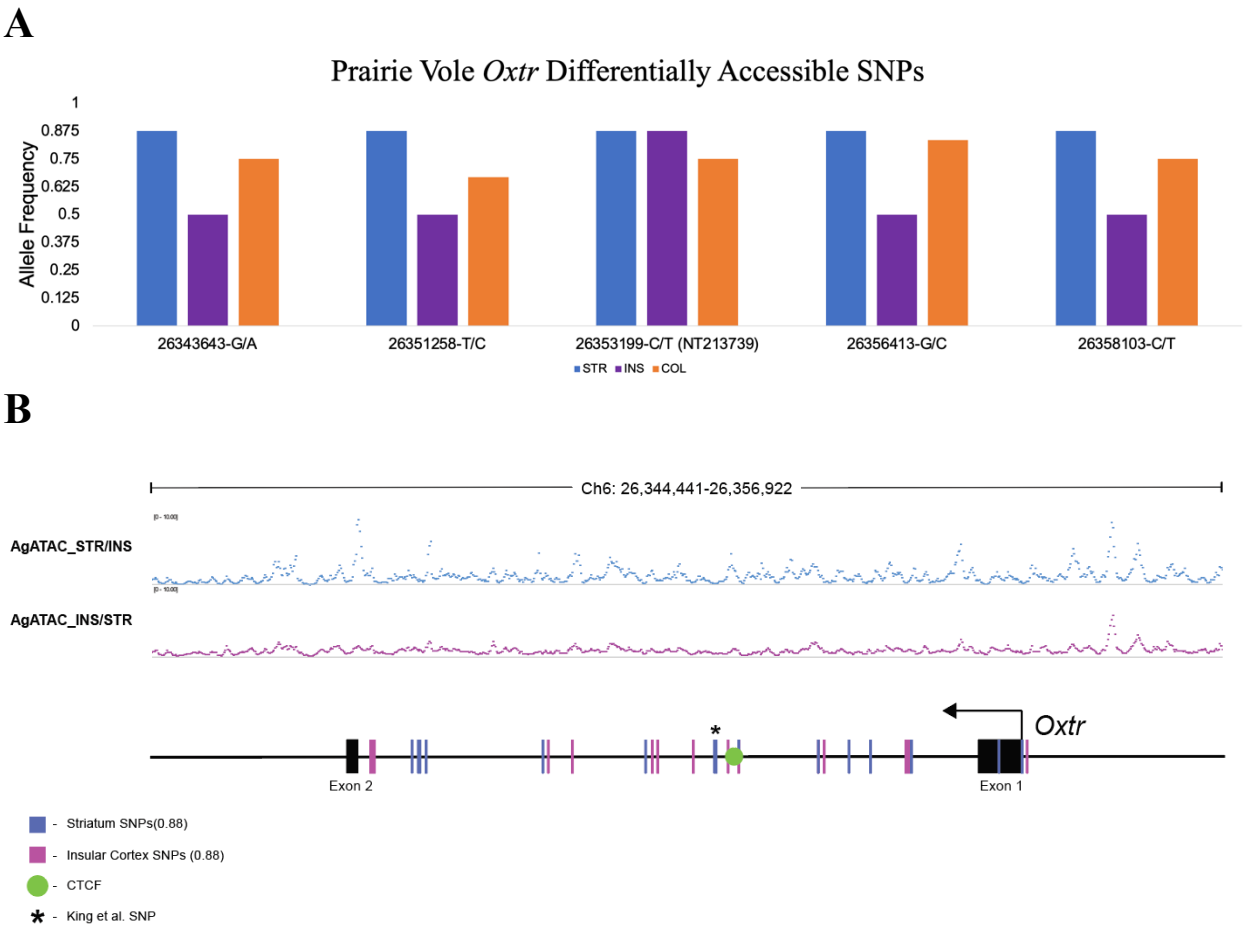


Figure 3.3. Prairie Vole Brain-region Specific Accessible SNPs at *Oxtr*.
(A) SNP allele frequencies for 5 representative intron SNPs with an allele frequency of 0.875 for striatum tissue samples. The frequencies for insular cortex and inferior colliculus are shown for comparison. No linear correlation between is identified between the three brain areas. The previously reported C/T SNP NT213739 is the only polymorphism in highest frequency that is shared between striatum and insular cortex. (B) Aggregate enrichment signals for STR (n=4) and INS (n=4) ATAC-Seq data. *Oxtr* gene is aligned to the aggregate signal with differentially accessible SNPs annotated. Merged ATAC-Seq reads are normalized by average depth and the result divided by the coverage from the compared brain area (i.e. STR/INS, INS/STR). The result indicates the magnitude of difference in coverage at each position for the two different brain areas. DNA binding sites for CTCF (green) can be identified within the sequence spanning differentially accessible SNPs and 3-4x magnitude of coverage difference between STR and INS ATAC-Seq data.

Given that we examined brain tissues that are known to express *Oxtr* in different magnitudes of variation, our findings may provide a molecular basis for highly dynamic vs static expression of *Oxtr* by means of variation in allelic accessibility. We also provide the exact location of the SNPs that may be involved in mediating striatum specific *Oxtr* regulation. Lastly, we sought to map the location of differentially accessible SNPs to the prairie vole *Oxtr* and overlay positions with the ATAC-Seq data for those brain areas. To appropriately represent the SNPs being defined from data that combines males and females, we also created an aggregate ATAC-Seq signal which combined DNA reads from males and females for the same tissue type. Additionally, we divided the read depth of these aggregates by the read depth of the opposing brain area, to highlight discrepancies in coverage at *Oxtr* between the striatum and insular cortex (Figure 3.3b). Although peak detection spanning *Oxtr* was inconsistent in a genome-wide analysis, plotting the coverage difference does show that *Oxtr* is more accessible at the promoter and intron of striatum tissue compared to insular cortex. Interestingly, we noticed the only shared polymorphism (SNP 26353199) with an allele frequency of 0.875 between striatum and insular cortex also corresponds to the location of a previously reported SNP that is believed to be associated with variation in prairie vole *Oxtr* expression (King et al., 2016). We hypothesize that despite observing signs of bias in DNA accessibility between entire *Oxtr* alleles, regulatory mechanisms at or near SNP 26353199 can leave the C variant of this portion of *Oxtr* to be accessible in a manner that is independent of differential gene regulation. The CTCF transcription factor also has a consensus binding sequence less than 500 base pairs away from this position. There could exist a potential for regulatory mechanisms that influence variation in *Oxtr* expression to interact somewhere around this region of the gene, though more experiments will need to be performed to confirm features of enhancer or repressor activity.

DISCUSSION

In the present study, we investigated the *Oxtr* ATAC-Seq signal from three areas of the prairie vole brain, to explore the signature for DNA accessibility among tissues known to express the gene and a region that does not. The ATAC-seq method allowed our laboratory to, for the first time, examine the landscape of open chromatin across the prairie vole genome, and specifically at *Oxtr*. Peaks near promoter regions are generally associated with gene expression, and previous work has demonstrated high levels of *Oxtr* expression in the prairie vole striatum (King et al., 2016). Our hypothesis was therefore that we would detect peaks at the *Oxtr* promoter for striatal and insular cortex samples, and no peaks in the inferior colliculus. In contrast, only 2 peaks in the gene were identified across 11 total ATAC-Seq samples. The inconsistency of peak detection can be partially explained by high levels of individual variation in DNA accessibility at the chromosomal region that spans *Oxtr*. ATAC-Seq data from human and mice studies share some commonality to our observations here in prairie voles, where *Oxtr* is generally less covered compared to other genes (Cusanovich et al., 2018; Fullard et al., 2017). In our research, creating an aggregate of the ATAC-Seq reads for each brain region did reveal approximately 100% more DNA reads aligning to the gene body of *Oxtr* in striatum samples than the inferior colliculus. From one interpretation, the entire *Oxtr* gene may be generally more accessible in tissues where it is more dynamically regulated, rather than only at the promoter region. This scenario would fit with the evolutionary view of neuropeptide receptor genes in voles and other species, where positive selection at hypervariable regulatory elements exists as a generation mechanism to drive phenotypic complexity in receptor expression, and respond differentially based on certain cell stages and environmental conditions. Beyond prairie voles, *Oxtr* can be thought of as a gene that may direct a multitude of transcription factors amid a transient chromatin state that can be highly

responsive in some environmental contexts, and much less responsive in others. Based on our findings, much work can now be done within the prairie vole model to identify a molecular genetic basis for *Oxtr* regulation. What remains to be established is exactly which factors may be interacting to mediate dynamic changes in Oxtr expression, and if these factors differ under varying conditions.

The quantity of *Oxtr* expressing cells may also be significantly different across different cell-types of the tissues that were examined. From the 100K nuclei quantified for each ATAC-seq sample, the highly accessible *Oxtr* fragments could be washed out by those with low accessibility. Further research of the ATAC-Seq signal in specific cell-types of the striatum, such as dopamine neurons and glia will need to be conducted to understand the full context of variation in *Oxtr* DNA accessibility. Our results do confirm that variant calling can be performed from prairie vole ATAC-Seq data, and that *Oxtr* allele frequencies can vary greatly between different brain tissues. We were able to detect many polymorphisms that are reported in previous studies, including NT213739 from the King et al. sequencing experiments. A larger cohort of animals may be necessary to confirm specific allele frequencies, especially as they pertain to DNA accessibility. We expect that SNP frequencies naturally vary across different laboratory colonies or field populations. Studying more genetically diverse animals would help to further deduce potential causative DNA variants from those that are in linkage disequilibrium.

From the perspective of chromatin accessibility, our results encourage the examination of a broader range of genomic regions for the study of genotypes that may contribute to social phenotypes. *Setd5*, for example, encodes a histone lysine methyltransferase that functions to directly regulate chromatin and is thought to be critical for normal brain development (Osipovich et al., 2016; Sessa et al., 2019). For the two female striatum samples, we found that the *Setd5* gene

body is generally more accessible than the two male striatum samples (Figure 3.1). Interestingly, De novo mutations in *Setd5* have been heavily implicated in studies of autism and intellectual disability, which present higher incidence in males (Fernandes et al., 2018; Szczaluba et al., 2016). Another gene, *Cntn6*, is a member of the immunoglobulin superfamily, and is known to function during nervous system development. In humans, *CNTN6* is also linked to autism and 3p25 syndrome (Hu et al., 2015; Kashevarova et al., 2014; Mercati et al., 2017). In prairie voles, approximately 8Mb from *Oxtr*, high signal to noise ratios are coupled with robust intron coverage and peak detection in 10/11 samples for the *Cntn6* gene. Although the observation of DNA accessibility alone does not place casual connection between these genes and any aspect of behavior in prairie voles, the detection of their apparent regulation may be considered a variable or biomarker among different experimental conditions that seek to understand connections between gene regulation and brain phenotypes that drive behavior. As one example, prairie voles are well known for socially monogamous mating strategies. Based on the high variability we observed in naïve prairie voles, we are curious to see next how the ATAC-Seq signal at *Oxtr* may change after pair bond formation and biparental care. Similar to individual differences in *Oxtr* receptor distribution that have been observed, we hypothesize some level of variation in chromatin accessibility will be seen under these behavioral conditions, which may be associated with the onset and frequency of socially specific behaviors.

In summary, the findings presented in the current study are supportive of a highly variable, and perhaps dynamic regulatory environment at *Oxtr* in prairie vole brain tissue. We speculate that this feature may contribute to the diversity in *Oxtr* expression both between individuals and across species. We propose that the lack of consistent peaks makes the gene expression profile of *Oxtr* highly sensitive to genetic variation, leading to evolvability in brain gene expression profiles.

Compared to the insular cortex and inferior colliculus, a higher total of ATAC-Seq reads align to *Oxtr* from striatum tissue. Additionally, a group of 36 polymorphisms are also in a high frequency that is unique to striatum samples. The differentially accessible striatum SNPs somewhat align with the positions at *Oxtr* that have higher coverage than insular samples. Our data provides some of the first functional knowledge about the regulation of *Oxtr* and nearby genes in fresh brain tissue, and confirms the presence of previously reported prairie vole SNPs in a less biased genome-wide analysis. Further, we demonstrate the execution of a novel form of variant calling to define previously undescribed *Oxtr* SNPs with the use of ATAC-Seq DNA reads. In conclusion, our assessment at the *Oxtr* locus supports the existence of high individual variability in DNA accessibility, which is consistent across brain regions that are known to express *Oxtr* to tissues that do not. Interestingly, we show that this phenomenon may be partially explained by differentially accessible alleles, which can comprise a majority of reads in one brain area and much less in another. Our results will have large implications for further research that seeks to manipulate the *Oxtr* sequence to alter its transcription, and the SNPs we have identified provide a much larger framework for assessing prairie vole *Oxtr* variation.

Chapter 4:

Discussion

ABSTRACT

Advancements in molecular biology and DNA sequencing have provided new opportunities for researchers to explore functional relationships between genetic variation and gene regulation. The Assay for Transposase Accessible Chromatin with DNA Sequencing (ATAC-Seq) is a popular method which can provide genome-wide information that represents open regions of DNA. This method also provides an opportunity to define single nucleotide polymorphisms (SNPs), which are common variables used in behavioral association studies. ATAC-Seq has allowed scientists to investigate DNA accessibility profiles for numerous unique tissues and cell types. ATAC-Seq in areas of the brain have been investigated in human and mouse, however, DNA accessibility is still not well documented for many other species. We employed ATAC-Seq in the socially monogamous prairie vole to examine chromatin accessibility in three areas of the brain: the striatum, insular cortex, and inferior colliculus. By gaining insights about the availability of DNA within prairie vole brain tissue, we bring attention to genomic regions that may be important for tissue specific gene regulation. Two important goals for this research were to obtain data from a diversified *M. Ochrogaster* sample, and to measure chromatin accessibility and further characterize SNPs at the oxytocin receptor gene (*Oxtr*). To meet these goals, we pooled individuals from distantly related groups of males and females, and generated ATAC-Seq data that can be mapped to the genome with sufficient coverage for defining genetic variants. Our findings provide evidence for substantial sex differences in open chromatin regions. At *Oxtr*, we show evidence for high individual variability across samples, as well as allele-specific DNA accessibility between different areas of the prairie vole brain. The ATAC-Seq results support a growing body of evidence that supports the flexibility of *Oxtr* regulation in different brain areas. The work described in this dissertation serves as only the first examination of chromatin accessibility in the prairie vole

species, and additional experiments should be performed with larger samples and an ability to discriminate between cell-types. Ideally, methods such as RNA-Seq and ChIP-Seq may also be used in combination with ATAC-Seq to examine the regulatory environment of specific genes such as *Oxtr*. In this final discussion, we focus on these other sequencing methods that can be applied to the study of gene regulation in the prairie vole, and we provide further interpretation of our own findings.

INTRODUCTION

(OXT) is a small and highly conserved neuropeptide that is known to modulate the neural circuitry associated with social behavior and the processing of social information. The OXT receptor is a g-protein coupled receptor that functions to drive cellular changes in response to OXT, and the receptor is encoded by the gene *Oxtr*. Many recent efforts to understand the neuro-anatomical and genetic correlates that are involved in regulating human social behavior have focused on the OT system, as discussed in Chapter 1. Human and animal studies have revealed that *OXTR/Oxtr* SNPs may be important factors that explain individual differences in socio-behavioral traits, such as trust, empathy, and affiliation (King et al., 2016; Kosfeld et al., 2005; Uzefovsky et al., 2015). The regulatory mechanisms by which SNPs drive phenotypic complexity to influence these behaviors is still largely unknown. Prairie voles have historically been researched to study the biological basis of specific social traits by examining characteristics of genes, neuroendocrine factors, and neural networks that may explain interspecies and intraspecies in social behavior (Amadei et al., 2017; Aragona et al., 2003; Lim et al., 2004). In this dissertation, we aimed to provide a foundation for studying chromatin regulation in the prairie vole brain by characterizing the landscape of DNA accessibility throughout the genome, with a specific interest in *Oxtr*. To achieve this, we developed a workflow for the Assay for Transposase Accessible Sequencing (ATAC-Seq), which allows precise measurement of DNA accessibility throughout the genome. Its advantages are small sample input requirements, short preparation time, and fast turnaround of data. As discussed in Chapters 2 and 3, we demonstrated how the ATAC-Seq method alone can provide unique insights about functional differences in DNA accessibility – between sexes and across different areas of the brain. Our results place an emphasis for further study of the prairie vole brain within the context of the growing field of chromatin research. We now seek to

highlight other ATAC-seq approaches and make suggestions for furthering our understanding of the molecular mechanisms of *Oxtr* regulation in the prairie vole research model.

MOLECULAR TECHNIQUES FOR STUDYING GENE REGULATION IN THE BRAIN

Expanding ATAC-Seq Techniques

ATAC-Seq was developed in 2015, and initially the method was designed to process nuclei from fresh tissues and cell cultures (Buenrostro et al., 2015). The technique has quickly grown in popularity within the neuroscience field due to sample input requirements that increase the feasibility of studying chromatin landscapes in small neuronal populations (Cusanovich et al., 2018; Fullard et al., 2018; Su et al., 2017). Varieties of sample types that have been used for ATAC-Seq has given rise to several derivative protocols that have been developed to accompany different storage and extraction conditions (Corces et al., 2017; Peng et al., 2021; Spektor et al., 2019), as well as for ATAC-seq of single cells (Cusanovich et al., 2018; Mezger et al., 2018).

Many of the reports using ATAC-Seq have been exploratory with the goal to simply characterize DNA accessibility in model organisms and various sample types. A handful of research has also focused on comparisons between the ATAC-Seq signal and other sequencing data, such as gene expression as measured by RNA-Seq (Ackermann et al., 2016; Hendrickson et al., 2018; Li et al., 2022; Narayanan et al., 2020). While in some organisms and sample types, chromatin accessibility and mRNA levels can be correlated, other research has shown contrasting results with very little correlation between variation in accessible peaks and gene expression (Sugino et al 2017). What is clear from these early studies is that gene expression and DNA accessibility alone do not give a complete picture of molecular genetic regulatory processes, but together the methods can help to build knowledge about the relationship between chromatin

structure and transcript levels in new ways that were previously less plausible. The current implication of ATAC-Seq studies combined with RNA-Seq is that it is more informative to investigate variation in measurements than to expect high correlations between the two types of data.

In addition to the increasing number of unique ATAC-Seq preparations, new forms of analyses are readily being developed in order to extrapolate more functional information from accessible sites, including transcription factor binding locations and de-convolution of cell types from bulk tissue samples (Chen et al., 2021; Li et al., 2019). With the use cases for the ATAC-Seq method continuing to expand into new areas of science, it appears to be a method that will last for several years and aid in the discovery of unique characteristics of chromatin availability. Some behavioral studies have begun to measure chromatin accessibility as biomarker for environmental influences on gene regulation. In rhesus macaques, for example, ATAC-Seq shows that some alterations in chromatin accessibility can be highly influenced by social status (Noah Snyder-Mackler et al. 2019). Some justification may already exist for including ATAC-Seq and other chromatin-based methods in prairie vole behavioral paradigms such as partner preference, or even comparisons to its closely related meadow vole species.

RNA-Seq

RNA-Seq is currently the most popular technique for quantifying gene transcripts, where sequencing of all mRNAs can provide an organism's transcriptome. Prior to the availability of Next Generation Sequencing methods, RNA molecules could be visually measured with in situ hybridization, and this technique is still widely used in new methods such as RNAscope, which allows single cell sequencing in combination with cell imaging. Hybridization approaches are also

still important for studying rapidly degrading RNAs, which includes the transcript for the OXT receptor. In recent years, more robust RNA-Seq datasets have been generated from tissues and cell types of the brain for the studying development, disease, and behavior (Hook et al., 2018; Luo et al., 2022; Tiklova et al., 2019). Many advancements in bioinformatics have also been centered working with RNA-Seq data, and the analysis for differential chromatin accessibility in ATAC-Seq is, in fact, derived from analysis of differential expression for RNA-Seq. Experimentally, limitations to replicability in RNA-Seq findings can include the predominance of regulatory RNAs, especially ribosomal RNA (rRNA). rRNA result from the expression of just a small handful of ribosomal genes but can account for up to 95% of total RNA during extraction. Consequently, many transcriptomics studies can be well executed experimentally but still yield inconsistent results due to rRNA. Many researchers are now including the use of rRNA removal kits to accommodate contamination of these types.

Gene encoding mRNA are usually the main target of RNA-Seq experiments, as many studies will compare transcript levels across different tissues, cell types, and treatment conditions. Methods have also been developed to capture unique classes of RNA that may play gene regulatory roles such as micro-RNA (miRNA) and long noncoding RNA (lncRNA). In one study, an artificial small RNA sequence known as a short hairpin RNA (shRNA), was used as an RNA interference (RNAi) mechanism to knockdown the prairie vole *OXTR* gene in the nucleus accumbens (Keebaugh et al., 2015). RNAi of the prairie vole *Oxtr* in female NAcc resulted in negative impacts on behavior related to social attachment, such as partner preference and alloparental care. Some efforts have also sought to characterize the transcriptomics that underly monogamy by looking at similar regions of the brain in distantly related monogamous species (Young et al., 2019). A similar approach was also used to compare expression in brain areas between prairie voles and the non-

monogamous meadow vole (Duclot et al., 2022). Both studies were able to demonstrate the existence of unique transcriptional profiles that separated monogamous and non-monogamous species, often at genes involved in regulation of neuronal signaling.

ChIP-Seq

In the cell nucleus, chromatin is mostly comprised of DNA that is packaged tightly around proteins called histones. Chemical modifications at the tails of histones influence the affinity between histone complexes and DNA. This system of chromatin organization serves also as a signature for certain gene states and the recruitment of other proteins to directly mediate expression of genes. Chromatin immunoprecipitation (ChIP) is a technique that can preserve nuclear DNA-protein complexes, and is typically performed by formaldehyde crosslinking of cells or tissues, chromatin isolation, and immunoprecipitation to select for proteins of interest (O'Neill and Turner, 1996). ChIP can be highly informative and has been used for decades to help to associate specific proteins, such as transcription factors, with specific sequences of DNA. Popular factors that are commonly studied with ChIP include cohesin, condensin, the CCCTC binding factor (CTCF), and lysine methylation/acetylation of histone positions H3K4me and H3K27ac. In ChIP-Seq, mapping of transcription factor binding to specific DNA sequences is characterized across the genome by comparing data from antibody preparations to non-antibody preparations. ChIP-Seq has also provided a pathway to understanding how stages of development and cell differentiation are mediated by different combinations of transcription factor binding and histone activity (Mundade et al., 2014).

Important considerations when performing ChIP-Seq experiments are properly sheared chromatin, antibody concentration, and PCR amplification steps. The laboratory that was used to

perform the ATAC-Seq has also recently acquired all equipment that is necessary for performing ChIP-Seq, and preliminary work identified CTCF to be enriched at the prairie vole *Oxtr* intron (Gardner and Young, Unpublished). One of the future aims of this laboratory is to generate ChIP-Seq signals for CTCF, H3K4 methylation, and H3K27 acetylation in various prairie vole brain regions. ChIP-Seq and ATAC-Seq can serve a straightforward experimental combination due to similar cell/tissue processing, protocol steps and DNA input requirements.

DISCUSSION; TOWARDS MOLECULAR MECHANISMS OF *OXTR* REGULATION

Social behavior and cognition are aspects of life that impact all humans and animals. Such an importance is demonstrated by the highly conserved processes that guide how the social brain is formed, and the adaptable processes in the brain that allow individuals to readily respond to the environment. Studying the molecular mechanisms that underlie how we perceive and interact with our social world will be essential to developing effective treatments for social deficits. Much research has been done to identify specific genes, genetic variants, and areas of the brain that are critical for social functioning. The prairie vole research model has specifically helped scientists to further build systems-level knowledge about the critical role of the oxytocin system in mediating social traits. This includes defining variation at genes such as *Oxtr*, linking polymorphisms to brain phenotypes involved in oxytocin signaling, and behavioral paradigms for quantifying components of social interaction that are mediated by oxytocin. Studying the chromatin environment of brain regions known to function in regulating social reward and reinforcement in prairie voles, such as the striatum, may be the next obvious step towards understanding the relationship between genes and behavior at the molecular level.

The primary goals for the research described in this dissertation was to, for the first time, characterize chromatin availability at *Oxtr* in the prairie vole striatum, and to be able to compare the ATAC signal between other tissues of the brain. We decided to use the Assay for Transposase Accessible Chromatin and DNA Sequencing (ATAC-Seq) because of its reputation as an unbiased technique for classifying features of the chromatin regulatory landscape on a genome-wide scale. Generating a meaningful dataset required careful dissection of fresh brain tissues that previously are known to express *Oxtr* in different in highly variable and invariable amounts. Comparing our results in prairie voles to similar approaches in humans and mice, a handful of factors limited the scope of the findings that have been presented here. Increased sample sizes of pooled animals, more dissected brain areas, distinguishment of cell-types, and an updated genome annotation may have contributed greatly to more refined peak detection and likely further demonstrated the extent of individual variability in accessibility at *Oxtr.*

Although there is a well-established interest in *OXTR* polymorphisms as predictors of behavioral phenotypes and/or biomarkers for psychiatric disorders, less often are *OXTR* SNPs examined for functional roles in gene regulatory processes. Establishing the ATAC-Seq method for prairie vole brain tissue provided us with a promising first peek into the extent of variation in how DNA is regulated across the genome, with specific attention to accessibility of genetic variants at *Oxtr*. Much of the science that surrounds the intersection between genes and behavior work from a common belief that neural gene expression fundamentally underlies the capacity for social bonding, as demonstrated in the early shRNA experiments, and comparisons between prairie and meadow vole transcriptome, as well as other species-species comparisons. Other methods have been used to show that environmental factors and epigenetic changes may also play a critical role in development and formation of social behaviors such as bonding and affiliation. Despite over a

decade of more closely elucidating the OXT system at the level of DNA and RNA, few efforts in prairie voles have demonstrated causative molecular factors such as transcription factors or chromatin modifications that can be studied from a perspective of inheritance and epigenetics. Considering that prairie vole research is an established behavioral model where single nucleotide polymorphisms (SNPs) and brain *Oxtr* expression are already well characterized, the combination of methods such as ATAC-Seq, RNA-Seq and ChIP-Seq should present a promising future for investigating the dynamics of gene by environment interactions at a molecular level.

In Summary, our research identified, for the first time, hundreds of thousands of open chromatin regions within prairie vole brain tissue. We then used this data to estimate differential accessibility of open chromatin between different brain areas. Our experiments provide evidence that the striatum is differentially accessible at several genes involved in DNA binding. We also show that sex explains most of the individual variation in peak magnitudes between the three brain areas that were studied. At *Oxtr*, the ATAC-Seq signal displays high individual variation in DNA accessibility. We defined novel prairie vole SNPs at the *Oxtr* intron and estimated individual SNP allele frequencies from the ATAC-Seq data for each brain area. Our data at *Oxtr* also provide evidence that unique groups of polymorphisms may be differentially accessible between different brain areas, which is promising for trying to uncover transcription factor activity that may be unique to areas such as the striatum. These findings provide a baseline for the ATAC-Seq signal in the prairie vole brain, and we can create new hypotheses about specific loci that may be involved in regulation of *Oxtr*.

REFERENCES

Ackermann, A.M., Wang, Z., Schug, J., Naji, A., Kaestner, K.H., 2016. Integration of ATAC-seq and RNA-seq identifies human alpha cell and beta cell signature genes. Mol Metab 5, 233-244.

Amadei, E.A., Johnson, Z.V., Jun Kwon, Y., Shpiner, A.C., Saravanan, V., Mays, W.D., Ryan, S.J., Walum, H., Rainnie, D.G., Young, L.J., Liu, R.C., 2017. Dynamic corticostriatal activity biases social bonding in monogamous female prairie voles. Nature 546, 297-301.

Aragona, B.J., Liu, Y., Curtis, J.T., Stephan, F.K., Wang, Z., 2003. A critical role for nucleus accumbens dopamine in partner-preference formation in male prairie voles. J Neurosci 23, 3483-3490.

Baker, M., Lindell, S.G., Driscoll, C.A., Zhou, Z., Yuan, Q., Schwandt, M.L., Miller-Crews, I., Simpson, E.A., Paukner, A., Ferrari, P.F., Sindhu, R.K., Razaqyar, M., Sommer, W.H., Lopez, J.F., Thompson, R.C., Goldman, D., Heilig, M., Higley, J.D., Suomi, S.J., Barr, C.S., 2017. Early rearing history influences oxytocin receptor epigenetic regulation in rhesus macaques. Proc Natl Acad Sci U S A 114, 11769-11774.

Balaton, B.P., Fornes, O., Wasserman, W.W., Brown, C.J., 2021. Cross-species examination of X-chromosome inactivation highlights domains of escape from silencing. Epigenetics Chromatin 14, 12.

Bales, K.L., Perkeybile, A.M., Conley, O.G., Lee, M.H., Guoynes, C.D., Downing, G.M., Yun, C.R., Solomon, M., Jacob, S., Mendoza, S.P., 2013. Chronic intranasal oxytocin causes long-term impairments in partner preference formation in male prairie voles. Biol Psychiatry 74, 180-188.

Barberis, C., Mouillac, B., Durroux, T., 1998. Structural bases of vasopressin/oxytocin receptor function. J Endocrinol 156, 223-229.

Baribeau, D.A., Dupuis, A., Paton, T.A., Scherer, S.W., Schachar, R.J., Arnold, P.D., Szatmari, P., Nicolson, R., Georgiades, S., Crosbie, J., Brian, J., Iaboni, A., Lerch, J., Anagnostou, E., 2017. Oxytocin Receptor Polymorphisms are Differentially Associated with Social Abilities across Neurodevelopmental Disorders. Sci Rep 7, 11618.

Bentsen, M., Goymann, P., Schultheis, H., Klee, K., Petrova, A., Wiegandt, R., Fust, A., Preussner, J., Kuenne, C., Braun, T., Kim, J., Looso, M., 2020. ATAC-seq footprinting unravels kinetics of transcription factor binding during zygotic genome activation. Nat Commun 11, 4267.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., Greenleaf, W.J., 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods 10, 1213-1218.

Buenrostro, J.D., Wu, B., Chang, H.Y., Greenleaf, W.J., 2015. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. Curr Protoc Mol Biol 109, 21 29 21-21 29 29.

Burkett, J.P., Andari, E., Johnson, Z.V., Curry, D.C., de Waal, F.B., Young, L.J., 2016. Oxytocin-dependent consolation behavior in rodents. Science 351, 375-378.

Carter, C.S., DeVries, A.C., Getz, L.L., 1995. Physiological substrates of mammalian monogamy: the prairie vole model. Neurosci Biobehav Rev 19, 303-314.

Carter, C.S., Getz, L.L., 1993. Monogamy and the prairie vole. Sci Am 268, 100-106.

Cesarini, D., Dawes, C.T., Fowler, J.H., Johannesson, M., Lichtenstein, P., Wallace, B., 2008. Heritability of cooperative behavior in the trust game. Proc Natl Acad Sci U S A 105, 3721-3726.

Chen, F.S., Kumsta, R., Dvorak, F., Domes, G., Yim, O.S., Ebstein, R.P., Heinrichs, M., 2015. Genetic modulation of oxytocin sensitivity: a pharmacogenetic approach. Transl Psychiatry 5, e664.

Chen, Z., Chen, W., Li, Y., Moos, M., Jr., Xiao, D., Wang, C., 2022. Single-nucleus chromatin accessibility and RNA sequencing reveal impaired brain development in prenatally e-cigarette exposed neonatal rats. iScience 25, 104686.

Chen, Z., Zhang, J., Liu, J., Zhang, Z., Zhu, J., Lee, D., Xu, M., Gerstein, M., 2021. SCAN-ATAC-Sim: a scalable and efficient method for simulating single-cell ATAC-seq data from bulk-tissue experiments. Bioinformatics.

Chini, B., Mouillac, B., Balestre, M.N., Trumpp-Kallmeyer, S., Hoflack, J., Hibert, M., Andriolo, M., Pupier, S., Jard, S., Barberis, C., 1996. Two aromatic residues regulate the response of the human oxytocin receptor to the partial agonist arginine vasopressin. FEBS Lett 397, 201-206.

Cho, M.M., DeVries, A.C., Williams, J.R., Carter, C.S., 1999. The effects of oxytocin and vasopressin on partner preferences in male and female prairie voles (Microtus ochrogaster). Behav Neurosci 113, 1071-1079.

Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., Kathiria, A., Cho, S.W., Mumbach, M.R., Carter, A.C., Kasowski, M., Orloff, L.A., Risca, V.I., Kundaje, A., Khavari, P.A., Montine, T.J., Greenleaf, W.J., Chang, H.Y., 2017. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat Methods 14, 959-962.

Corley, S.M., MacKenzie, K.L., Beverdam, A., Roddam, L.F., Wilkins, M.R., 2017. Differentially expressed genes from RNA-Seq and functional enrichment results are affected by the choice

of single-end versus paired-end reads and stranded versus non-stranded protocols. BMC Genomics 18, 399.

Creswell, K.G., Wright, A.G., Troxel, W.M., Ferrell, R.E., Flory, J.D., Manuck, S.B., 2015. OXTR polymorphism predicts social relationships through its effects on social temperament. Soc Cogn Affect Neurosci 10, 869-876.

Cusanovich, D.A., Hill, A.J., Aghamirzaie, D., Daza, R.M., Pliner, H.A., Berletch, J.B., Filippova, G.N., Huang, X., Christiansen, L., DeWitt, W.S., Lee, C., Regalado, S.G., Read, D.F., Steemers, F.J., Disteche, C.M., Trapnell, C., Shendure, J., 2018. A Single-Cell Atlas of In Vivo Mammalian Chromatin Accessibility. Cell 174, 1309-1324 e1318.

Cushing, B.S., Martin, J.O., Young, L.J., Carter, C.S., 2001. The effects of peptides on partner preference formation are predicted by habitat in prairie voles. Horm Behav 39, 48-58.

Dale, H.H., 1909. The Action of Extracts of the Pituitary Body. Biochem J 4, 427-447.

Dayanithi, G., Sabatier, N., Widmer, H., 2000. Intracellular calcium signalling in magnocellular neurones of the rat supraoptic nucleus: understanding the autoregulatory mechanisms. Exp Physiol 85 Spec No, 75S-84S.

Dayanithi, G., Widmer, H., Richard, P., 1996. Vasopressin-induced intracellular Ca2+ increase in isolated rat supraoptic cells. J Physiol 490 (Pt 3), 713-727.

De Dreu, C.K., Baas, M., Roskes, M., Sligte, D.J., Ebstein, R.P., Chew, S.H., Tong, T., Jiang, Y., Mayseless, N., Shamay-Tsoory, S.G., 2014. Oxytonergic circuitry sustains and enables creative cognition in humans. Soc Cogn Affect Neurosci 9, 1159-1165.

de la Torre-Ubieta, L., Stein, J.L., Won, H., Opland, C.K., Liang, D., Lu, D., Geschwind, D.H., 2018. The Dynamic Landscape of Open Chromatin during Human Cortical Neurogenesis. Cell 172, 289-304 e218.

Di Napoli, A., Warrier, V., Baron-Cohen, S., Chakrabarti, B., 2014. Genetic variation in the oxytocin receptor (OXTR) gene is associated with Asperger Syndrome. Mol Autism 5, 48.

Do-Rego, J.L., Acharjee, S., Seong, J.Y., Galas, L., Alexandre, D., Bizet, P., Burlet, A., Kwon, H.B., Luu-The, V., Pelletier, G., Vaudry, H., 2006. Vasotocin and mesotocin stimulate the biosynthesis of neurosteroids in the frog brain. J Neurosci 26, 6749-6760.

Du Vigneaud, V., 1954. Hormones of the posterior pituitary gland: oxytocin and vasopressin. Harvey Lect 50, 1-26.

Du Vigneaud, V., Ressler, C., Trippett, S., 1953. The sequence of amino acids in oxytocin, with a proposal for the structure of oxytocin. J Biol Chem 205, 949-957.

Duclot, F., Sailer, L., Koutakis, P., Wang, Z., Kabbaj, M., 2022. Transcriptomic Regulations Underlying Pair-bond Formation and Maintenance in the Socially Monogamous Male and Female Prairie Vole. Biol Psychiatry 91, 141-151.

Dulac, C., 2010. Brain function and chromatin plasticity. Nature 465, 728-735.

Fang, R., Wang, C., Skogerbo, G., Zhang, Z., 2015. Functional diversity of CTCFs is encoded in their binding motifs. BMC Genomics 16, 649.

Ferguson, J.N., Aldag, J.M., Insel, T.R., Young, L.J., 2001. Oxytocin in the medial amygdala is essential for social recognition in the mouse. J Neurosci 21, 8278-8285.

Fernandes, I.R., Cruz, A.C.P., Ferrasa, A., Phan, D., Herai, R.H., Muotri, A.R., 2018. Genetic variations on SETD5 underlying autistic conditions. Dev Neurobiol 78, 500-518.

Freedman, M.L., Monteiro, A.N., Gayther, S.A., Coetzee, G.A., Risch, A., Plass, C., Casey, G., De Biasi, M., Carlson, C., Duggan, D., James, M., Liu, P., Tichelaar, J.W., Vikis, H.G., You, M., Mills, I.G., 2011. Principles for the post-GWAS functional characterization of cancer risk loci. Nat Genet 43, 513-518.

Frey, W.D., Sharma, K., Cain, T.L., Nishimori, K., Teruyama, R., Kim, J., 2018. Oxytocin receptor is regulated by Peg3. PLoS One 13, e0202476.

Fuchs, A.R., Saito, S., 1971. Pituitary oxytocin and vasopressin content of pregnant rats before, during and after parturition. Endocrinology 88, 574-578.

Fullard, J.F., Giambartolomei, C., Hauberg, M.E., Xu, K., Voloudakis, G., Shao, Z., Bare, C., Dudley, J.T., Mattheisen, M., Robakis, N.K., Haroutunian, V., Roussos, P., 2017. Open chromatin profiling of human postmortem brain infers functional roles for non-coding schizophrenia loci. Hum Mol Genet 26, 1942-1951.

Fullard, J.F., Hauberg, M.E., Bendl, J., Egervari, G., Cirnaru, M.D., Reach, S.M., Motl, J., Ehrlich, M.E., Hurd, Y.L., Roussos, P., 2018. An atlas of chromatin accessibility in the adult human brain. Genome Res 28, 1243-1252.

Gallagher, M.D., Chen-Plotkin, A.S., 2018. The Post-GWAS Era: From Association to Function. Am J Hum Genet 102, 717-730.

Gegenhuber, B., Tollkuhn, J., 2020. Signatures of sex: Sex differences in gene expression in the vertebrate brain. Wiley Interdiscip Rev Dev Biol 9, e348.

Gegenhuber, B., Wu, M.V., Bronstein, R., Tollkuhn, J., 2022. Gene regulation by gonadal hormone receptors underlies brain sex differences. Nature 606, 153-159.

Gorewit, R.C., Wachs, E.A., Sagi, R., Merrill, W.G., 1983. Current concepts on the role of oxytocin in milk ejection. J Dairy Sci 66, 2236-2250.

Gray, J.D., Kogan, J.F., Marrocco, J., McEwen, B.S., 2017. Genomic and epigenomic mechanisms of glucocorticoids in the brain. Nat Rev Endocrinol 13, 661-673.

Halstead, M.M., Kern, C., Saelao, P., Wang, Y., Chanthavixay, G., Medrano, J.F., Van Eenennaam, A.L., Korf, I., Tuggle, C.K., Ernst, C.W., Zhou, H., Ross, P.J., 2020. A

comparative analysis of chromatin accessibility in cattle, pig, and mouse tissues. BMC Genomics 21, 698.

Harabula, I., Pombo, A., 2021. The dynamics of chromatin architecture in brain development and function. Curr Opin Genet Dev 67, 84-93.

Hendrickson, D.G., Soifer, I., Wranik, B.J., Botstein, D., Scott McIsaac, R., 2018. Simultaneous Profiling of DNA Accessibility and Gene Expression Dynamics with ATAC-Seq and RNA-Seq. Methods Mol Biol 1819, 317-333.

Hodes, G.E., Pfau, M.L., Purushothaman, I., Ahn, H.F., Golden, S.A., Christoffel, D.J., Magida, J., Brancato, A., Takahashi, A., Flanigan, M.E., Menard, C., Aleyasin, H., Koo, J.W., Lorsch, Z.S., Feng, J., Heshmati, M., Wang, M., Turecki, G., Neve, R., Zhang, B., Shen, L., Nestler, E.J., Russo, S.J., 2015. Sex Differences in Nucleus Accumbens Transcriptome Profiles Associated with Susceptibility versus Resilience to Subchronic Variable Stress. J Neurosci 35, 16362-16376.

Hook, P.W., McClymont, S.A., Cannon, G.H., Law, W.D., Morton, A.J., Goff, L.A., McCallion, A.S., 2018. Single-Cell RNA-Seq of Mouse Dopaminergic Neurons Informs Candidate Gene Selection for Sporadic Parkinson Disease. Am J Hum Genet 102, 427-446.

Howell, W.H., 1898. The Physiological Effects of Extracts of the Hypophysis Cerebri and Infundibular Body. J Exp Med 3, 245-258.

Hu, J., Liao, J., Sathanoori, M., Kochmar, S., Sebastian, J., Yatsenko, S.A., Surti, U., 2015. CNTN6 copy number variations in 14 patients: a possible candidate gene for neurodevelopmental and neuropsychiatric disorders. J Neurodev Disord 7, 26.

Insel, T.R., Shapiro, L.E., 1992. Oxytocin receptor distribution reflects social organization in monogamous and polygamous voles. Proc Natl Acad Sci U S A 89, 5981-5985.

Israel, S., Lerer, E., Shalev, I., Uzefovsky, F., Reibold, M., Bachnermelman, R., Granot, R., Bornstein, G., Knafo, A., Yirmiya, N., 2008. Molecular genetic studies of the arginine vasopressin 1a receptor (AVPR1a) and the oxytocin receptor (OXTR) in human behaviour: from autism to altruism with some notes in between, Advances in Vasopressin and Oxytocin — From Genes to Behaviour to Disease, 435-449.

Israel, S., Lerer, E., Shalev, I., Uzefovsky, F., Riebold, M., Laiba, E., Bachner-Melman, R., Maril, A., Bornstein, G., Knafo, A., Ebstein, R.P., 2009. The oxytocin receptor (OXTR) contributes to prosocial fund allocations in the dictator game and the social value orientations task. PLoS One 4, e5535.

Jauregui-Lozano, J., Bakhle, K., Weake, V.M., 2021. In vivo tissue-specific chromatin profiling in Drosophila melanogaster using GFP-tagged nuclei. Genetics 218.

Johnson, Z.V., Walum, H., Jamal, Y.A., Xiao, Y., Keebaugh, A.C., Inoue, K., Young, L.J., 2016. Central oxytocin receptors mediate mating-induced partner preferences and enhance correlated activation across forebrain nuclei in male prairie voles. Horm Behav 79, 8-17.

Karabacak Calviello, A., Hirsekorn, A., Wurmus, R., Yusuf, D., Ohler, U., 2019. Reproducible inference of transcription factor footprints in ATAC-seq and DNase-seq datasets using protocol-specific bias modeling. Genome Biol 20, 42.

Kashevarova, A.A., Nazarenko, L.P., Schultz-Pedersen, S., Skryabin, N.A., Salyukova, O.A., Chechetkina, N.N., Tolmacheva, E.N., Rudko, A.A., Magini, P., Graziano, C., Romeo, G., Joss, S., Tumer, Z., Lebedev, I.N., 2014. Single gene microdeletions and microduplication of 3p26.3 in three unrelated families: CNTN6 as a new candidate gene for intellectual disability. Mol Cytogenet 7, 97.

Keebaugh, A.C., Barrett, C.E., Laprairie, J.L., Jenkins, J.J., Young, L.J., 2015. RNAi knockdown of oxytocin receptor in the nucleus accumbens inhibits social attachment and parental care in monogamous female prairie voles. Soc Neurosci 10, 561-570.

Keebaugh, A.C., Young, L.J., 2011. Increasing oxytocin receptor expression in the nucleus accumbens of pre-pubertal female prairie voles enhances alloparental responsiveness and partner preference formation as adults. Horm Behav 60, 498-504.

Kelly, A.M., Goodson, J.L., 2014. Hypothalamic oxytocin and vasopressin neurons exert sex-specific effects on pair bonding, gregariousness, and aggression in finches. Proc Natl Acad Sci U S A 111, 6069-6074.

Kim, H.S., Sherman, D.K., Sasaki, J.Y., Xu, J., Chu, T.Q., Ryu, C., Suh, E.M., Graham, K., Taylor, S.E., 2010. Culture, distress, and oxytocin receptor polymorphism (OXTR) interact to influence emotional support seeking. Proc Natl Acad Sci U S A 107, 15717-15721.

King, L.B., 2016. Specialized Networks for Social Cognition: A Defining Role for the Oxytocin Receptor. J Neurosci 36, 8283-8285.

King, L.B., Walum, H., Inoue, K., Eyrich, N.W., Young, L.J., 2016. Variation in the Oxytocin Receptor Gene Predicts Brain Region-Specific Expression and Social Attachment. Biol Psychiatry 80, 160-169.

Klatt, J.D., Goodson, J.L., 2013. Oxytocin-like receptors mediate pair bonding in a socially monogamous songbird. Proc Biol Sci 280, 20122396.

Kosfeld, M., Heinrichs, M., Zak, P.J., Fischbacher, U., Fehr, E., 2005. Oxytocin increases trust in humans. Nature 435, 673-676.

Kraaijenvanger, E.J., He, Y., Spencer, H., Smith, A.K., Bos, P.A., Boks, M.P.M., 2019. Epigenetic variability in the human oxytocin receptor (OXTR) gene: A possible pathway from early life experiences to psychopathologies. Neurosci Biobehav Rev 96, 127-142.

Ku, C.Y., Qian, A., Wen, Y., Anwer, K., Sanborn, B.M., 1995. Oxytocin stimulates myometrial guanosine triphosphatase and phospholipase-C activities via coupling to G alpha q/11. Endocrinology 136, 1509-1515.

Leng, G., Ludwig, M., 2016. Intranasal Oxytocin: Myths and Delusions. Biol Psychiatry 79, 243-250.

Lerer, E., Levi, S., Salomon, S., Darvasi, A., Yirmiya, N., Ebstein, R.P., 2008. Association between the oxytocin receptor (OXTR) gene and autism: relationship to Vineland Adaptive Behavior Scales and cognition. Mol Psychiatry 13, 980-988.

Li, G., Fu, S., Wang, S., Zhu, C., Duan, B., Tang, C., Chen, X., Chuai, G., Wang, P., Liu, Q., 2022. A deep generative model for multi-view profiling of single-cell RNA-seq and ATAC-seq data. Genome Biol 23, 20.

Li, Z., Schulz, M.H., Look, T., Begemann, M., Zenke, M., Costa, I.G., 2019. Identification of transcription factor binding sites using ATAC-seq. Genome Biol 20, 45.

Lim, M.M., Wang, Z., Olazabal, D.E., Ren, X., Terwilliger, E.F., Young, L.J., 2004. Enhanced partner preference in a promiscuous species by manipulating the expression of a single gene. Nature 429, 754-757.

Liu, C., Wang, M., Wei, X., Wu, L., Xu, J., Dai, X., Xia, J., Cheng, M., Yuan, Y., Zhang, P., Li, J., Feng, T., Chen, A., Zhang, W., Chen, F., Shang, Z., Zhang, X., Peters, B.A., Liu, L., 2019. An ATAC-seq atlas of chromatin accessibility in mouse tissues. Sci Data 6, 65.

LoParo, D., Waldman, I.D., 2015. The oxytocin receptor gene (OXTR) is associated with autism spectrum disorder: a meta-analysis. Mol Psychiatry 20, 640-646.

Luo, P.X., Zakharenkov, H.C., Torres, L.Y., Rios, R.A., Gegenhuber, B., Black, A.M., Xu, C.K., Minie, V.A., Tran, A.M., Tollkuhn, J., Trainor, B.C., 2022. Oxytocin receptor behavioral effects and cell types in the bed nucleus of the stria terminalis. Horm Behav 143, 105203.

Maher, C.A., Palanisamy, N., Brenner, J.C., Cao, X., Kalyana-Sundaram, S., Luo, S., Khrebtukova, I., Barrette, T.R., Grasso, C., Yu, J., Lonigro, R.J., Schroth, G., Kumar-Sinha, C., Chinnaiyan, A.M., 2009. Chimeric transcript discovery by paired-end transcriptome sequencing. Proc Natl Acad Sci U S A 106, 12353-12358.

Manoli, D.S., Tollkuhn, J., 2018. Gene regulatory mechanisms underlying sex differences in brain development and psychiatric disease. Ann N Y Acad Sci 1420, 26-45.

Marazziti, D., Baroni, S., Mucci, F., Piccinni, A., Moroni, I., Giannaccini, G., Carmassi, C., Massimetti, E., Dell'Osso, L., 2019. Sex-Related Differences in Plasma Oxytocin Levels in Humans. Clin Pract Epidemiol Ment Health 15, 58-63.

Mardis, E.R., 2007. ChIP-seq: welcome to the new frontier. Nat Methods 4, 613-614.

Marlin, B.J., Mitre, M., D'Amour J, A., Chao, M.V., Froemke, R.C., 2015. Oxytocin enables maternal behaviour by balancing cortical inhibition. Nature 520, 499-504.

McGraw, L.A., Davis, J.K., Thomas, P.J., Program, N.C.S., Young, L.J., Thomas, J.W., 2012. BAC-based sequencing of behaviorally-relevant genes in the prairie vole. PLoS One 7, e29345.

Mercati, O., Huguet, G., Danckaert, A., Andre-Leroux, G., Maruani, A., Bellinzoni, M., Rolland, T., Gouder, L., Mathieu, A., Buratti, J., Amsellem, F., Benabou, M., Van-Gils, J., Beggiato, A., Konyukh, M., Bourgeois, J.P., Gazzellone, M.J., Yuen, R.K., Walker, S., Delepine, M.,

Boland, A., Regnault, B., Francois, M., Van Den Abbeele, T., Mosca-Boidron, A.L., Faivre, L., Shimoda, Y., Watanabe, K., Bonneau, D., Rastam, M., Leboyer, M., Scherer, S.W., Gillberg, C., Delorme, R., Cloez-Tayarani, I., Bourgeron, T., 2017. CNTN6 mutations are risk factors for abnormal auditory sensory perception in autism spectrum disorders. Mol Psychiatry 22, 625-633.

Mezger, A., Klemm, S., Mann, I., Brower, K., Mir, A., Bostick, M., Farmer, A., Fordyce, P., Linnarsson, S., Greenleaf, W., 2018. High-throughput chromatin accessibility profiling at single-cell resolution. Nat Commun 9, 3647.

Mitre, M., Kranz, T.M., Marlin, B.J., Schiavo, J.K., Erdjument-Bromage, H., Zhang, X., Minder, J., Neubert, T.A., Hackett, T.A., Chao, M.V., Froemke, R.C., 2017. Sex-Specific Differences in Oxytocin Receptor Expression and Function for Parental Behavior. Gend Genome 1, 142-166.

Mundade, R., Ozer, H.G., Wei, H., Prabhu, L., Lu, T., 2014. Role of ChIP-seq in the discovery of transcription factor binding sites, differential gene regulation mechanism, epigenetic marks and beyond. Cell Cycle 13, 2847-2852.

Narayanan, A., Blanco-Carmona, E., Demirdizen, E., Sun, X., Herold-Mende, C., Schlesner, M., Turcan, S., 2020. Nuclei Isolation from Fresh Frozen Brain Tumors for Single-Nucleus RNA-seq and ATAC-seq. J Vis Exp.

Nickerson, K., Bonsness, R.W., Douglas, R.G., Condliffe, P., Du Vigneaud, V., 1954. Oxytocin and milk ejection. Am J Obstet Gynecol 67, 1028-1034.

O'Neill, L.P., Turner, B.M., 1996. Immunoprecipitation of chromatin. Methods Enzymol 274, 189-197.

Okhovat, M., Berrio, A., Wallace, G., Ophir, A.G., Phelps, S.M., 2015. Sexual fidelity trade-offs promote regulatory variation in the prairie vole brain. Science 350, 1371-1374.

Olazabal, D.E., Alsina-Llanes, M., 2016. Are age and sex differences in brain oxytocin receptors related to maternal and infanticidal behavior in naive mice? Horm Behav 77, 132-140.

Olazabal, D.E., Young, L.J., 2006. Species and individual differences in juvenile female alloparental care are associated with oxytocin receptor density in the striatum and the lateral septum. Horm Behav 49, 681-687.

Oliver, G., Schafer, E.A., 1895. On the Physiological Action of Extracts of Pituitary Body and certain other Glandular Organs: Preliminary Communication. J Physiol 18, 277-279.

Osipovich, A.B., Gangula, R., Vianna, P.G., Magnuson, M.A., 2016. Setd5 is essential for mammalian development and the co-transcriptional regulation of histone acetylation. Development 143, 4595-4607.

Oumi, T., Ukena, K., Matsushima, O., Ikeda, T., Fujita, T., Minakata, H., Nomoto, K., 1996. Annetocin, an annelid oxytocin-related peptide, induces egg-laying behavior in the earthworm, Eisenia foetida. J Exp Zool 276, 151-156.

Penagarikano, O., Lazaro, M.T., Lu, X.H., Gordon, A., Dong, H., Lam, H.A., Peles, E., Maidment, N.T., Murphy, N.P., Yang, X.W., Golshani, P., Geschwind, D.H., 2015. Exogenous and evoked oxytocin restores social behavior in the Cntnap2 mouse model of autism. Sci Transl Med 7, 271ra278.

Peng, S., Bellone, R., Petersen, J.L., Kalbfleisch, T.S., Finno, C.J., 2021. Successful ATAC-Seq From Snap-Frozen Equine Tissues. Front Genet 12, 641788.

Pobbe, R.L., Pearson, B.L., Defensor, E.B., Bolivar, V.J., Young, W.S., 3rd, Lee, H.J., Blanchard, D.C., Blanchard, R.J., 2012. Oxytocin receptor knockout mice display deficits in the expression of autism-related behaviors. Horm Behav 61, 436-444.

Porsch, R.M., Middeldorp, C.M., Cherny, S.S., Krapohl, E., van Beijsterveldt, C.E., Loukola, A., Korhonen, T., Pulkkinen, L., Corley, R., Rhee, S., Kaprio, J., Rose, R.R., Hewitt, J.K., Sham, P., Plomin, R., Boomsma, D.I., Bartels, M., 2016. Longitudinal heritability of childhood aggression. Am J Med Genet B Neuropsychiatr Genet 171, 697-707.

Quintana, D.S., Rokicki, J., van der Meer, D., Alnaes, D., Kaufmann, T., Cordova-Palomera, A., Dieset, I., Andreassen, O.A., Westlye, L.T., 2019. Oxytocin pathway gene networks in the human brain. Nat Commun 10, 668.

Rault, J.L., Carter, C.S., Garner, J.P., Marchant-Forde, J.N., Richert, B.T., Lay, D.C., Jr., 2013. Repeated intranasal oxytocin administration in early life dysregulates the HPA axis and alters social behavior. Physiol Behav 112-113, 40-48.

Rizzardi, L.F., Hickey, P.F., Rodriguez DiBlasi, V., Tryggvadottir, R., Callahan, C.M., Idrizi, A., Hansen, K.D., Feinberg, A.P., 2019. Neuronal brain-region-specific DNA methylation and chromatin accessibility are associated with neuropsychiatric trait heritability. Nat Neurosci 22, 307-316.

Rocks, D., Jaric, I., Tesfa, L., Greally, J.M., Suzuki, M., Kundakovic, M., 2022. Cell type-specific chromatin accessibility analysis in the mouse and human brain. Epigenetics 17, 202-219.

Ross, H.E., Freeman, S.M., Spiegel, L.L., Ren, X., Terwilliger, E.F., Young, L.J., 2009. Variation in oxytocin receptor density in the nucleus accumbens has differential effects on affiliative behaviors in monogamous and polygamous voles. J Neurosci 29, 1312-1318.

Sanborn, B.M., Dodge, K., Monga, M., Qian, A., Wang, W., Yue, C., 1998. Molecular mechanisms regulating the effects of oxytocin on myometrial intracellular calcium. Adv Exp Med Biol 449, 277-286.

Saphire-Bernstein, S., Way, B.M., Kim, H.S., Sherman, D.K., Taylor, S.E., 2011. Oxytocin receptor gene (OXTR) is related to psychological resources. Proc Natl Acad Sci U S A 108, 15118-15122.

Schaschl, H., Huber, S., Schaefer, K., Windhager, S., Wallner, B., Fieder, M., 2015. Signatures of positive selection in the cis-regulatory sequences of the human oxytocin receptor (OXTR) and arginine vasopressin receptor 1a (AVPR1A) genes. BMC Evol Biol 15, 85.

Sessa, A., Fagnocchi, L., Mastrototaro, G., Massimino, L., Zaghi, M., Indrigo, M., Cattaneo, S., Martini, D., Gabellini, C., Pucci, C., Fasciani, A., Belli, R., Taverna, S., Andreazzoli, M., Zippo, A., Broccoli, V., 2019. SETD5 Regulates Chromatin Methylation State and Preserves Global Transcriptional Fidelity during Brain Development and Neuronal Wiring. Neuron 104, 271-289 e213.

Skuse, D.H., Lori, A., Cubells, J.F., Lee, I., Conneely, K.N., Puura, K., Lehtimaki, T., Binder, E.B., Young, L.J., 2014. Common polymorphism in the oxytocin receptor gene (OXTR) is associated with human social recognition skills. Proc Natl Acad Sci U S A 111, 1987-1992.

Smearman, E.L., Almli, L.M., Conneely, K.N., Brody, G.H., Sales, J.M., Bradley, B., Ressler, K.J., Smith, A.K., 2016. Oxytocin Receptor Genetic and Epigenetic Variations: Association With Child Abuse and Adult Psychiatric Symptoms. Child Dev 87, 122-134.

Spektor, R., Tippens, N.D., Mimoso, C.A., Soloway, P.D., 2019. methyl-ATAC-seq measures DNA methylation at accessible chromatin. Genome Res 29, 969-977.

Su, Y., Shin, J., Zhong, C., Wang, S., Roychowdhury, P., Lim, J., Kim, D., Ming, G.L., Song, H., 2017. Neuronal activity modifies the chromatin accessibility landscape in the adult brain. Nat Neurosci 20, 476-483.

Szczaluba, K., Brzezinska, M., Kot, J., Rydzanicz, M., Walczak, A., Stawinski, P., Werner, B., Ploski, R., 2016. SETD5 loss-of-function mutation as a likely cause of a familial syndromic intellectual disability with variable phenotypic expression. Am J Med Genet A 170, 2322-2327.

Tabak, B.A., McCullough, M.E., Carver, C.S., Pedersen, E.J., Cuccaro, M.L., 2014. Variation in oxytocin receptor gene (OXTR) polymorphisms is associated with emotional and behavioral reactions to betrayal. Soc Cogn Affect Neurosci 9, 810-816.

Tiklova, K., Bjorklund, A.K., Lahti, L., Fiorenzano, A., Nolbrant, S., Gillberg, L., Volakakis, N., Yokota, C., Hilscher, M.M., Hauling, T., Holmstrom, F., Joodmardi, E., Nilsson, M., Parmar, M., Perlmann, T., 2019. Single-cell RNA sequencing reveals midbrain dopamine neuron diversity emerging during mouse brain development. Nat Commun 10, 581.

Trabzuni, D., Ramasamy, A., Imran, S., Walker, R., Smith, C., Weale, M.E., Hardy, J., Ryten, M., North American Brain Expression, C., 2013. Widespread sex differences in gene expression and splicing in the adult human brain. Nat Commun 4, 2771.

Uzefovsky, F., Shalev, I., Israel, S., Edelman, S., Raz, Y., Mankuta, D., Knafo-Noam, A., Ebstein, R.P., 2015. Oxytocin receptor and vasopressin receptor 1a genes are respectively associated with emotional and cognitive empathy. Horm Behav 67, 60-65.

Van Kesteren, R.E., Smit, A.B., De Lange, R.P., Kits, K.S., Van Golen, F.A., Van Der Schors, R.C., De With, N.D., Burke, J.F., Geraerts, W.P., 1995. Structural and functional evolution of the vasopressin/oxytocin superfamily: vasopressin-related conopressin is the only member

present in Lymnaea, and is involved in the control of sexual behavior. J Neurosci 15, 5989-5998.

Verbeke, W., Bagozzi, R.P., van den Berg, W.E., Lemmens, A., 2013. Polymorphisms of the OXTR gene explain why sales professionals love to help customers. Front Behav Neurosci 7, 171.

Visscher, P.M., Brown, M.A., McCarthy, M.I., Yang, J., 2012. Five years of GWAS discovery. Am J Hum Genet 90, 7-24.

Wang, J., Zibetti, C., Shang, P., Sripathi, S.R., Zhang, P., Cano, M., Hoang, T., Xia, S., Ji, H., Merbs, S.L., Zack, D.J., Handa, J.T., Sinha, D., Blackshaw, S., Qian, J., 2018. ATAC-Seq analysis reveals a widespread decrease of chromatin accessibility in age-related macular degeneration. Nat Commun 9, 1364.

Warrier, V., Grasby, K.L., Uzefovsky, F., Toro, R., Smith, P., Chakrabarti, B., Khadake, J., Mawbey-Adamson, E., Litterman, N., Hottenga, J.J., Lubke, G., Boomsma, D.I., Martin, N.G., Hatemi, P.K., Medland, S.E., Hinds, D.A., Bourgeron, T., Baron-Cohen, S., 2018. Genome-wide meta-analysis of cognitive empathy: heritability, and correlates with sex, neuropsychiatric conditions and cognition. Mol Psychiatry 23, 1402-1409.

Wermter, A.K., Kamp-Becker, I., Hesse, P., Schulte-Korne, G., Strauch, K., Remschmidt, H., 2010. Evidence for the involvement of genetic variation in the oxytocin receptor gene (OXTR) in the etiology of autistic disorders on high-functioning level. Am J Med Genet B Neuropsychiatr Genet 153B, 629-639.

Wijmenga, C., Zhernakova, A., 2018. The importance of cohort studies in the post-GWAS era. Nat Genet 50, 322-328.

Williams, J.R., Catania, K.C., Carter, C.S., 1992. Development of partner preferences in female prairie voles (Microtus ochrogaster): the role of social and sexual experience. Horm Behav 26, 339-349.

Williams, J.R., Insel, T.R., Harbaugh, C.R., Carter, C.S., 1994. Oxytocin administered centrally facilitates formation of a partner preference in female prairie voles (Microtus ochrogaster). J Neuroendocrinol 6, 247-250.

Xu, J., Burgoyne, P.S., Arnold, A.P., 2002. Sex differences in sex chromosome gene expression in mouse brain. Hum Mol Genet 11, 1409-1419.

Xu, J., Deng, X., Watkins, R., Disteche, C.M., 2008. Sex-specific differences in expression of histone demethylases Utx and Uty in mouse brain and neurons. J Neurosci 28, 4521-4527.

Yang, J., Ferreira, T., Morris, A.P., Medland, S.E., Genetic Investigation of, A.T.C., Replication, D.I.G., Meta-analysis, C., Madden, P.A., Heath, A.C., Martin, N.G., Montgomery, G.W., Weedon, M.N., Loos, R.J., Frayling, T.M., McCarthy, M.I., Hirschhorn, J.N., Goddard, M.E., Visscher, P.M., 2012. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet 44, 369-375, S361-363.

Young, K.A., Liu, Y., Gobrogge, K.L., Wang, H., Wang, Z., 2014. Oxytocin reverses amphetamine-induced deficits in social bonding: evidence for an interaction with nucleus accumbens dopamine. J Neurosci 34, 8499-8506.

Young, L.J.H., B.; Nilsen, R.; Wang,Z.; Insel,T.R., 1996. Species Differences in Central Oxytocin Receptor Gene Expression: Comparative Analysis of Promoter Sequences. Neuroendocrinology 8, 777-783.

Young, R.L., Ferkin, M.H., Ockendon-Powell, N.F., Orr, V.N., Phelps, S.M., Pogany, A., Richards-Zawacki, C.L., Summers, K., Szekely, T., Trainor, B.C., Urrutia, A.O., Zachar, G., O'Connell, L.A., Hofmann, H.A., 2019. Conserved transcriptomic profiles underpin monogamy across vertebrates. Proc Natl Acad Sci U S A 116, 1331-1336.

Zheng, J.J., Li, S.J., Zhang, X.D., Miao, W.Y., Zhang, D., Yao, H., Yu, X., 2014. Oxytocin mediates early experience-dependent cross-modal plasticity in the sensory cortices. Nat Neurosci 17, 391-399.

Zou, Y., Wu, Z., Fan, Z., Liang, D., Wang, L., Song, Z., You, F., 2020. Analyses of mRNA-seq and miRNA-seq of the brain reveal the sex differences of gene expression and regulation before and during gonadal differentiation in 17beta-estradiol or 17alpha-methyltestosterone-induced olive flounder (Paralichthys olivaceus). Mol Reprod Dev 87, 78-90.