Engineering Microfluidic Systems for Low-Input Multi-Omic Profiling of Brain Regulatory States

Thomas Hadlock

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In

Chemical Engineering

Chang Lu, Chair

Steven Wrenn

Aaron Goldstein

William Ducker

February 12, 2026

Blacksburg, VA

Keywords: Microfluidics, Epigenome, Transcriptome, Chromatin Immunoprecipitation, Next generation sequencing, RNA sequencing, Bisulfite Sequencing, Viral Diagnostics, Nanopore Sequencing, Psychedelics

Engineering Microfluidic Systems for Low-Input Multi-Omic Profiling of Brain Regulatory States

Thomas Hadlock

Abstract

Microfluidic systems enable sensitive application of sequencing based “omic” assays to rare sample sets. In this thesis, we developed and applied these systems to address fundamental limitations in existing assays and to investigate biological questions that are inaccessible using traditional bulk protocols. First, we introduced a microfluidic platform for rapid field-ready library preparation of viral samples for nanopore sequencing. Our microreactor enabled laboratory independent viral diagnostics of field-collected Senecavalley virus A samples with accuracy in-line with gold standard RT-PCR assays. Additionally, our system enabled real-time mutation identification, detecting two high-confidence consensus single nucleotide variants within the SVA positive cohort. We then applied microfluidic MOWChIP platform to two epigenomic investigations into the effects of psilocybin on mouse cortical and subcortical regions. First, we examined the sex-specific enhancer alterations following psilocybin exposure in neurons extracted from the frontal cortex and nucleus accumbens to elucidate therapeutic mechanisms for combating opioid use disorder. Here we describe localized increases in enhancer activity in the mesolimbic nucleus accumbens compared to the frontal cortex. Additionally, we demonstrate significant recovery of key enhancer linked gene pathways disrupted by prolonged opioid use following acute psilocybin exposure

in male mice absent in female cohort. Finally, we investigated epigenetic inheritance of prolonged psilocybin exposure in prenatal dams on their offspring. We uncover significant transcription factor regulatory network disruptions that persist through to the F1 generation. Specifically, reduction in regulatory efficacy of *Egr2* and *JunB* transcription factors of which direct psilocybin exposure mediates significant expressional increases in the prefrontal cortex. We further uncover a sex specific nature of these alterations, in which minimized reach of early gene *Egr2* is found primarily in female offspring following maternal exposure. These multi-omic investigations uncover significant sex-specific implications of psilocybin exposure on brain epigenome.

Engineering Microfluidic Systems for Low-Input Multi-Omic Profiling of Brain Regulatory States

Thomas Hadlock

General audience abstract

Microfludic devices allow for miniaturization of traditional laboratory techniques into closed, automatable systems that enable downscaling of initial input requirements. Among its various uses, the compact single pot nature of microfluidic systems removes requirements around large, immobile equipment anchoring sample handing and assay protocols from traditional laboratory environments. As a demonstration of these benefits, we begin this thesis with the development of a field-deployable microfluidic platform that enables preparation of nanopore sequencing libraries from viral samples. This portable system enables rapid viral pathogen diagnostics and variant monitoring at the point-of-care. We demonstrated the efficacy of this platform using the swine pathogen Senecavalley virus A, from which our system successfully diagnosed all 7 of our patient samples and 3 healthy controls. Additionally, our microreactor platform distinguished itself from gold-standard RT-PCR assays by identifying 2 single nucleotide variations within our patient cohort while simultaneously generating diagnostic calls. We then turned our attention to application of microfluidic devices to biological investigations of rare cell types within the mouse brain affected by psychedelic exposure made possible by their significantly reduced input requirements. Previous research on psychedelic compounds has demonstrated unique abilities to alter cell-signaling within the brain.

These studies have shown psychedelic exposure leads to increased synaptic plasticity and dendritic density leading to gains in cognitive function. These neurogenic mechanisms may underly observed therapeutic qualities in treating myriad psychological disorders including PTSD, as well as driving positive behavioral changes associated with addiction and drug seeking behavior. Previous work in our lab has shown that these gains are accompanied by long-lasting epigenomic alterations in brain regions responsible for high-order executive function. We thus carried out two studies to better understand the epigenetic alterations associated with psychedelic exposure within these critical brain regions. First, we performed the first ever sex-specific investigation into the epigenomic changes associated with psilocybin exposure and its associated recovery in opioid addicted mice, probing brain regions driving drug seeking behavior. Our study found that psilocybin effectively recovered significant epigenomic alterations induced by prolonged opioid use, with observed recovery far more prevalent in males than in females. We then performed a study into epigenomic alterations inherited in offspring of females following prolonged psilocybin exposure to observe inherited marks escaping epigenetic reprograming. This study showed that maternal psilocybin exposure minimizes the regulatory impact of several key transcription factors driving the positive neurogenic properties of primary psychedelic exposure. Thus, downstream effects of increased expression of these key transcription factors are reduced in offspring. We also found these regulatory network minimizations to be localized within female offspring populations, potentially decreasing the efficacy of psychedelic therapeutics within these mice in the future.

Acknowledgements

I would like to thank my advisor, co-workers, collaborators, family and friends. The work presented here would not have been possible without them. They have all played key roles in my professional and personal growth over the past five years. To the success I have had here at my time at Virginia Tech, and all future success I may find off the back of this degree, all of it was made possible by the doors they opened.

I would like to start by thanking my advisor and head of my committee Dr. Chang Lu for everything he has done for me during my degree. He has been a key driver of development in my research capabilities. He has provided ample trust and space when necessary yet is always there with sturdy guidance and motivation when necessary. I appreciate the knowledge he has imparted to me, and I really appreciated the great conversation, both professional and personal, over the years.

I would like to thank my committee members, Dr. Steven Wrenn, Dr. William Ducker, and Dr. Aaron Goldstein, for their guidance and support during my preliminary examination. This discussion helped guide my final projects and approach future work in a new, more inquisitive light. (Placeholder)

I would like to thank my collaborators, Dr. Javier Gonzalez-Maeso, Dr. Alaina Jaster, Dr. Song Li, and Dr. Sunil Mor. Their knowledge of animal models and tireless sourcing of interesting samples enabled these projects to reach their full potential.

I would like to thank my colleagues and friends in the lab who have supported me over the years: Dr. Bohan Zhu, Dr. Qiang Zhang, Dr. Jerry (ZhengZhi) Liu, Dr. Zirui Zhou, Gaoshan Li, Jacob Niece, Xin Zhang, Jenna Catalano, and Rui Han. They are a group of bright, inquisitive and thoughtful scientists whose support has been invaluable. Their insightful feedback and advice have helped shaped my professional growth, while their lively conversations and camaraderie have made even the longest days enjoyable.

I would like to thank my family and friends for their tireless support through this journey. I am especially grateful to my parents John and Amy, and my brother, Mark, who have always been there to encourage, motivate, and support me in pursuing my goals in life. I am also deeply thankful for the friends who have provided constant encouragement, perspective and humor during both the challenging and rewarding moments along the way.

Most importantly, I would like to thank my wife, Jessie, for her unwavering love, patience, and belief in me. Her constant support, understanding and encouragement have been a source of strength throughout this process, and I could not have reached this milestone without her by my side.

Table of contents

List of abbreviations

ATAC-seq	Assay for transposase-accessible chromatin coupled with sequencing
ChIP-seq	Chromatin immunoprecipitation coupled with sequencing
CNS	Central nervous system
CUT&RUN	Cleavage under targets and release using nuclease
DEG	Differentially expressed gene
DMR	Differentially methylated region
DNA	Deoxyribonucleic acid
DOI	1-(2,5-dimethoxy-4-iodophenyl)-2-aminopropane
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EGR1	Early growth response factor 1
EGR2	Early growth response factor 2
EGTA	Ethylene glycol-bis(β-aminoethyl ether)-N,N,N′,N′-tetraacetic acid
ENCODE	Encyclopedia of DNA Elements
FACS	Fluorescence-activated cell sorting
FRiP	Fraction of reads in peaks
GO	Gene ontology
HC	Healthy control
JunB	Jun B proto-oncogene, AP-1 transcription factor subunit
KEGG	Kyoto encyclopedia of genes and genomes
LAMP	Loop mediated isothermal amplification

LSD	lysergic acid diethylamide
MNase	Micrococcal nuclease
MOR	μ-opioid receptor
MOWChIP	Microfluidic oscillatory washing–based chromatin immunoprecipitation
NAc	Nucleus accumbens
OUD	Opioid use disorder
OXY	Oxycodone
PCR	Polymerase chain reaction
PDMS	Polydimethylsiloxane
PFC	Prefrontal cortex
PIC	Protein inhibitor cocktail
PMSF	Phenylmethylsulfonyl fluoride
PPR	Personalized page rank
PSI	Psilocybin
qPCR	Quantitative polymerase chain reaction
RdRp	RNA-dependent RNA polymerases
RFU	Relative fluorescence unit
RNA	Ribonucleic acid
RREB1	Ras responsive element binding protein 1
SDS	Sodium dodecyl sulfate
SNV	Single nucleotide variation
SUD	Substance use disorder
SVA	Senecavalley virus A

TE	Tris-EDTA
Tris	Tris(hydroxymethyl)aminomethane
TSS	Transcription start site
TF	Transcription factor
UHC	Unsupervised hierarchical clustering
UMAP	Uniform manifold approximation and projection
VEH	Vehicle
VFE	Vesicular lesion fluid extract
VTA	Ventral tegmental area

1. Project Summary

Nucleic acid sequencing has expanded the scope of biological inquiry by enabling high-resolution analyses of molecular processes. Beginning with the widespread adoption of Sanger platform in the late 1970's and accelerated by cost reductions and throughput increases associated with next-generation Illumina systems, and now nanopore sequencing, a diverse array of sequencing based *omic* assays has become accessible for studying gene expression and regulatory networks. As technological improvements have expanded the capabilities of sequencing, they have also shifted limitations toward upstream stages of the workflow, revealing new bottlenecks in sample handling and assay input requirements. In this study, we demonstrate microfluidic systems that mitigate critical bottlenecks in sequencing assay workflows and demonstrate their utility in supporting highly sensitive biological investigations of rare sample populations.

To demonstrate real world application of the current cutting edge of sequencing technology, we begin with a look at a microreactor platform that enables real-time diagnostics and variant monitoring by taking advantage of the speed, portability, and cost effectiveness third generation nanopore sequencers. The global experience of the COVID-19 pandemic has brought to the forefront the need for rapid viral diagnostics and variant tracking, even in locations that lack traditionally laboratory setups. While on the surface it may appear that laptop powered nanopore sequencers can perform this role, the nucleic acid handling requirements to generate the readable input, or library, necessitate immobile lab-based equipment. In this project, we designed a microreactor platform that generates libraries from extracted viral RNA, unlocking the true portability of these third-generation devices. To minimize equipment and footprint, the platform

uses a built-in microscope stage heater to isothermally generate amplicons at moderate temperature using a loop mediated isothermal amplification system. This system was expanded to incorporate 6 distinct primer groups to cover up to 14% of the target genome, a significant expansion on the 150-200bp typically targeted during LAMP. This amplification system was paired with a tagmentation based library preparation approach to create the rapid, portable platform.

To validate our system, we first utilized cultured Senecavalley Virus A (SVA) viral RNA collected by our collaborator Dr. Sunil Mor. The design and operational protocol were optimized to reliably and robustly generate sequence ready samples. Thus, we further tested our platform with real-world collected vesicular lesion fluid samples from SVA positive pigs. Our device provided RT-PCR matched positive diagnostics amongst all samples within 27 minutes. Additionally, our system identified 2 high-confidence consensus level single nucleotide variants (SNVs) within our population. By simply changing primer sets, this platform can be applied to target any LAMP ready viral pathogen, providing a critical tool for combating future pandemic outbreaks.

While this initial project yielded valuable experience in sequencing technology development, this application was somewhat peripheral to the central biological focus of our lab’s work: epigenetics. Epigenetics involves changes in gene expression that do not alter the underlying DNA sequence. These reversible modifications are often triggered by environmental factors and play a critical role in a myriad of conditions, including neurological disorders, cancer, developmental disorders, and addiction. These modifications can be inherited across generations, highlighting the complexity of epigenetic inheritance, as germline alterations can be passed down both directly and

indirectly. Notably, epigenomic alterations have been shown to affect males and females differently, suggesting that these modifications may play a role in sex-specific disease susceptibilities and responses to treatments. Despite the absence of changes to the DNA itself, these epigenetic modifications carry significant implications for understanding heredity and the development of various diseases.

To investigate the mechanisms underlying epigenetic changes and their role in a wide range of diseases, researchers have employed several advanced methods. Among these, genomic sequencing technology has emerged as a particularly powerful tool for acquiring high-quality information on epigenetic modifications. Techniques such as bisulfite sequencing, chromatin immunoprecipitation, assay for transposase-accessible chromatin (ATAC), and methylated DNA immunoprecipitation have been utilized, tailored to the specific modifications being studied. These assays can also be paired with RNA sequencing protocols in a multi-omic manner to generate direct links between the observed gene expression and epigenomic modifications for a more comprehensive understanding of how the epigenome regulates the transcriptome.

Prior to the work presented here, our lab has developed several important microfluidic technologies for probing epigenomic modifications. These include low-input bulk strategies, such as the SurfaceChIP, as well as single-cell tools such as the Drop-BS chip for individual cell investigations of DNA methylation state. For the purposes of this work, we utilized another in-house developed low-input platform, the MOWChIP, for high-quality chromatin immunoprecipitation followed by sequencing identify key state-specific histone modifications. Specifically, this device was used to probe histone acetylation of the H3K27 lysine residue. This epigenetic mark, generally referred to as

an “active enhancer” typically leads to an open chromatin state through charge neutralization, thus loosening the DNA surrounding the histone allowing increased access to transcriptional machinery.

In this work, we began by applying MOWChIP platform to an evaluation of epigenetic alterations following psilocybin exposure in the context of a therapeutic for opioid use disorder. The low-input nature of this device enabled ChIP-seq of neuron populations extracted from slices of prefrontal cortex (PFC) and nucleus accumbens (NAc). These regions were chosen due to overrepresentation of 5-HT2A serotonin receptors and higher order executive function (PFC), as well as critical roles in reward and drug seeking behavior (NAc). High quality ChIP-seq datasets were generated, and differential analysis was performed comparing drug and therapeutic states. Concurrently, our collaborator Dr. Javier Maeso-Gonzalez performed behavioral studies of the oxycodone addicted and psilocybin treated mice, with saline injected mice as a control, and find disparities in therapeutic efficacy between male and female cohorts. With this knowledge, along with extensive literature highlighting sex-specific epigenomic alterations, we expanded our ChIP-seq study to compare male and female active enhancer binding. Our findings show significant increase in differential enhancer activity in the NAc compared to the PFC amongst our opioid use model and opioid use model followed by psilocybin treatment, regardless of sex. The PFC did show significant sex differences, however, with males demonstrating larger recovery of oxycodone altered enhancer linked genes to baseline levels following psilocybin exposure. This was in line with our behavioral findings, which demonstrated increased efficacy of psilocybin in reducing drug-seeking behavior in males compared to females.

Building off this study, along with previous findings in the lab demonstrating psychedelic induced alterations are not just acute in nature, we began another investigation into inherited gene regulatory alterations following prolonged psilocybin exposure in prenatal dams. Previous studies have shown parental drug use has deleterious heritable implications in offspring that cannot be fully explained genetically. Yet to this point no such study has been performed examining the implications involving psychedelics, and more specifically, psilocybin. In anticipation of lower volume of inherited epigenomic alterations compared to those induced by primary exposure, we expanded the scope of our investigation in a multi-omic manner, performing whole genome bisulfite sequencing (WGBS) and RNA-seq along with H3K27ac MOWChIP to profile multiple epigenomic marks and link them to observed gene expression changes.

Utilizing mice provided from Dr. Maeso-Gonzalez' lab, we performed all three assays on neurons extracted from frontal cortex tissue. While our initial findings yielded observed epigenomic and transcriptomic changes, we found little overlap between differentially expressed genes and our annotated histone marks. This drove us to utilize more robust multi-modal analysis methods for quantifying less obvious connections between our epigenomic data and gene expression. Utilizing Taiji analysis, we linked our MOWChIP data to nearby transcription factors, and identified alterations in downstream regulatory networks of several TFs within our dataset. We identified 11 transcription factors with significant changes to their regulatory networks, including two (Egr2 and JunB) heavily altered by primary psilocybin exposure. We also identified 114 significantly weighted genes within the differentially altered networks, 113 of which have decreased role of Egr2 in regulating their expression following maternal psilocybin exposure. This

suggests maternal psilocybin exposure may modulate downstream Egr2 mediated effects induced by primary treatment in offspring.

We then combined our WGBS data with our TF regulatory set to locate regions co-regulated by multiple epigenetic marks. We identified 13 genes with altered transcription factor regulatory activity that also displayed significant differentially methylated regions. These co-regulated genes were distinctive in their involvement in pathways observed in morphine addiction and oxytocin signaling, suggesting inherited alterations related to those recovered from primary exposure in our initial study. One of these genes which displayed decreased importance of Egr2 TF in its regulation, Pde4d, is a canonical imprinted gene region and contains a significant DMR in its promotor region. The hypermethylation at this region can inhibit protein binding and may be responsible for the decreased regulatory importance of Egr2 at this location. Additionally, as this is a canonical imprinted gene, it suggests a pathway of signal transmission by circumventing demethylation of inherited DNA methylation marks during epigenomic reprograming.

Finally, we completed this project by examining sex-specific inherited epigenomic alterations induced by maternal psilocybin exposure. We identified 5 male and 7 female specific transcription factors with significantly altered regulatory roles. The reduced role of Egr2 TF regulatory activity is localized amongst female cohorts. Interestingly, females also display an increase in another early growth family transcription factor, Egr1, which is not seen in males. However, the prevalence and magnitude of gene weights linked to these transcription factors suggests Egr2 TF activity is the most impactful inherited alteration. The diminished activity of Egr2 may inhibit females from eliciting the neurogenic benefits of primary psilocybin exposure brought about by binding of the 5-

HT2A receptor in cortical neurons. This project is ongoing, and we look forward to confirmation of this hypothesis by future examination of gene expression and behavioral studies of maternally treated offspring given primary psilocybin exposure.

2. A microreactor system for point-of-care viral genome sequencing

2.1 Introduction

Outbreaks of vesicle causing diseases in swine populations can wreak havoc on American agriculture[1-3]. As the 5 known diseases that generate vesicles in swine vary in prognostic severity, rapid identification of the of virus at hand is critical for determining proper interventions and prevent disastrous outbreak. While vesicular diseases such as Foot and Mouth disease (FMD) have been largely eradicated from the United States[4], others such Senecavalley Virus A (SVA), have seen repeated outbreaks both domestic and abroad over the last decade[5-9]. This non-enveloped, single-stranded, positive-sense RNA picornavirus transmits rapidly through populations, leading to formation of ulcerative lesions around the coronary band and fluid filled vesicles around the snout and oral mucosa[10].

As the vesicular lesions formed by the presence of SVA are virtually indistinguishable from those caused by other vesicular diseases, molecular diagnostics have become a critical tool for resolving stain identification[11]. Among molecular diagnostic tools, RT-PCR has been the gold standard of viral testing for decades due to its high sensitivity[12]. Several successful RT-PCR based assays have been reported for SVA detection in pigs[13-16]. However, this method is not without its drawbacks, such as the general requirement of expensive, bulky thermocyclers and lack of individual nucleotide level information provided[17]. Alternative sequence specific amplification methods that operate at isothermal conditions have been introduced, with Loop Mediated Isothermal Amplification (LAMP) by Notomi and colleagues amongst the most popular[18-20]. This method utilizes sets of 4-6 primers targeting short (150-250bp) pre-specified regions to

form long, repeated looped amplicon structures. Amplification is accomplished at moderate temperatures (~65°C) with its isothermal nature allowing for operation with just a simple hotplate. However, this fluorescent monitoring of LAMP amplicons in real time still lacks information provided at the single nucleotide level, limiting its capacity to monitor genomic alterations. Additionally, LAMP amplification for viral diagnostics has been hindered by primer dimer formation leading to increased false positive rates[21, 22].

Genome sequencing remains the most robust molecular diagnostic tool, providing nucleotide level resolution, but has traditionally been limited to laboratory settings due to cost and size of Sanger and NGS sequencers[23, 24]. This has changed in recent years with the emergence nanopore sequencers, such as those offered by Oxford Nanopore. These small, field deployable sequencers can be operated anywhere with just a laptop, and at significant upfront cost reduction[25, 26]. However, as sequencing technology and cost have improved, library preparation has emerged as a procedural bottleneck[27, 28]. In most cases, sequencing of low yield sample quantities still requires PCR amplification to generate enough library for accurate sequencing, with necessary equipment tethering nanopore sequencing back to the laboratory setting[24]. Additionally, the multiple steps required during library preparation must be performed carefully by a trained operator using tubes and pipettes, with ample opportunity for contamination or error[29, 30].

Microfluidic devices, which allow for automated operation of multiple laboratory processes in a single closed miniaturized container, significantly reduce contamination and human error at the point of application[27]. Microfluidic devices have previously shown success at generating NGS libraries, typically through ligation-based approaches, but on-chip nanopore library preparation has not received the same degree

of attention[27, 29-32]. An alternative, simpler library preparation approach that could be more easily conducted on-chip is available in the form of TN5 transposon mediated tagmentation, but this method requires significantly higher input quantities as well as longer fragment DNA to function properly. LAMP amplification could help address these two requirements, considering the longer fragment size of its looped amplicons. LAMP amplification on-chip has also been explored, but typically for assays with rapid diagnostic readouts where generated amplicons are not suitable for additional downstream applications[33-37].

In this study, we develop a field deployable microfluidic device for generation of nanopore sequencing libraries utilizing LAMP amplification for rapid SVA diagnostics and variant identification. Our system provides automation that leads to a reduction in manual labor requirements as well as breaks away from traditional need for expensive thermocyclers to generate amplified, sequencing-ready libraries.

2.2 Results/Discussion

Increased specificity of RT-LAMP is driven by its 4-6 primer sets in uses for amplification, targeting 6-8 regions of the genome (Fig 2.1A). However, due to mutations in the SVA genome potentially nullifying primer sets, along with the generally small, targeted regions, we sought to use multiple primer sets within the same assay. LAMP primers were designed to target 3 distinct regions within the SVA genome: 5’ UTR, VP2, and 3DPol. These regions were specifically targeted due to their differing degrees of conservation, with the VP2 and 3DPol protein coding regions in particular

offering moderate variance allowing us to display variant monitoring potential while maintaining diagnostic capabilities. The distance between these regions was also a consideration, as dispersed regions aides an even amplicon generation. Additionally, primer groupings that contained regional overlaps were explored to lengthen LAMP amplicons and provide even further genome coverage. Several dozen primer groups were designed using primer explore (V4) and tested for amplification time to ensure simultaneous amplification (and thus non-cannibalization of reagents). Of the primers that showed similar amplification time, 5 total groups allowed for successful amplification with relatively even single-pot dispersion (fig 2.1b), with ideal mixture ratios also established. Finalized LAMP primer sets are shown in table 1, with the mixture comprised of two overlapping groups in the 5'UTR, 3 overlapping groups in the 3DPol region, and a single primer group in the VP2 region, covering ~14% of the total SVA genome.

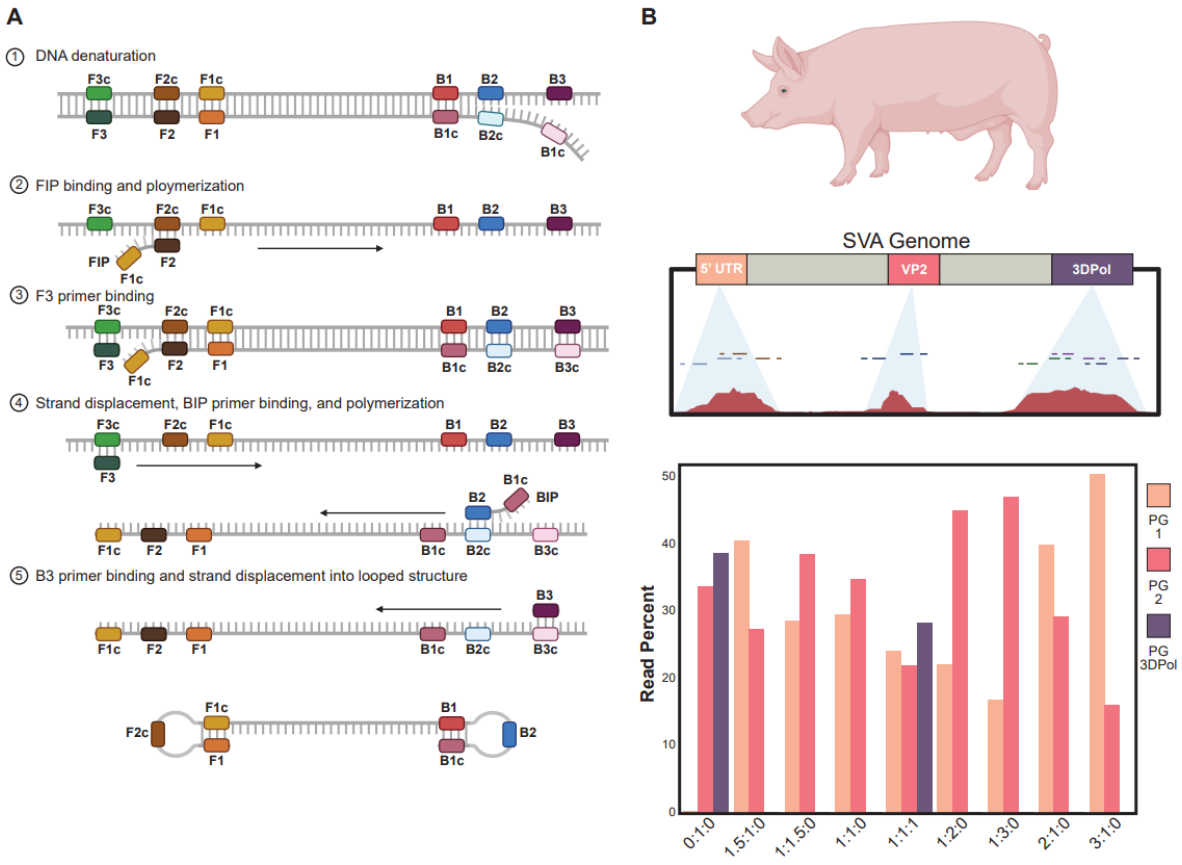

Fig 2.1. (A) Schematic overview of loop-mediated isothermal amplification (LAMP) amplicon generation. (B) Off-chip comparison of LAMP primer mix ratios targeting 3 distinct regions of SVA genome.

Here, we demonstrate the operation of our microfluidic chip to enact LAMP amplification and library preparation on target viral RNA. The device is setup as depicted (Fig 2.2A). As expansion of trapped air within the microfluidic chamber can occur during amplification heating, it is imperative to remove all gas from the chamber before operation. To ensure proper gas purging, fabricated microfluidic devices undergo an additional desiccation step and plasma treatment to increase hydrophilic interactions between applied solution and the inner chamber. A single syringe containing 75%

ethanol attached to PFA tubing is inserted into inlet b as shown (Fig 2.2A). 75% Ethanol is pumped into the device to remove all air from the chamber. A second syringe attached to PFA tubing containing degassed ultrapure water is then injected through (inlet a) to remove all 75% EtOH. Following 300ul of H2O wash, U-shaped tubing junction is placed in the outlet. The remaining inlets are then plugged by PFA tubing attached to syringes as shown (Fig 2.2A).

To begin processing a sample, the target sample is mixed with a LAMP mastermix and primer solution and slowly injected by syringe pump (Fig 2.2B step 2). The device is then moved over to the hotplate, set at 65°C for LAMP amplification for 1 hour (Fig 2.2B step 3). Following this amplification period, the device is moved back over to the handling platform and allowed to cool to room temperature. Once cooled, SPRI paramagnetic beads are injected slowly into both side inlets, with magnetic bars under the device sporadically applied to induce more thorough mixing. After a short (5 minute) incubation, the magnetic bars are applied to the top chamber, to collect all beads into a single mass before washing with 75% EtOH (Fig 2.2B step 5). The beads are then pulled against channeling of inlet 6, followed by slow injection of water from both side inlets and 75% ethanol inlet, to remove residual ethanol, using upstream gap bubble as a marker for progression. As shown in Fig 2.2B step 6, fragmentation solution is then injected slowly directly into the mass of collected beads. The underneath magnetic bar is removed, and the beads are permitted to flow along with the injected fragmentation solution towards the outlet. As the initial beads approach the outlet, the U-shaped PFA tubing is again placed into the outlet for collection of the fragmented library. The underneath magnet is applied one final time to retain the paramagnetic beads from the

solution as it is pumped into a collection tube. 1ul of sequencing adapter is then inputted into this collected library, although to avoid additional pipette stages, operators can instead opt to have the sequencing adaptor already waiting in the tube, as long as the tubing remains frozen until sample collection. For cost saving purposes, sequencing adaptor is not applied on the device, as efficient usage of our sequencing run for our purposes requires several (12) barcoded libraries.

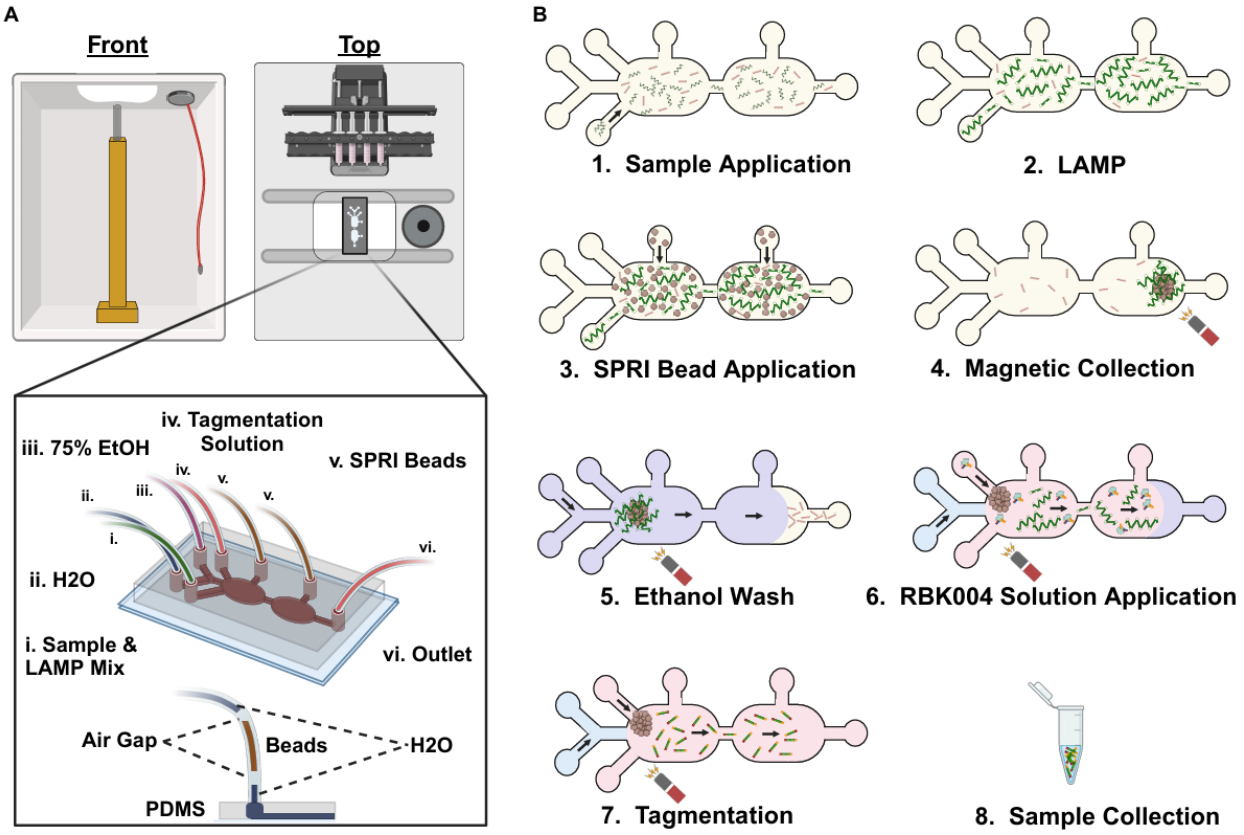

Fig 2.2. (A) Overview of microfluidic library preparation device. Basic Workflow (B): 1) Bubble removal and chamber prep 2) Sample application 3) LAMP 4) Sample collection

by paramagnetic beads 5) Bead clean-up with 75% ethanol 6) Library tagmentation and elution

Our system was first applied to cultured SVA viral RNA extracts to examine feasibility. LAMP amplification was examined on-chip at various process conditions (Fig 2.6B&D), with optimal operation established at 65°C (as confirmed by internalized probe) for 1 hour. Comparison of amplicon length by tapestation confirmed characteristic LAMP amplicon repeat structure formation matched that seen off chip (Fig 2.6F). Extracted SVA RNA was then subjected to amplification and library preparation protocols both on- and off-chip. Although reduction in aligned reads and alignment rate is observable in (Fig 2.3A&B) when compared to off-chip process, on-chip viral libraries generated 154.8x more regional alignments and a 616.2-fold increase in alignment rate over the similarly produced processed controls. Interestingly, on-chip controls produced libraries with roughly negligible regionally aligned reads and alignment rate compared to off-chip controls, which yielded 2.6x more aligned reads at 5.6% alignment.

It appears slight disparities in the on- vs off-chip generated libraries may arise during the amplification stage. While characteristic LAMP structures are generated in adequate quantity, sequencing of libraries shows the on-chip process yields libraries with smaller fragment sizes with an 8.9% decrease on average (Fig 2.3C). Additionally, on-chip libraries produced sequenced reads of slightly lower per base quality (Phred Q-score) with 16.7% of reads falling below a Q-score of 9, as opposed to off-chip generated libraries where less than 4% of reads were of similar quality (Fig 2.3D). However, despite the slight drop-off in quality, on-chip generated libraries displayed ample reads

at all targeted regions with relatively even distribution (Fig 2.3E&G). Libraries produced with the device provided high average identity with reference genome (all samples >95%), as was the case with those generated off-chip. With several hundred 1st hour reads per region, our system provides ample information for multiple downstream diagnostics and mutation identification.

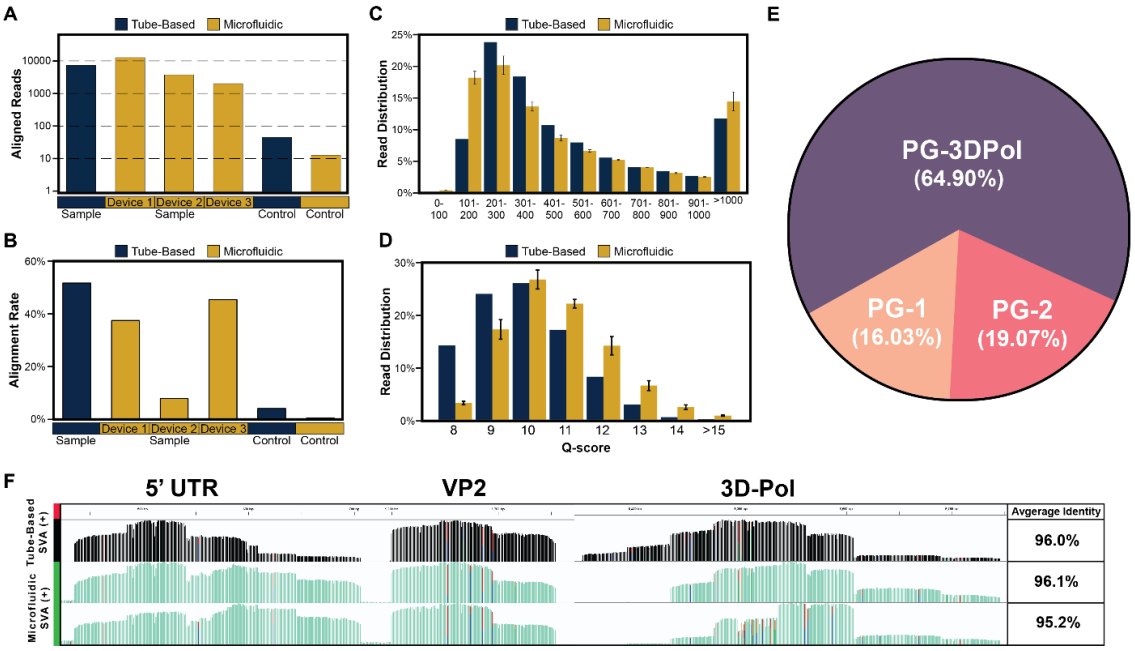

Fig 2.3. Comparison of generated nanopore sequencing libraries from cultured SVA viral extracts. Comparison of on-chip vs manual library preparations by (A) aligned reads, (B) alignment rate, (C) sequenced read length, (D) basecalling Phred Qscore. (E) On-chip library alignment dispersion amongst targeted regions. (F) Aligned sequencing tracks overlayed by integrative genomics viewer

We next sought to test our system using complex sample types it would encounter during field operation. Oral vesicular lesion fluid samples were extracted from SVA

positive pigs with saliva from healthy pigs taken as control. Samples determined SVA (+) by RT-PCR were confirmed with RT-LAMP in house. We first sought to establish the diagnostic capabilities of our field deployable library preparation system. Results from within the first hour (Fig 2.4A) showed the 7 SVA-positive samples had an average target region aligned read count of 1064, a 32.6x increase over that seen in healthy controls (avg 31.7 aligned reads). Individual region examination showed all positive samples had at least 2 regions with minimum 104 aligned reads within the first hour of sequencing. Of the controls, only a single region of one sample yielded more than 17 regionally aligned reads (PG1 of HC-VFE-15, 44 reads). Fig 2.4B shows maximum alignment rate of healthy controls was 58% lower (4.7%, 2% average alignment) than the lowest observed alignment rate from the positive samples (average 19.7%). Thus, first hour diagnostic cutoffs of >50 aligned reads within at least 2 target regions along with an alignment rate >5% were established for a positive sample determination.

Sequencing time requirements for positive diagnosis was compared to real- time fluorescence RT-LAMP using qPCR. Fig 4C shows the time requirement to diagnostic call for the two systems, with a score of 100 meaning the sample has reached the system-specific threshold for positive diagnostic call. Positive SVA samples yielded average Ct values of 52 minutes during real-time RT-LAMP using our designed primer sets, 51.6% longer than average time requirement (34.3 minutes) to meet our established LAMP sequencing threshold.

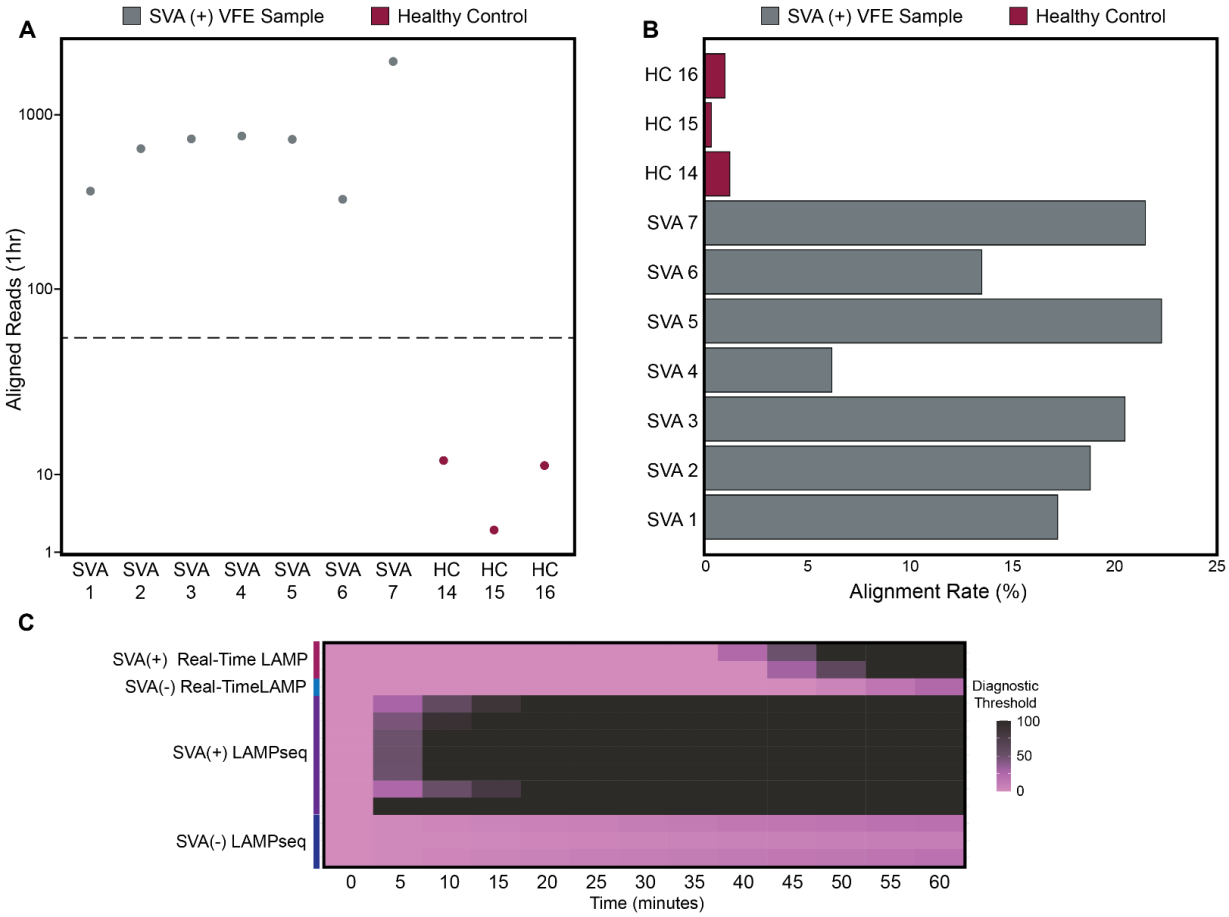

Fig 2.4. Diagnostic performance of microfluidic library preparation of SVA(+)-VFE samples including the first hour (A) read alignment and (B) alignment rate to SVA genome (NC_011349.1). (C) Time point diagnostic determination comparison of RT-LAMP amplicon sequencing compared to in-house fluorescent RT-LAMP assay.

Sequencing alignment tracks for libraries generated from VFE samples were then overlayed on IGV to more closely investigate library quality (Fig. 2.5A). Despite amplification of all targeted regions, primer set 2 of PG1 targeting the 5' UTR failed to amplify in any sample produced off chip and only appeared in 14.2% of libraries produced with the device. However, successfully amplified regions comprised >10% of

the SVA genome in libraries produced with and without the device. On-chip SVA (+) VFE libraries showed high degree of agreement with target genome (95.8% average identity) in line with those generated in tubes (95.6% average identity). When accounting for alignment artifact correction, both on SVA (+) libraries produced with both methods displayed high degree of consistency, with high average accuracy measurements (both groups >91%) within 0.1% of each other. Tested healthy controls showed significantly diminished average alignment accuracy in comparison (86%). Thus, despite generating near negligible quantities of alienable amplicons, those generated by healthy controls are still distinguishable by their notable read mismatch from the SVA genome, while positive samples show strong agreement and consistency regardless of method of library preparation.

To harness the full capabilities of sequencing, we sought to ensure our system can be used to rapidly identify sequencing variations that arise within populations of interest. Regions of significant variance from the reference genome were identified in samples containing at least 100x coverage of the target region. Specific base divergence from the genome composing >75% of the aligned reads was taken as the threshold to identify SNVs. Overall, the VP2 region contained the most observed variance from the reference, with two identified highly altered base locations. A single polymorphic SNP observed in 2 samples was identified at location 1674, while the other location (1623) did not contain enough consensus variance to meet our SNV distinction despite multiple samples displaying >50% differentiation. An additional singleton SNV was observed in the 3DPol region (6460), which interestingly showed a high degree of agreement (89.3%) with reference amongst all other tested samples. Previous investigation of SVA

mutation prevalence shows identified SNVs display a wide array of nucleotide compositions, with allelic frequencies ranging from dispersed to fully homogenous[38]. The sensitivity of our system provides accurate information to distinguish SNPs regardless of prevalence. This is confirmed by our Qscore average (Q>9) which indicates generated data is not distorted by high levels of basecalling error. The 5'UTR region tested did not contain any locations containing a high degree of variation.

RNA viruses are more likely to spawn genetic mutations compared to their DNA counterparts owing to lack of proofreading capabilities in RNA-dependent RNA polymerases (RdRp)[39, 40]. This is especially true in picornaviruses, which display some of the highest mutation rates amongst RNA viruses, and SVA is no exception. SVA has shown a tremendous mutation potential, even amongst narrow populations, with variants playing a direct role in virulence. Taken together, this highlights the critical importance of variant monitoring of SVA outbreaks. Our system has shown its capacity to generate sequence ready amplicons with alignment identity in line with previously reported interpopulation variance[41]. With its ability to accurately display single nucleotide resolution at multiple regions, our system stands out as a valuable tool for tracking viral progression from the frontlines.

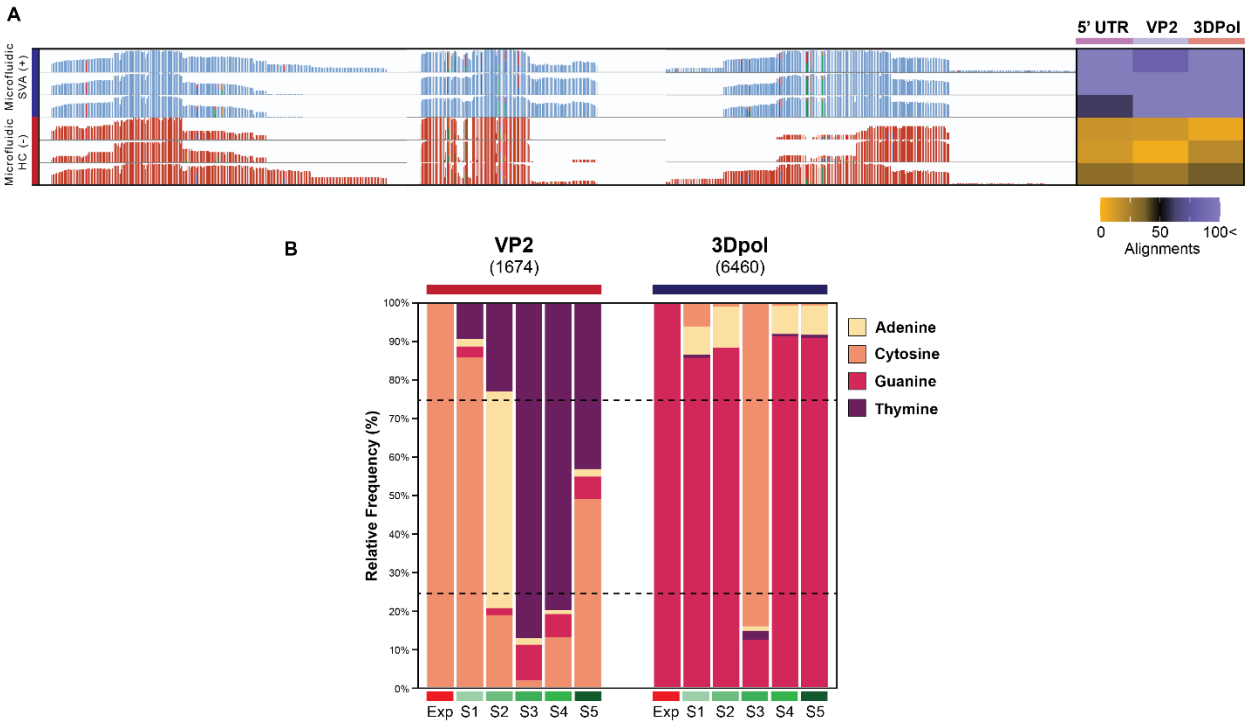

Fig 2.5. (a) Sequence specific overlay of SVA(+)-VFE samples with libraries produced manually and on-chip. (b) Identification of potential single nucleotide variations at three separate base locations

Region	Name	Primer	Sequence
5' UTR	P1	FIP	CCTGGGTCGAGAGGGGAGAACAAGCTCCGACACAGAGTC
		BIP	ATCCACGGCTCGATCCAGAGGCGCTAAGGCCTAGCTACAGT
		F3	GAACCGAGAGGCCTTCGT
		B3	GTCGCCCTATCAGGCAGTA
5' UTR	P2	FIP	AAGGCACGCTAAGGCCTAGCTTTTTGGCATGATCCCCTAGCA
		BIP	GACGGCCTAGTCGTGTCGGTTTTTTTGCCAGAGGCTGTATCGAAAG
		F3	TATCCACGGCTCGATCCA
		B3	CCGGTTACRTCTTCAAAGGT
VP2	P33A	FIP	GGCCGATAGTATGTGCCAAGAGAAATGTCTGACGATTACCGG
		BIP	TGGACTTGGGGTCCCAATTTCGTAGAGGTTCTCGCGTTCA
		F3	GAATGAAGAACAGTGGGTGG
		B3	GGGACGTTTATGTCTACCGA
3DPol	P5	FIP	GCAAAGCGCCCAAGCAACATTGTTGATGTGCCCTCCCT
		BIP	TCCCATCCTGGCTTTCTCCTCGCCTCGAGTTGAGCCCCTAT
		F3	ACCCTCAGAGAAAGTCCGA
		B3	GTCCACGTCATACGTGTTCT
3DPol	P6	FIP	GCCCCTATGACGGTCCAGAAAACTGGCTTTCTCCTCGGCT
		BIP	GCCTTTGACTCTTCACACGGCATGTCCACGGTGAAAAAGTGA
		F3	GCTGCCAAGTTTCAATCCCA
		B3	GCGCAGGGCTAAAACCAT
3DPol	P14	FIP	GTCCCAGCGCAGGGCTAAAACTCCTTCGAGGCTCTCATCTC
		BIP	GGCGAGCGTCGCATCAAGATTTGAGCACTGTGTTCAGCA
		F3	CTCTTCACACGGCACTGG
		B3	TCAATGCCAGAGCAGTCCT

Table 2.1 Primer sets used for loop mediated isothermal amplification

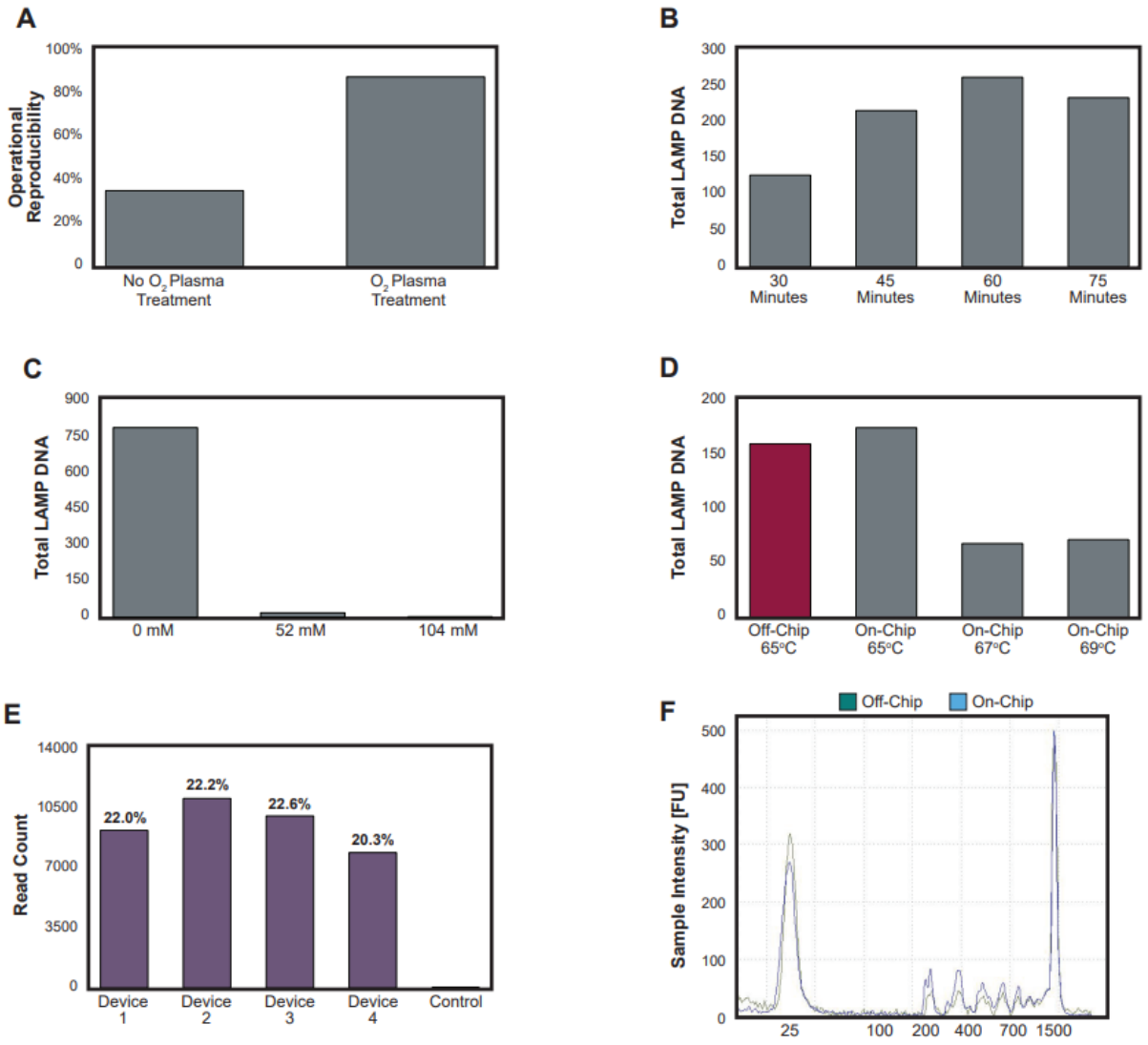

Fig 2.6. On-chip protocol optimization: (a) Sample maintenance within reaction chambers during amplification (no bubble formation), (b) LAMP amplicon generation at different operational timepoints, (c) Impact of NaCl/PEG solution (nM solution) in LAMP reaction mix during isothermal amplification, (d) Amplification performance at relevant operating temperatures, (e) Library preparation reproducibility from a single viral extract sample across 4 identical devices, (f) Size and structure of generated amplicons by Tapestation

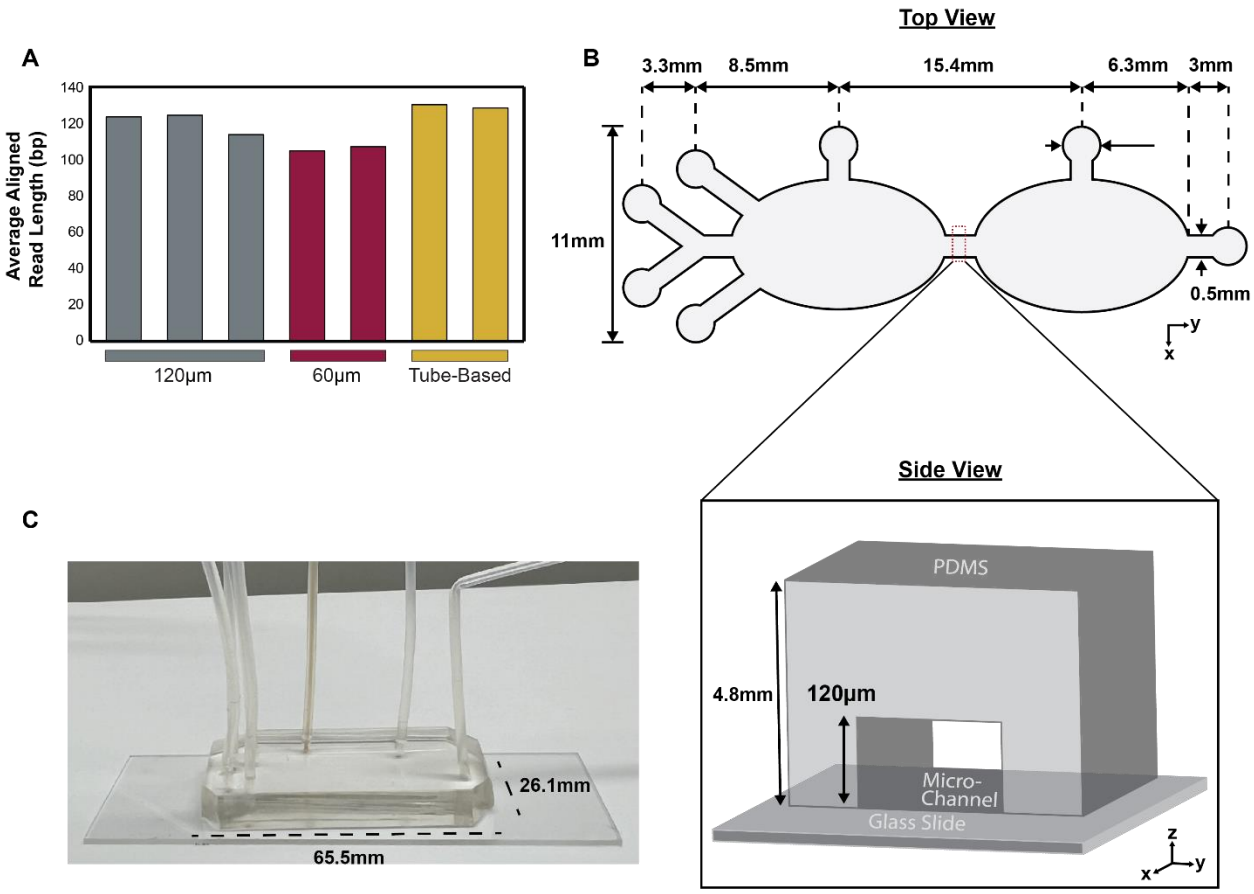

Fig 2.7. Microreactor dimensions. (A) Chamber height determination based on alignment length agreement with off-chip preparations using SVA viral extract, **(B)** Top-down overview of microfluidic channeling and side-view displaying channel and PDMS height (z-axis), **(C)** Image of microreactor with inserted tubing.

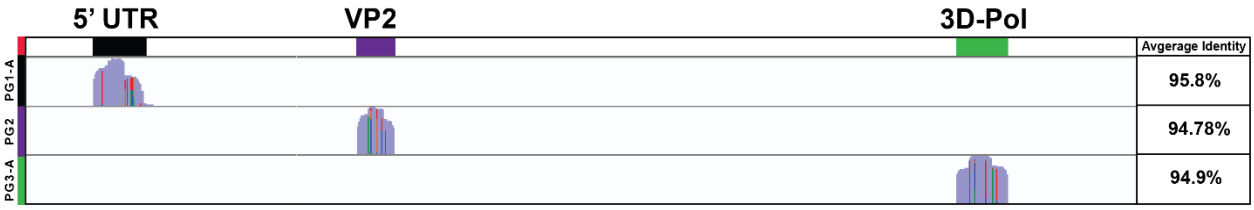

Fig 2.8. Single primer region identity. Sequence specific overlay of viral RNA extracts produced using microreactor targeting utilizing single primer sets. Purple bars indicate consensus alignment with the reference genome.

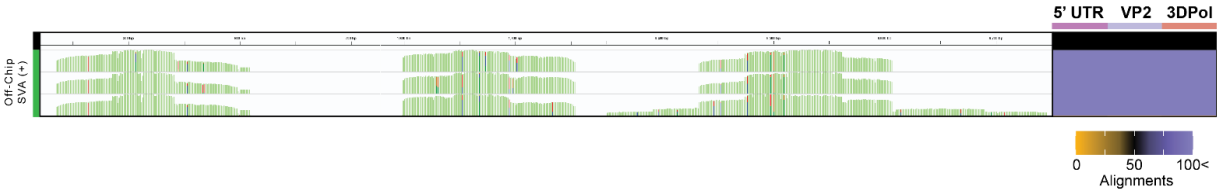

Fig 2.9. Tube based preparation sequence comparison. Sequence specific overlay of SVA(+)-VFE samples produced manually using tube-based method. Green bars indicate consensus alignment with the reference genome.

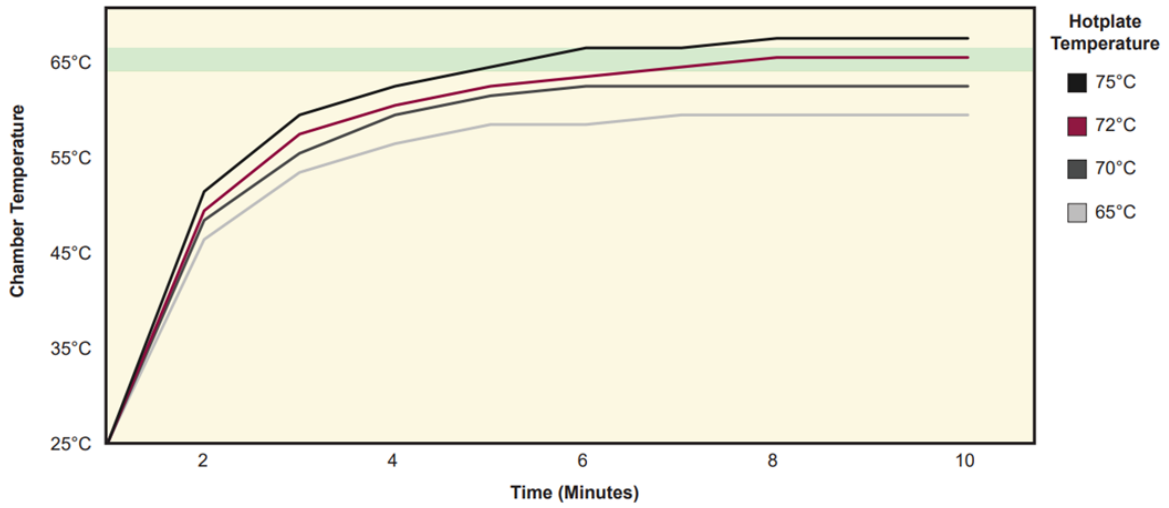

Fig 2.10. Hotplate protocol optimization. Interior microfluidic channel heat curve over time based on hotplate temperature setting. Generated using channel-integrated temperature probe. Green bar indicates targeted internal temperature (65±1°C).

2.4 Conclusion

Recent progress in nanopore sequencing technology has enabled real time detection at resolution not before seen outside of traditional laboratory settings. Here, we have demonstrated the capabilities of a microfluidic device to generate nanopore sequencing libraries with field deployability to match that of the sequencers themselves and unlock their full point of care potential. Our automated system limits human interaction, minimizing sample contamination while offering operational cost savings. We can isothermally amplify multiple target regions along the SVA viral genome at high specificity using our developed, overlapping LAMP primer panel. By sequencing generated amplicons, we are able to avoid the largest hurdle of traditional RT-LAMP assays, false positives by producing data with single nucleotide resolution.

Our system was validated by correctly identifying all tested RT-PCR authenticated SVA positive vesicular fluid samples as well as examined healthy controls. Libraries generated by our system provide diagnostics within 30 minutes of sequencing, allowing for rapid clinical interventions. Libraries generated using our device identified two potential SNVs within our sample set, demonstrating the capability of our device to monitor mutations within the viral genome between populations.

2.5 Materials and Methods

Sample Collection

Senecavalley virus A (SVA) virus was cultured in human lung carcinoma (LC) cell line (both gifts from University of Minnesota Department of Veterinary Medicine) using RPMI

at 37 °C and 5% CO2, passaged every 2–3 days to maintain desired growth.) Whole RNA was extracted using Trizol (Invitrogen) following manufacturers instructions. Extracted RNA was spiked into low-EDTA TE buffer and concentration adjusted using Qubit (Invitrogen, Gubit 2.0 Fluorometer) with RNA quantification kit and used immediately following extraction. SVA positive vesicular fluid samples were taken from SVA positive pigs. Collected fluid was immersed in saliva from the test subject. Saliva was collected from healthy pigs as controls. These samples were stored at -80C until processing. RNA from VFE samples underwent the same Trizol extraction procedure with libraries generated the same day to minimize RNA degradation.

Microfluidic Device Fabrication

Microfluidic device masted mold design was performed using Layout Editor with photomask containing desired pattern printed on 10000 dpi films (Fineline Imaging, Colorado Springs, CO). Photomasks were used to generate microfluidic molds by multilayer soft lithography. Molds were crafted on 5” silicon wafers by spinning SU-8 2025 (Microchem, Newton, MA) to a desired chamber height of 115µm. Following spinning, molds were soft baked at 95°C for 12 minutes, before exposure to UV light for 13 seconds utilizing the crafted photomask for pattern generation. Molds were baked once more at 95°C for 10 minutes before introduction of SU-8 developer (Kayakli) for 5 minutes with moderate agitation to remove undesired channeling. Devices were cleaned with acetone, isopropanol, and water, followed by 25-minute bake at 155°C to smooth cracks in the etched channeling. A 10:1 mixture of PDMS was added to the wafer (RTV615A:RTV615B), degassed by desiccation for 1 hour, then baked at 80°C

overnight to create finished microfluidic mold. Microfluidic inlets and outlets were established by 1.2mm biopsy punch.

Microfluidic Device Setup

Setup of the microfluidic device was performed as depicted (Fig 2.2A). Fluidic handling within the device was performed using automated syringe pumps (Chemyx Fusion 400), where operation was performed either manually or by LabView software automating the pump control. Syringes are connected to 10” flexible C-flex tubing (Cole-Parmer 06433-01) for ease of use, with 1” PFA tubing (IDEX 1622L) inserted into the outlet end of the C-flex tubing serving as the interface between the syringe setup and the device inlets. A 4” U-shaped bent PFA tubing elbow is inserted into the outlet for controlled sample collection and to limit bubble generation during heating. Devices are placed on an elevated platform with the surface directly below the flow chambers removed. A moveable handling block containing a neodymium magnet is then inserted within the platform beneath the device for paramagnetic bead handling when required. A hotplate (Thermo Scientific Super-nova) was set up directly next to the platform, with both handling platform and hotplate surfaces aligned so the device can smoothly slide between the two. Syringe tubing required to carry more than a single fluid had H2O as the base (liquid in syringe and into the initial segment of tubing), with any additional liquids being separated by small air gap (~0.5cm). Air gap size was designed to be captured in the intel channeling without passing into the reaction chambers.

Library Preparation

The same kits were used for LAMP amplification (New England BiO WarmStart LAMP Kit.) and library preparation (Oxford Nanopore SQK-RBK004), regardless of if they were performed on-chip or manually. Manual preparations followed kit manufacturer's instructions. For on-chip amplification, 1ul of sample was added to premade mixture of 12.5ul LAMP mastermix, 9ul H2O, and 2.5ul LAMP primer (IDT) mix then pumped onto the chamber at 5µl/min for 4 minutes. U-shaped outlet tubing was applied 2.5 minutes into sample application to ensure only sample mix in the tubing. For proper amplification, at least 5µl sample should be in the tubing before amplification to prevent gas expansion in the outlet channeling. LAMP is performed on the hotplate for 60 minutes at 65°C. The device is then moved back to the handling platform and cooled at room temperature for 3 minutes before injection of paramagnetic SPRI bead solution (Beckman Coulter SPRIselect) at 3µl/min. During this step, the magnetic stir bar is sporadically applied in varied chamber locations to induce mixing of the NaCl/PEG solution and the amplified sample. During the amplification, it is critical that the bead solution is separated from the microfluidic chamber (0.5cm air gap in PFA tubing) as the NaCl/PEG solution will inhibit LAMP amplification (supplemental Fig S1). The mixed beads allowed to incubate 5 minutes in the reaction chamber, before collection by the neodymium magnet, which corals all beads into the first reaction chamber. The beads are then washed with 75% EtOH at 10ul/min for 3 minutes, with the beads then corralled against the fragmentation solution inlet channeling using the magnet. The remaining 75% EtOH solution is removed by mixture of water through the side inlets (2.5µl/min) and ethanol inlet, which allows a small bubble to enter the chamber and

track water progression. With the beads aggregated by the fragmentation inlet, the fragmentation solution of 2.5ul rapid barcode frag mix (Oxford Nanopore SQK-RBK00) plus 7.5ul ultrapure low EDTA TE buffer is injected at 2.5µl/min, with the magnetic bar removed once injection begins. The emptied outlet elbow is then inserted into the outlet channel, and the free end is placed into a fresh 2.5ml collection tube. Water from inlet 1 is injected at 10µl/min for 4 minutes, with ~10ul of the barcoded library collected in the tube. For single sample processing, 1µl sequencing adaptor is stored cold in the tube, or can be applied through the final side inlet. However, to save costs, fragmented libraries were pooled before application of sequencing adaptor.

Nanopore Sequencing and Bioinformatics

Nanopore sequencing was performed using Oxford Nanopore MinION device (Mk1B) utilizing the Flongle flow cells (FLO-FLG001) following manufacturer's instructions with slight modification to ensure more reliable pore health. 25µl flush solution was pooled on the flow cell inlet. Tape was then placed over one of the two air ports, and a pipette with tip cut for wider bore fit inserted into the other air port. The pipette volume was gently spun to generate back pressure and pull the pooled solution through the flow cell. The tape and pipette were then removed and remaining flush solution applied dropwise to the inlet. Up to 12 barcoded libraries were pooled, and 5ul were combined with sequencing mix of 15µl ONT Sequencing Buffer II and 10µl ONT Loading Beads II, and applied dropwise to the flow cell. Sequencing was controlled using MinKNOW software and ran for 15 hours and with real-time Qscore cutoff of 8. Raw reads were base-called and trimmed into pass reads using the Guppy basecaller (ONT, v5.0.16). Realtime

alignment was performed and monitored using EPI2ME with custom alignment to SVA genome (NC_011349.1). For non-real time alignment following sequencing, minimap2 was used for reference alignment using FASTQ files generated by Guppy. Fraction of total reads at 1 hour over the full course of sequencing as displayed in MinKNOW was used to determine first-hour sequencing results in our analysis. Integrated genome viewer (IGV) was used for visualization of aligned tracks.

2.5 Contributions

T.M.H, under the direction of Dr. Chang Lu, designed and fabricated the microreactor and established operational protocol. T.M.H performed all bioinformatics discussed in this chapter. SVA virus isolates and VFE extracts were provided by Dr. Sunil Mor.

3. Epigenomic investigation of psilocybin for the treatment of opioid addicted mice

3.1 Introduction

Abuse of prescription drugs, particularly pain-relieving opiates such as oxycodone, has increased greatly in recent years among both adult and adolescent populations in the United States. While opioids are among the most effective analgesics in treating acute and chronic pain, long-term use can lead to development of complications such as tolerance, physical dependence, and Opioid Use Disorder (OUD) in some individuals. In 2021 alone, OUD affected over 2.5 million American adults but less than 25% received pharmacological treatment.

Current pharmacotherapies for treatment of OUD exist, but they specifically target the μ-opioid receptor (MOR) to reduce or block the effects of higher efficacy MOR agonists, rather than the underlying mechanisms behind maladaptive behaviors[42]. Examples of these pharmacotherapies include MOR agonists with a lower efficacy, such as methadone or buprenorphine. While the use of opioid-based pharmacotherapies have proven effective for many, they solely target the opioid receptor system and can also cause physical dependence. Therefore, in considering translational perspectives towards the treatment of OUD, it is imperative to identify the molecular targets and neural circuits governing the impact of opioids within the reward pathway, as opposed to a direct targeting of the opioid receptor system.

Classical psychedelics, such as lysergic acid diethylamide (LSD), psilocybin (PSI) and 1-(2,5-dimethoxy-4-iodophenyl)-2-aminopropane (DOI), are compounds that produce

profound changes in perception, sensory processes, and cognition. In the past decade, psychedelics have been studied as potentially transforming therapeutics for a variety of neuropsychiatric disorders such as depression and substance use disorders (SUDs) including tobacco, alcohol and most recently, opioids[43, 44]. In survey studies assessing data from the National Survey on Drug Use and Health (NSDUH), it was found that use of PSI was associated with a 30% decrease in odds of developing an OUD[45]. Further, clinical studies indicate that 80% of subjects reported quitting smoking at the 6-month follow-up following psilocybin administration. This classical psychedelic has also been shown to decrease percentage of drinking days in those with alcohol use disorder. Despite these striking effects, their acute symptoms and potentially uncontrolled recreational use preclude the routine use of classical psychedelics in daily clinical practice. Additionally, although these results have interesting implications for clinical psychiatry and SUD research, many mechanistic questions regarding psychedelics as fast-acting therapeutics remain open.

Epigenetics is the study of how environmental factors can influence gene expression without altering the underlying DNA sequence. Epigenetics involve chemical modifications to DNA and histone proteins, which can regulate gene expression. These changes can be heritable and impact various biological processes, including development, health, and response to environmental stressors[46]. Amongst the most common types of epigenomic modification are histone modifications. Histones are basic proteins that cluster into octamers around which genomic DNA is wound, forming a structure known as a nucleosome. Histones can be affected by dynamic post-translational modifications which typically involve acetylation, phosphorylation, or

methylation of N-terminal tails[47]. These modifications impact DNA winding around the histone octamer, which in-turn affects transcription activity. H3K27ac, a histone modification commonly associated with many neurological disorders, is marked by the presence of acetyl groups on the lysine residue at position 27 of the H3 histone[48]. These acetyl groups neutralize the positive charge of the lysine residue, leading to loosening of DNA winding at this location. The separation from the histone generated by the negatively charged DNA allows increased access of transcriptional machinery to the neighboring sequencies, thus regulating gene expression without sequence alteration. As acetylation leads to increased transcription, these marks are typically denoted as active enhancers and can typically be found on either end of transcriptional initiation sites (TSS). Other histone modifications can lead to tightening of DNA winding (H3K9me3, H3K27me3), transcriptional activation (H3K4me3, H3K36 me3), and other forms of gene regulation.

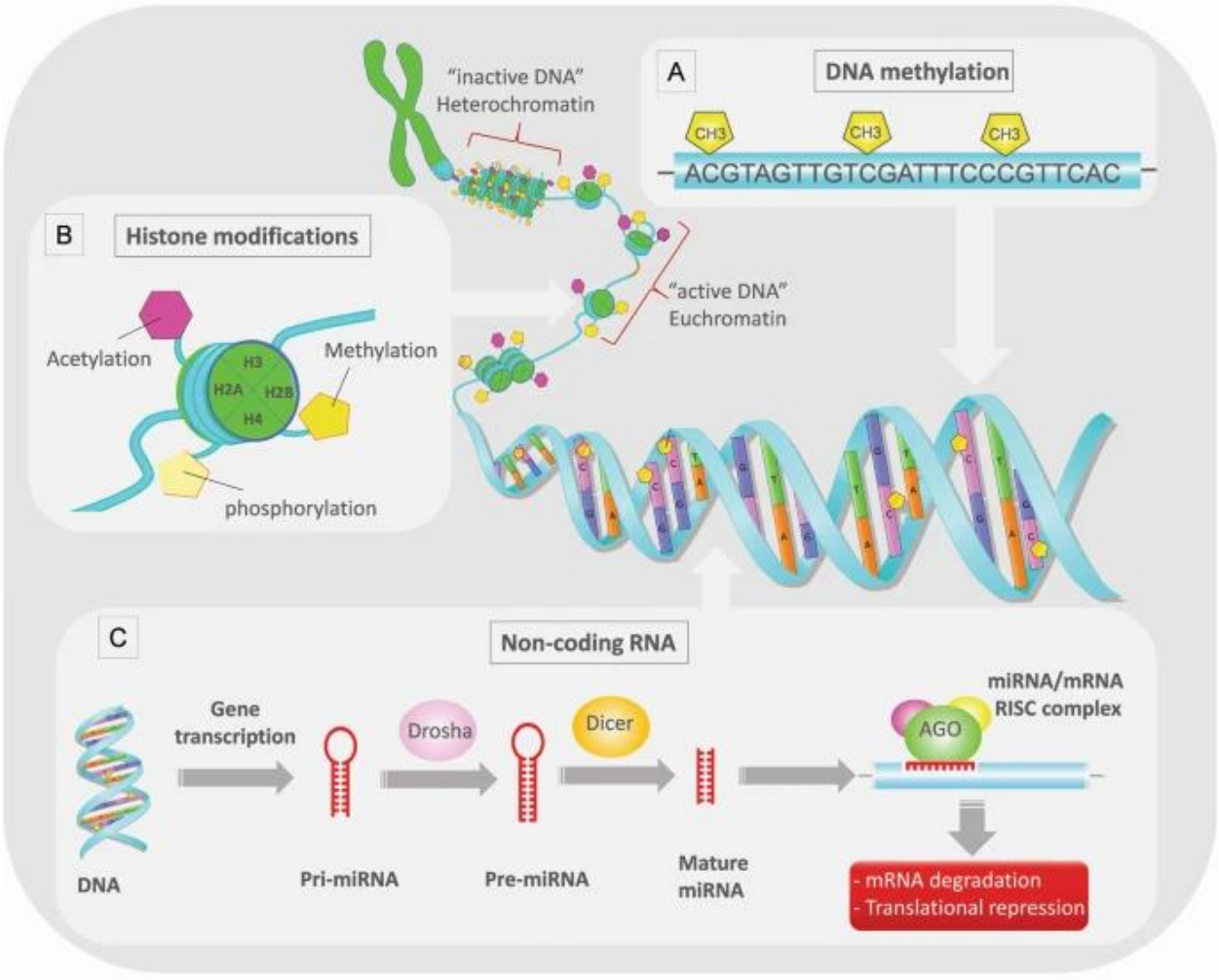

Fig 3.1. Overview of common epigenomic modifications. Reprinted with permission. Springer 2019[49].

Epigenetic factors play an important role in vulnerability to addiction to drugs of abuse[50]. Additionally, a growing body of literature shows that drugs of abuse induce significant epigenomic modifications in the brain. Prolonged abuse of opioids, for example, has been shown to induce epigenomic modifications such as histone acetylation, methylation, DNA methylation, and microRNA (miRs) activity. Of these, histone H3 acetylation has received the most attention, as repeated studies have

confirmed global H3K27 acetylation alterations within the mesolimbic dopamine system[51, 52]. This system includes brain regions such as the nucleus accumbens (NAc) and prefrontal cortex (PFC) and mediates the motivational effects of opioids through mechanisms including MOR stimulated dopamine release. Investigations involving rats and humans showed increased H3 acetylation at discrete loci in the NAc induced open chromatin states, allowing for increased expression of the *GRIA1* glutamatergic gene heavily implicated in drug-seeking behavior[53]. Interestingly, this study showed the degree of H3 acetylation was directly linked to duration of opioid abuse. These findings are supported by similar work showing drastically increased acetylation at multiple lysine residues in histones H3 and H4 in NAc core and shell of rats following sustained opioid self-administration[54]. This study also showed that injection of the histone deacetylase inhibitor sodium butyrate led to increased heroine seeking behavior, further demonstrating the involvement of histone acetylation on opioid addiction.

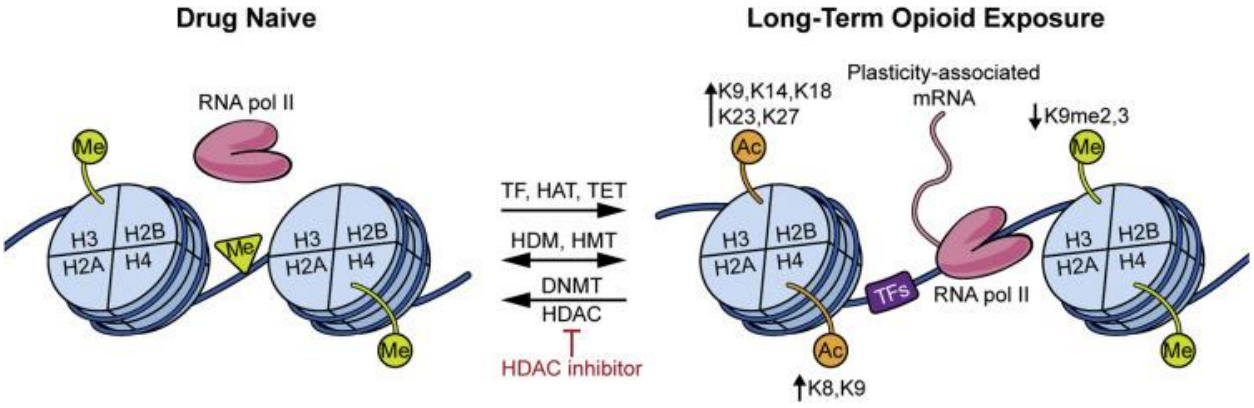

Fig 3.2. Epigenomic alterations induced by opioid exposure. Reprinted with permission. Elsevier 2020[55].

Chromatin immunoprecipitation (ChIP) is a vital tool for resolving epigenomic modifications involving histone modifications. ChIP assays involve shearing chromatin lysed form cell nuclei to generate manageable sized fragments. Then, antibody-coated paramagnetic beads targeting and binding proteins of interest within chromatin fragments are added to solution, isolating them from the bulk chromatin. These beads can be pulled from solution, with targeted fragments then eluted generating a purified enriched sample of the DNA that was bound to the histone mark of interest. Sequencing libraries can then be generated involving PCR amplification of these enriched regions creating highly concentrated samples for sequencing.

Chromatin immunoprecipitation assays are performed using millions of cells in tubes. While bulk ChIP-seq assays of this form are simple and easy to perform with low training requirements, the low sensitivity and large sample requirements limit the feasibility of applying this platform to rare tissue and cell types. Several strategies have been explored to increase sensitivity and lower ChIP-seq input requirements[56-58]. Amongst these methods, microfluidic technologies have shown tremendous promise in increasing ChIP sensitivity. The microfluidic SurfaceChIP, for example, removed the bead requirements, opting instead to utilize antibodies functionalized to the inner surface of the glass slide[59]. However, perhaps the most functional application of microfluidic technology to ChIP assays is the microfluidic oscillatory washing-based ChIP (MOWChIP) platform pioneered in our lab[60, 61]. This process utilizes on-chip packing of coated IP beads at high density followed by thorough microfluidic washing to generate ChIP-seq libraries from as few as 100 cells. The versatility of this platform has

been demonstrated by its successful high-sensitivity targeting of several histone marks, including H3K27ac. The efficiency of this device allows one to obtain high-quality ChIP-seq data from the low nuclei inputs of sorted NAc and PFC neuron populations.

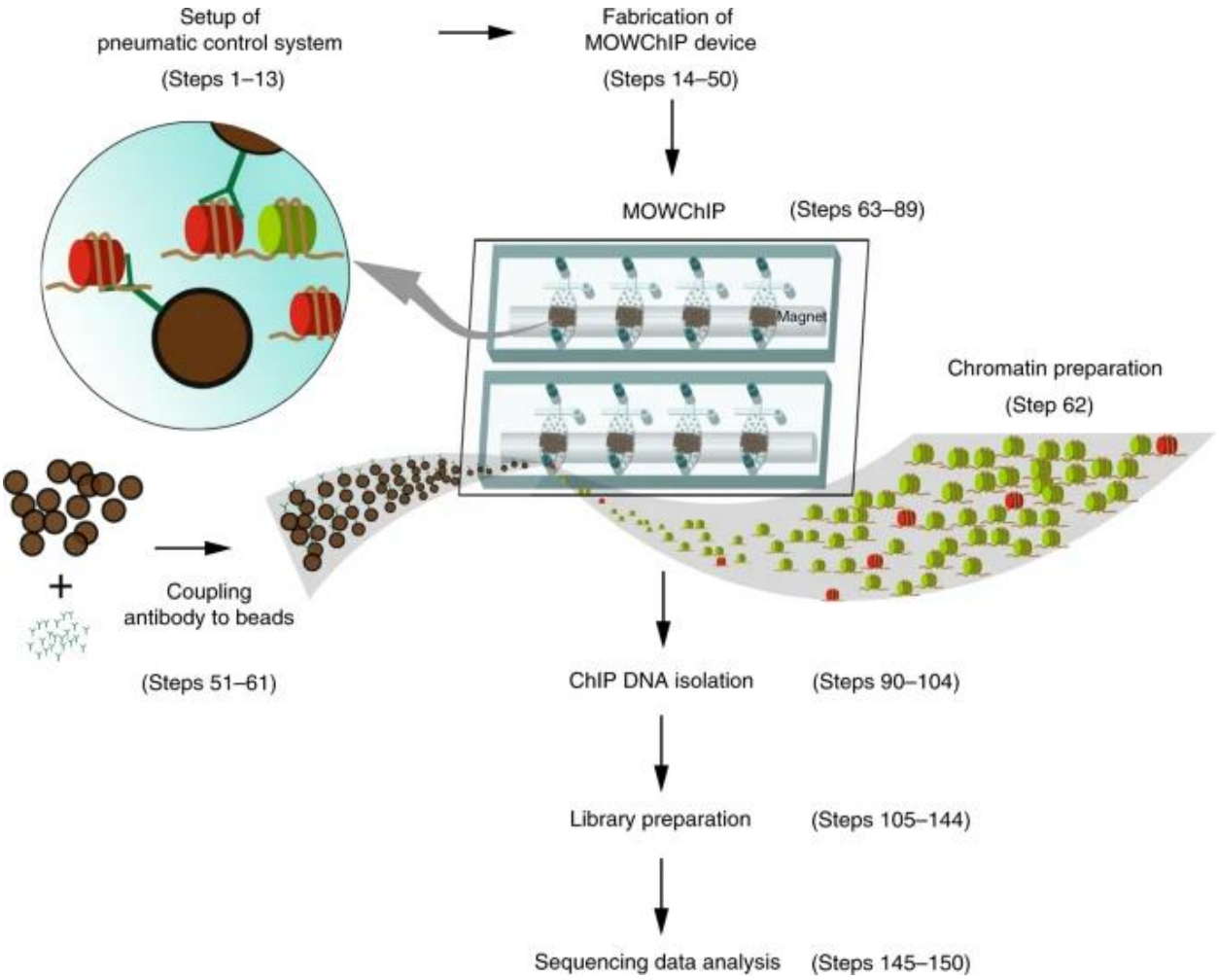

Fig 3.3. Overview of chromatin immunoprecipitation (ChIP) carried out on microfluidic MOW-ChIP device. Reprinted with permission. Springer 2019[61].

3.2 Results

Previous studies have demonstrated that opioids promote an open chromatin state via histone acetylation at H3K27, which facilitates over expression of genes specifically

involved in opioid addiction[53, 62, 63]. Our previous study suggested psychedelic treatment has lasting impact on epigenomic landscape, as opposed to more transient transcriptomic alterations, in frontal cortex of adult male mice[64]. Existing literature suggests notable differences in epigenomic mechanisms between sexes[65], while both human and animal studies demonstrate significant differences in behavioral and synaptic response to psychostimulants[66]. We thus sought to examine potential epigenomic differences between sexes after oxycodone treatment followed by psychedelic exposure. Mice were treated in four conditions: VEH-PSI: mice that received vehicle injection followed by psilocybin administration; OXY-VEH: mice that received 3 separate oxycodone injections with 24 h intervals in between followed by vehicle injection; OXY-PSI: mice that received the oxycodone treatment followed by psilocybin administration at 24 h after the final oxycodone dose; VEH-VEH: control mice that received vehicle injection only. Mice were sacrificed the same day as the last treatment, and pre-frontal cortex (PFC) and nucleus accumbens (NAc) tissue samples were extracted for H3K27ac profiling. The unsupervised hierarchical clustering shows that the individual brain samples largely cluster within each group, with exception in female PFC (PFC/F) and male NAc (NAc/M) samples (Fig 3.6).

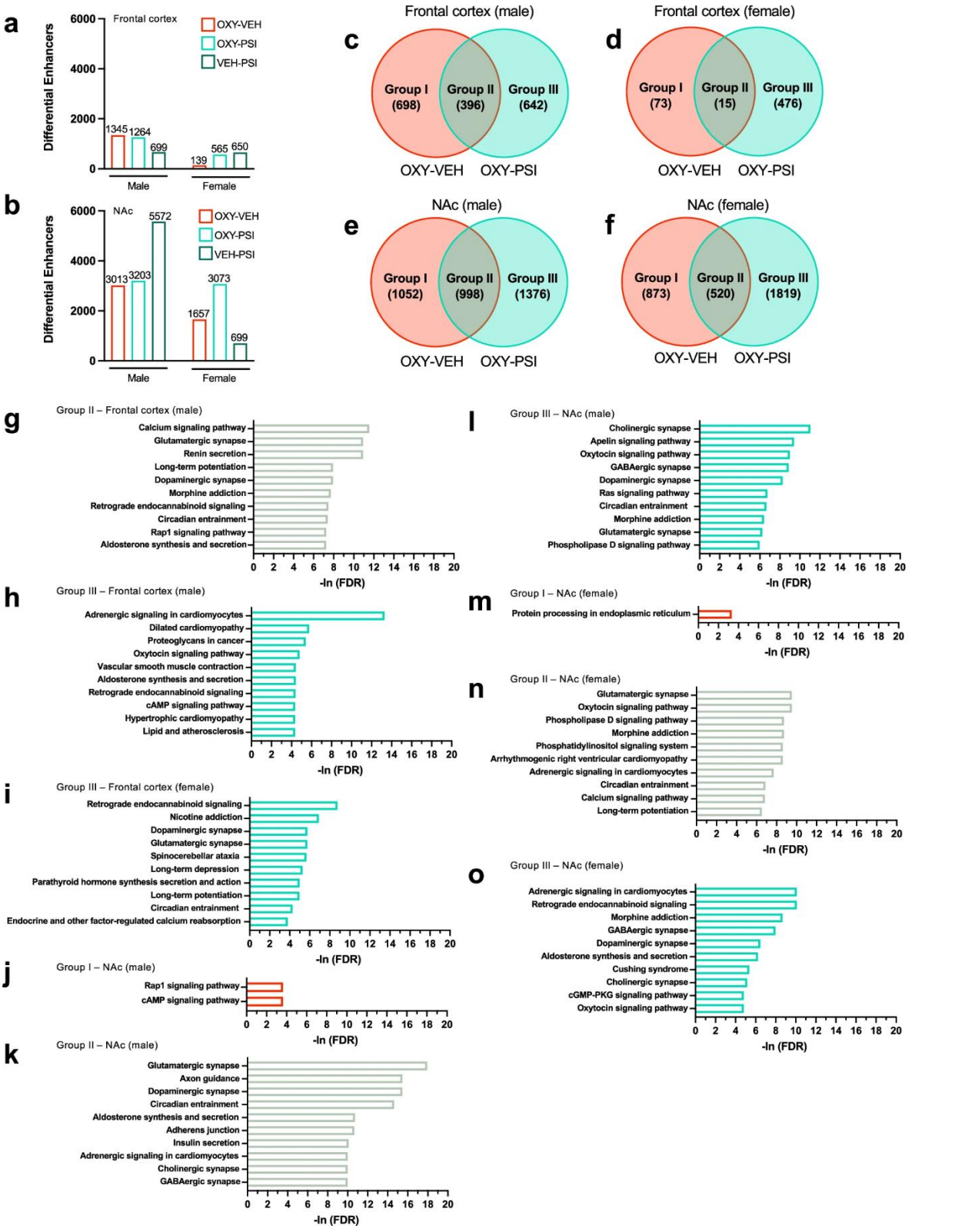

Fig 3.4. Differential enhancer linked genes due to OXY-VEH and OXY-PSI treatment.

Next, we used VEH-VEH as the reference to create lists of differential enhancers that are affected by psilocybin treatment (VEH-PSI), oxycodone treatment (OXY-VEH), and oxycodone treatment followed by psilocybin treatment (OXY-PSI) (Fig. 3.4A&B). Here enhancers are defined by H3K27ac peaks that do not overlap with regions near transcription start sites (TSS)[64, 67]. For all treatment conditions, there are substantially fewer differential enhancers in the female than in the male. Between the two brain regions, the number of differential enhancers is generally much larger in the NAc than in the PFC for both sexes.

We further linked differential enhancer peaks to genes based on Hi-C data[68], and proximity. In order to further understand how psilocybin treatment affects oxycodone addictive brain, we divided differential-enhancer-linked genes associated with OXY-VEH and OXY-PSI into three groups (Fig. 3.4C-F): i) Genes exclusively in the OXY-VEH group. They are altered by oxycodone treatment but recovered to the reference state after psilocybin treatment; ii) Overlapped genes. They are altered by oxycodone treatment and remain altered after psilocybin treatment; iii) Genes exclusively in the OXY-PSI group. They are additional genes that are not altered by oxycodone treatment but altered by psilocybin treatment. We first conducted GO analysis of various groups of genes. For male PFC, group I genes are mostly enriched in GO terms related to cell-substrate adhesion. Cell adhesion molecules are implicated for involvement in addiction[69]. Thus, our results suggest the possibility that psilocybin treatment may counteract the effects of oxycodone treatment via cell-adhesion-associated pathways in male PFC. Group I genes of both sexes in PFC present connections with

morphogenesis, dendrite development, synapse terms. Changes in morphogenesis has known association with psychedelic exposure[64, 70, 71]. Previous studies have demonstrated the effect of opioids on cognitive processes such as learning and memory, typically exhibited by formation of strong connections and cue-triggered behaviors when alerted to the drug[72]. Recovery of genes related to learning and memory may play a role in the observed recovery in conditioned place preference of male mice following psilocybin treatment. KEGG pathway analysis was also performed for the recovered genes (group I) of NAc of both sexes (Fig. 3.4B). For male mice, psilocybin treatment leads to recovery in genes involved in both the Rap1 and cAMP signaling pathways. RAP1 dependent signaling pathways have previously been shown to alter neuronal excitability and behaviors associated with drug responses in the NAc of mice[73], while cAMP signaling pathway up-regulation in the NAc is a prominent feature of opioid addiction[74]. In contrast, genes recovered following psilocybin treatment (group I) in female NAc yielded no significant terms related to addiction or neuron activity. This result suggests that male NAc is affected more strongly by psilocybin treatment in the context of reversing oxycodone effects than the female one. While both male and female PFCs yield no significantly enriched KEGG terms for group I genes, group III genes of female PFC are uniquely enriched in long-term depression. Previously disclosed metanlysis reported highly enriched addiction related genes amongst four drug classifications were most significanlly involved in long-term synaptic depression associated pathways[75].

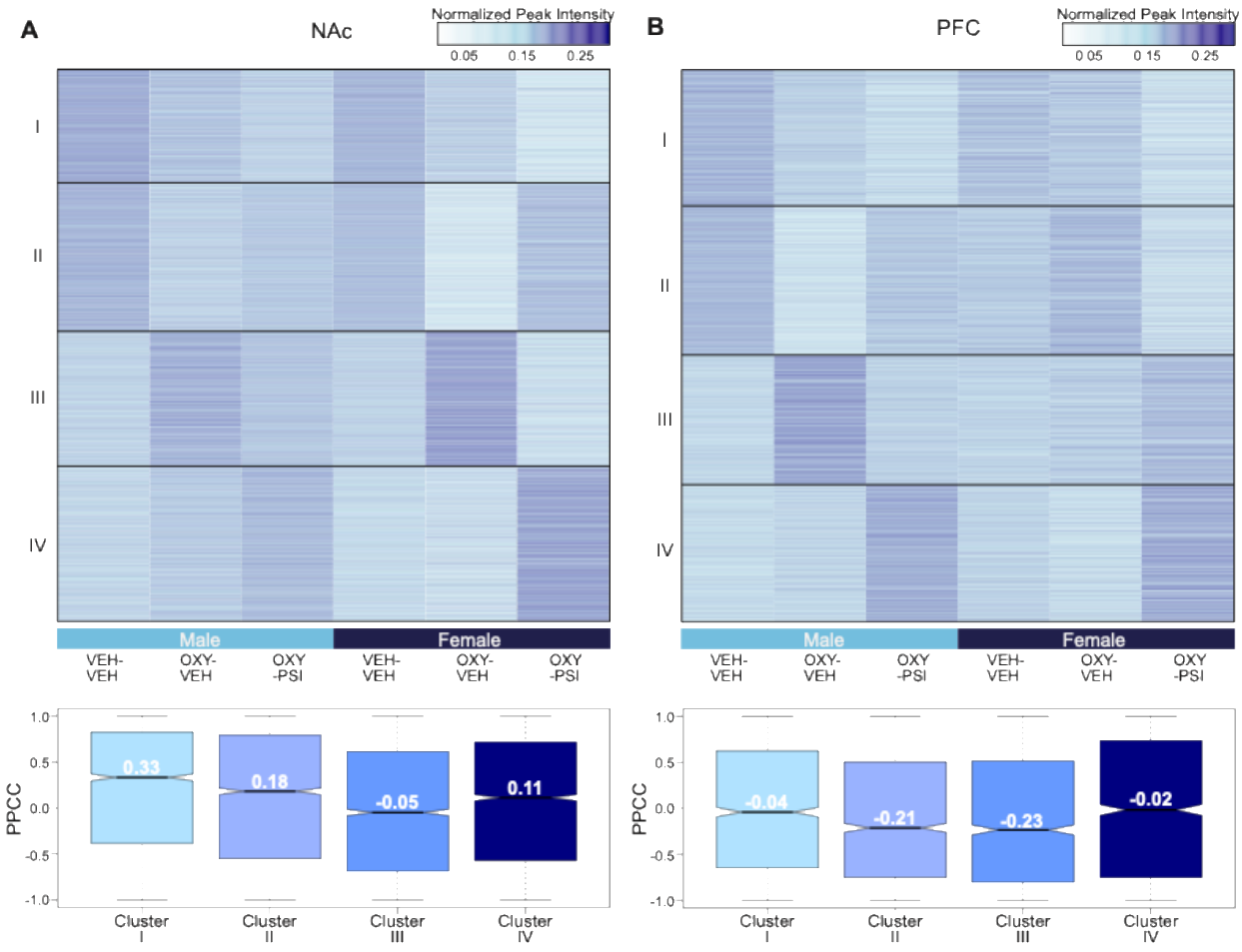

Fig 3.5. K-means clustering of differential enhancers over three conditions and two sexes.

In order to highlight the overall difference between the two sexes in terms of response to OXY and OXY-PSI treatment, we constructed heatmaps of differential enhancers containing both sexes and three conditions (VEH-VEH, OXY-VEH, and OXY-PSI) after K-means clustering for the two brain regions (Fig. 3.4). Similar to our previous work[76], we examined the similarity in the pattern of variation across the three conditions between sexes and quantified it using the average pair-wise Pearson correlation

coefficients (PPCC) for each K-means clusters. As visualized in the heatmap (Fig. 3.5A), the pattern of variation is similar between two sexes for all four clusters of NAc. This is also corroborated by the fairly high PPCCs ranging from -0.05 to 0.33. In contrast, two out of four clusters (clusters II and III) of PFC exhibit very different patterns of variation with PPCCs of -0.21 and -0.23, respectively (Fig. 3.5B). In cluster II of PFC, the two sexes show opposite trends of change when VEH-VEH and OXY-VEH are compared. In cluster III of PFC, the two sexes exhibit opposite trends when OXY-VEH and OXY-PSI are compared. Overall, PFC appears to be the brain region presenting substantially more differences between the two sexes than NAc, with regard to enhancers and OXY/OXY-PSI treatment.

3.3 Discussion

In this study we elucidate the impact on epigenomic regulations brought about by psychedelic administration to opioid-treated mice in two critical brain regions. Previous literature has established that dopamine signaling in the NAc leads to the outsized role of the region in drug addiction compared to other regions in the brain[77, 78]. Our results provide insights into epigenomic changes that are consistent with these previous findings, as both oxycodone and psilocybin, used individually or sequentially, lead to outsized effects in the NAc compared to the PFC in terms of the number of altered enhancers.

Our epigenomic profiling shows a lower number of altered enhancers in the female brain (in both PFC and NAc) than the male counterpart after the same oxycodone treatment. This provides an epigenomic basis for previous observations that female

rodents take in more oxycodone than the male in self-administration to elicit a response[79, 80]. In addition, the female brain also exhibits less altered enhancers than the male one with psilocybin or combined oxycodone/psilocybin treatment. Additionally, we report that male mice show psilocybin induced differences in genes involved in cognitive processes in both PFC and NAc, while similar cognition linked epigenomic alterations are absent in female mice. There has been widespread reporting of the involvement of these brain regions in cognitive function and flexibility[81, 82] and the effects of psilocybin on these processes[83-85]. Recovery of cognition-associated genes in their epigenomic status in male mice after psilocybin treatment may be the underlying mechanism for their change in addiction behaviors. Genes in the male NAC with their epigenomic regulatory status altered by oxycodone administration and then recovered by psilocybin treatment significantly link with biological processes related to cognition, as well as with signaling pathways (Rap1 and cAMP) heavily implicated in drug abuse and addiction.

Global mapping of differential enhancers demonstrates that the more significant differences between the two sexes occur in the PFC instead of the NAc, in spite of the fact that the NAc exhibits more alterations epigenomically than PFC in response to oxycodone or psilocybin treatment. As a part of the reward circuit, PFC is thought to be critical in etiology and persistence of addiction due to its key role in shaping cognition and behavior[86]. Various subdivisions of PFC (e.g. the more dorsal prelimbic and the more ventral infralimbic subregions) also potentially play different roles in driving drug-seeking behaviors [87, 88]. Our results on the substantial sex difference in the PFC

epigenomics underscores the potential importance of PFC in producing response to oxycodone and psilocybin treatment.

3.4 Materials and Methods

Nuclei isolation and sorting via FACS

Nuclei extraction was performed on frozen tissue using an adapted published protocol[89]. Samples were selected using a random number generator to limit downstream batch effects. Each step was performed on ice unless otherwise noted. A small (6-10mg) slice of frozen PFC or NAc tissue was excised from the original sample over dry ice, and immediately plunged into a tissue grinder ((D9063, Sigma-Aldrich) filled with 3ml of ice-cold nuclei extraction buffer (NEB) [[0.32M sucrose, 5mM CaCl2, 3mM Mg(Ac)2, 0.1mM EDTA, 10 mM tris-HCl, and 0.1%(v/v) Triton X-100] along with 30µl of freshly prepared protease inhibitor cocktail (PIC, Sigma-Aldrich), 3µl of 1 M dithiothreitol (DTT, Sigma-Aldrich), 3µl of 100mM phenylmethylsulfonyl fluoride (PMSF, Sigma-Aldrich) in isopropyl alcohol, and 4.5µl of recombinant ribonuclease (RNase) inhibitor (2313A,Takara Bio). The tissue was then homogenized by douncing 25 times with pestle A and 15 times with pestle B. A 40-µm cell strainer (22-363-547, ThermoFisher Scientific) was then used to filter the homogenate, which was then collected in a 15 ml centrifuge tube. Next, the cell suspension was centrifuged at 1000RCF for10 minutes. The supernatant was then discarded, and the collected pellet was resuspended in 0.5 ml of ice-cold NEB with freshly added 5µl of PIC, 0.5µl of PMSF, 0.5µl of DTT, and 0.75µl of RNase inhibitor. 500µl of this sample was mixed with 750µl of 50%(w/v) iodixanol (a mixture of 4ml of OptiPrep^{TM} gradient (Sigma-Aldrich))

and 0.8 ml of diluent [150 mM KCl, 30 mM MgCl2, and 120 mM Tris-HCl]), along with 7.5µl of PIC, 0.75µl of PMSF, 0.75µl of DTT, and 1.125µl of RNase inhibitor. The mixture was centrifuged at 10,000 RCF at 4°C for 20 minutes. The supernatant was then removed and 300µl of 2% (w/v) normal goat serum (50062Z, Life technologies) in Dulbecco's PBS (DPBS, Life technologies), and 0.3µl of RNase inhibitor was gently added to the sample and incubated on ice for 10 minutes, before resuspended the nuclei pellet by pipetting. In order to label and separate NeuN+ and NeuN- fractions by FACS, 10µl of 2ng/ml anti-NeuN antibody conjugated with Alexa 488 (MAB377X, Millipore) in DPBS was added into the nuclei suspension. This suspension was mixed well by pipetting before being placed on a roto-mixer (Labnet) without light at 4°C for 1 hour. Next, NeuN+ and NeuN- fractions were sorted using a BD FACSAria™ cell sorter (BD Biosciences). Around 192µl of the sorted solution containing ~24,000 NeuN+ nuclei was then added into 192µl of 2x lysis buffer [4% (vol/vol) Triton X-100, 100mM Tris-HCl (pH 8.0), 100mM NaCl and 30mM MgCl2 in Milli-Q water], along with 1.92µl of PIC and 1.92µl PMSF. This solution was mixed well and incubated at room temperature for 10 minutes, before addition of 38.4µl of 100mM $CaCl_2$ and 4.8µl 100U/ul MNase (88216, ThermoFIsher). The solution was mixed by vortex and incubated at room temperature for 10 minutes. 42.24µl of 0.5M UltraPure EDTA (15575-038, Life Technologies) and allowed to incubate on ice for 10 minutes to halt the digestion. The resulting mixture was centrifuged at 16,100g for 10 minutes at 4°C and all of the supernatant was collected as sample.

ChIP-seq library construction

Approximately 110ul of digested chromatin fragments from 10,000 sorted nuclei were used for each ChIP assay, with an additional 2,000ul used to generate an input library. ChIP was carried out using a multiplexed MOWChIP assay[61] and anti-H3K27ac (39135, Active Motif) antibody. ChIP-seq libraries were prepared using Accel-NGS 2 S Plus DNA Library kit (Swift Biosciences) from the outputted DNA following MOWChIP. A small modification was made to the manufacturer's procedures as detailed below. To ensure minimal loss of sample throughout numerous rounds of bead purification, a bead solution containing 50% (vol/vol) SPRI beads (Beckman Coulter SPRIselect) and 50% NaCl/PEG solution (IDT, cat #10009854) was used in place of full SPRI beads in the protocol. Following each cleanup step, beads were not discarded following elution, but instead retained throughout the entire library prep processes until the second ligation step was completed.

Library quality control and sequencing

Libraries were assessed for quality control purposes prior to sequencing. Library enrichment was evaluated using qPCR (Bio-Rad CFX Connect) and region-specific primers as previously described[61]. The library fragment size was then assessed using high sensitivity DNA analysis kit (5067-4626, Agilent) on a TapeStation system (2200, Agilent), and any primer-dimers were removed using SPRI-bead purification. Library concentrations were examined using a KAPA library quantification kit (KK4809, KapaBiosystems), and quantified libraries were pooled to 10 nM. Illumina HiSeq 4000 with single-end 50-nt read was used to sequence each library, generating ~15 million reads for each sample and input.

Data Processing

Read trimming was performed using default settings by Trim Galore! (Babraham Institute, https://www.bioinformatics.babraham. ac.uk/projects/trim_galore/). Raw ChIP-seq reads and input data were then mapped to mouse mm10 genome (GRCm38) using Bowtie2 (2.2.5), and peaks were called using MACS2 (2.2.7.1) using a q-value cutoff of 0.05. In order to improve data quality, MM10 blacklisted regions (ENCODE) were removed using SAMtools[90]. Mapped reads were extended by100bp on either side (250 bp total) using BEDtools, with reads then normalized using their respective input [91].

Differential Analysis

Each library was randomly sub-sampled twice to 95% of its original raw reads to generate replicates for analysis using the seqtk package[92]. These subsampled replicates were used as input into diffbind R package for generating a consensus peakset. Consensus peaksets were created independently for each brain region (PFC and NAc) using dba.peakset in diffbind (DBA_BLACKLIST_MM10, consensus = DBA_TREATMENT, minOverlap = 2). The raw read counts were extracted using the diffbind function dba.count (data_blacklist_remove, summits = FALSE, peaks = ConsensusPeaks, filter=1, bScaleControl = TRUE, minCount=1, score=DBA_SCORE_TMM_MINUS_FULL), before removing peaks with less than 20 reads in over 50% of the samples. This count file was then used as the input to Deseq2 package in R[93]. Uniform Manifold Approximation and Projection (UMAP) was used to identify principal components with outsized impact. The following covariates were

regressed out: demographic covariates (sex, condition) and technical covariates (align rate, unique rate, FRiP, NSC, the number of identified peaks, and sequencing batch) by correlating the top 6 principal components (Pearson correlation > 0.7) with these covariates[94]. Differential peak analysis was then performed to identify differential peaks between the 4 conditions (VEH-VEH, VEH-PSI, OXY-VEH, OXY-PSI), and sexes for each brain region (FDR <0.05). The p-values were adjusted by performing a standard Bonferroni correction. For more accurate annotation, enhancers (defined as identified peaks that don't fall within ±2000bp of TSS) were first checked for regional overlap with published Hi-C data on neurons in mice[68]. Overlapped peak regions were annotated based on these existing Hi-C gene links, and the rest of the enhancers were associated with their nearest-neighbor genes.

K-means clustering, UHC, GO term analysis, KEGG pathway analysis

K-means clustering and unsupervised hierarchical clustering (UHC) were performed on differential enhancers for each sex and region. Count files were filtered based on differential peak callouts to form a new consensus peakset comprising the count files of all differential peaks between any three groups. To normalize the counts based on the number of samples, each row of the count file was divided by $\sqrt{sum(X)^2}$. The optimal number of clusters for k-means was determined using the silhouette method[95]. UHC was performed using the default settings of the pheatmap library in R, with cluster position on the x-axis rotated about their dendrogram tentacle for continuity. Gene ontology (GO) biological enrichment analysis was performed on annotated peaks using the bioconducter clusterProfiler package on R[96]. KEGG pathway analysis was

performed using shinyGO version 0.8[97]. Both GO and KEGG analysis were performed using FDR adjusted P values < 0.05 to ensure generation of high-confidence terms.

3.5 Supplemental Figures

Sample	Sex	Condition	Treatment
946-1	MALE	Saline	Saline
949-1	MALE	Saline	Saline
950-1	MALE	Saline	Saline
947-4	MALE	Saline	Saline
948-4	MALE	Saline	Saline
945-4	MALE	Saline	Saline
949-2	MALE	Saline	Psilocybin
946-2	MALE	Saline	Psilocybin
947-3	MALE	Saline	Psilocybin
950-2	MALE	Saline	Psilocybin
948-3	MALE	Saline	Psilocybin
945-3	MALE	Saline	Psilocybin
946-3	MALE	Oxycodone	Saline
950-3	MALE	Oxycodone	Saline
947-1	MALE	Oxycodone	Saline
945-1	MALE	Oxycodone	Saline
948-1	MALE	Oxycodone	Saline
949-3	MALE	Oxycodone	Saline
946-4	MALE	Oxycodone	Psilocybin
950-4	MALE	Oxycodone	Psilocybin
947-2	MALE	Oxycodone	Psilocybin
949-4	MALE	Oxycodone	Psilocybin
948-2	MALE	Oxycodone	Psilocybin
945-2	MALE	Oxycodone	Psilocybin
Sample	**Sex**	**Condition**	**Treatment**
968-1	FEMALE	Saline	Saline
969-1	FEMALE	Saline	Saline
967-1	FEMALE	Saline	Saline
966-1	FEMALE	Saline	Saline
971-1	FEMALE	Saline	Saline
970-1	FEMALE	Saline	Saline
969-2	FEMALE	Saline	Psilocybin
967-2	FEMALE	Saline	Psilocybin

966-2	FEMALE	Saline	Psilocybin
968-2	FEMALE	Saline	Psilocybin
970-2	FEMALE	Saline	Psilocybin
971-2	FEMALE	Saline	Psilocybin
967-3	FEMALE	Oxycodone	Saline
971-3	FEMALE	Oxycodone	Saline
966-3	FEMALE	Oxycodone	Saline
968-3	FEMALE	Oxycodone	Saline
969-3	FEMALE	Oxycodone	Saline
970-3	FEMALE	Oxycodone	Saline
968-4	FEMALE	Oxycodone	Psilocybin
966-4	FEMALE	Oxycodone	Psilocybin
971-4	FEMALE	Oxycodone	Psilocybin
969-4	FEMALE	Oxycodone	Psilocybin
967-4	FEMALE	Oxycodone	Psilocybin
970-4	FEMALE	Oxycodone	Psilocybin

Table 3.1 Information on psilocybin treated and non-treated OUD model

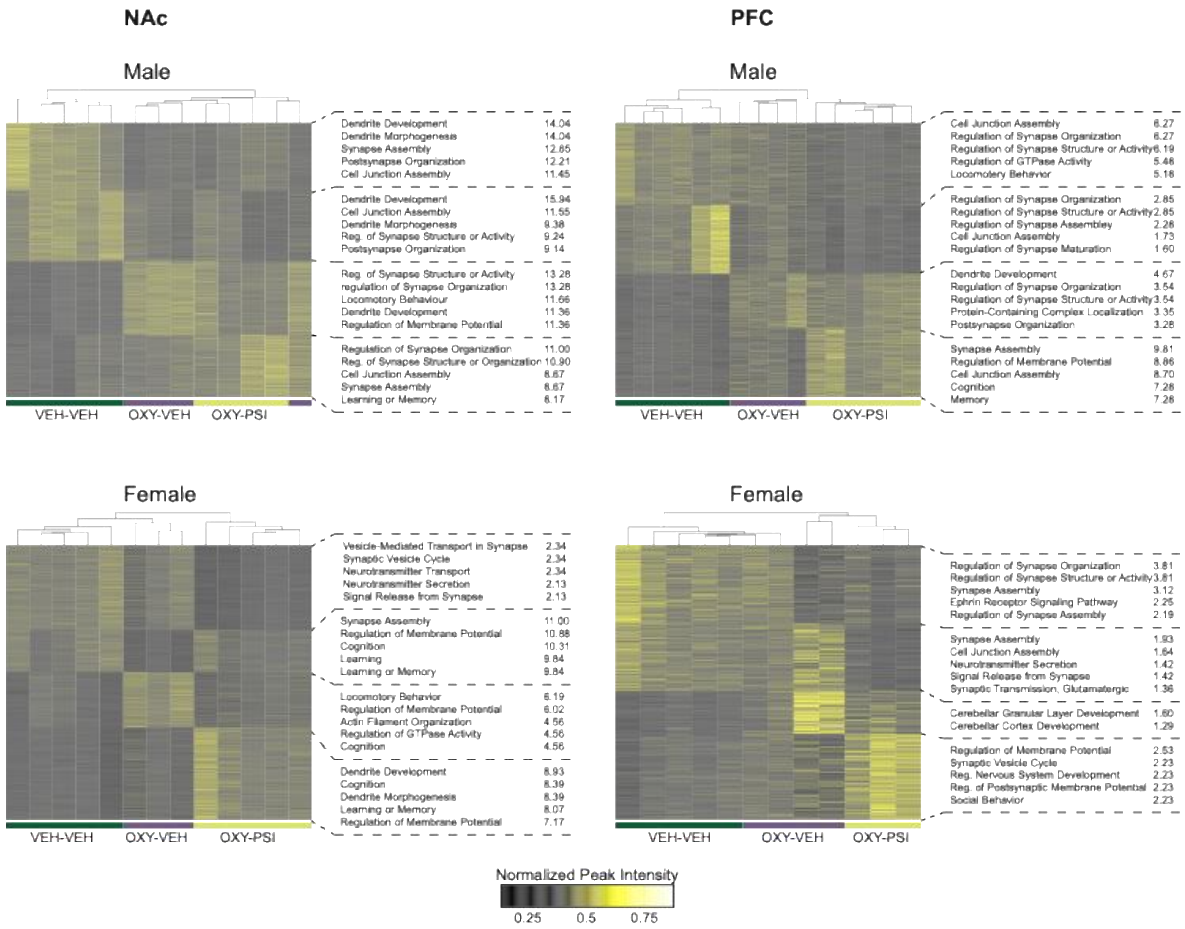

Fig 3.6. Unsupervised hierarchical clustering of differential enhancer data for three conditions, both sexes, and two brain regions.

3.6 Contributions

T.M.H, under the direction of Dr. Chang Lu, performed ChIP-seq assays from mouse brains collected by Dr. Javier Gonzalez-Maeso and Alaina Jaster. T.M.H performed all bioinformatics discussed in this chapter. Belle Buzzi performed animal behavioral studies under the direction of Dr. M. Imad Damaj.

4. Multi-omic evaluation of multigenerational effects of sustained psilocybin exposure on mouse frontal cortex

4.1 Significance

Though existing research on epigenomic effects of psychedelics on the brain is limited, the present body of knowledge suggests a lasting, important impact on multiple regions. While a few studies have been able to observe psychedelic induced epigenomic alterations in tissue collected from target brain regions, the rare nature of these cells makes it difficult for most labs to conduct high quality ChIP assays, limiting further explorations. However, the microfluidic ChIP-seq tools designed and optimized in our lab allow a unique opportunity to follow-up on these previous findings, providing highly sensitive sequencing of these critical marks in rare neuron populations. This capability was demonstrated in our preliminary work, utilizing MOW-ChIP microfluidic technology to perform epigenomic evaluation of psilocybin as a potential treatment option for opioid addiction.

Epigenetic inheritance is characterized as the passing of epigenomic modifications through generations, regulating gene expression without altering the primary nucleic acid structure. This phenomenon is common in plants but observed less frequently in mammals due to epigenetic correction through the process of germline reprogramming[98]. Epigenomic inheritance can be divided into two distinct categories based on generational removal from the F0 mother (Fig 4.2). Intergenerational epigenomic effects arise due to environmental stress on the mother, fetus, or either F0 or F1 gametes, and do not carry through generations beyond the F1 offspring. Transgenerational epigenomic alterations, on the other hand, are germline mediated

and are denoted as existing in the F3 generation and beyond (with maternal origination) or F2 generation and beyond (if paternally inherited).

Links between prenatal use of drugs of abuse and next generational phenotype are well documented, but these are likely due to direct effects upon the fetus. In fact, early investigations into direct embryonic application of various hallucinogenic compounds has a strong impact on behavioral outcomes[99]. Indirect effects of parental drug use on gametes are less understood. However, existing investigations have uncovered links between indirect effects of parental usage of some drugs of abuse, such as paternal alcohol abuse and alcohol abuse in children. Results of these studies indicate that genetic factors cannot explain observed inheritance alone[100]. Recent studies have begun to uncover multigenerational epigenomic effects following sustained use of drugs of abuse which may account for the non-genetic phenotypic patterns. Mechanistic investigation revealed decreased DNA methylation at the *Bdnf* promoter in VTA tissue of alcohol-sired may be the epigenomic driving force behind increased VTA *Bdnf* expression suspected of driving the variance in alcohol consumption patterns[101]. Similar studies involving psychostimulants have been performed examining the offspring of cocaine sired mice, where increased acetylation of histone H3 with *Bdnf* promotors was uncovered as a potential epigenomic mechanism driving the observed multi-generational effects[102]. Yet to this point, multigenerational epigenomic effects from psychedelic administration have not been investigated. As a changing regulatory landscape is leading to a rise in interest in psychedelics for a wide host of disorders, gaining a full understanding of the cross-generational effects and downstream risks assumed by the patient becomes critical for an informed treatment decision.

In this study, we build off the knowledge of our preliminary work and apply microfluidic MOW-ChIP for enhancer profiling of frontal cortex neurons of the offspring of adult, preconception dam mice following sustained psilocybin exposure. We broadened the scope of this study by taking a multi-omic approach, performing RNA-seq and bulk bisulfite sequencing to fully interrogate the inheritance of the most well-documented epigenomic marks associated with primary psilocybin exposure, as well as their associated effects on gene expression.

4.2 Introduction

Psychdelics

Psychedelics are a class of psychoactive drugs that induce changes in perception and emotional condition, generally leading to a significantly transformed state of consciousness. After a wave of research in the 1950s, psychedelic studies slowed until a revival in interest in clinical application of psychedelics for treating myriad diseases in the 1990's. While the method of action of psychedelics is complex and not fully mapped, many psychedelics are known to act as agonists upon 5HT2A serotonin receptors, promoting neuroplasticity[103]. This is the case for psilocybin, the most common and, perhaps, most ancient psychedelic substance known. This naturally occurring substance is produced in various species of fungi, and once metabolized into psilocin in the body, generates common psychedelic hallucinogenic effects. Since loosening regulations in 1992, studies have utilized the psychoactive effects of psilocybin to treat mental health conditions such as depression, anorexia, obsessive compulsive disorder,

alcohol dependence, and tobacco cessation with varying degrees of success.[43, 44, 104-106]. Therapeutic success of these compounds is commonly attributed to a restructuring of neural circuitry due to reopening of neural plasticity windows, especially when paired with guided behavioral training[107].

Frontal Cortex Neurons

The frontal cortex in mammals is responsible for the control of many critical aspects of emotion, behavior, and motor skills. Frontal cortical subregions, such as the PFC, are heavily affected by many brain disorders. Analysis of this region yields important insight into the mechanism of those mental disorders, as well as their potential treatments. Thorough investigation of psychedelic compounds has shown effects localized in the frontal cortex, which may explain some of their therapeutic potential. Specifically, psychedelics compounds appear to affect frontal cortex through promotion of neuron strengthening and their connections throughout the region. This is supported by the large concentration of 5-HT2A serotonin receptors within the region, as well as observed increases in brain-derived neurotrophic factor (*Bdnf*) expression, which is known to support neuron growth following psychedelic use[108]. At the circuit level, psilocybin induced plasticity may rebalance prefrontal-limbic connections to strengthen top-down regulatory control over emotional processing. This reconfiguring of cortical networks is increasingly viewed as a key contributor to their rapid and durable therapeutic effects.

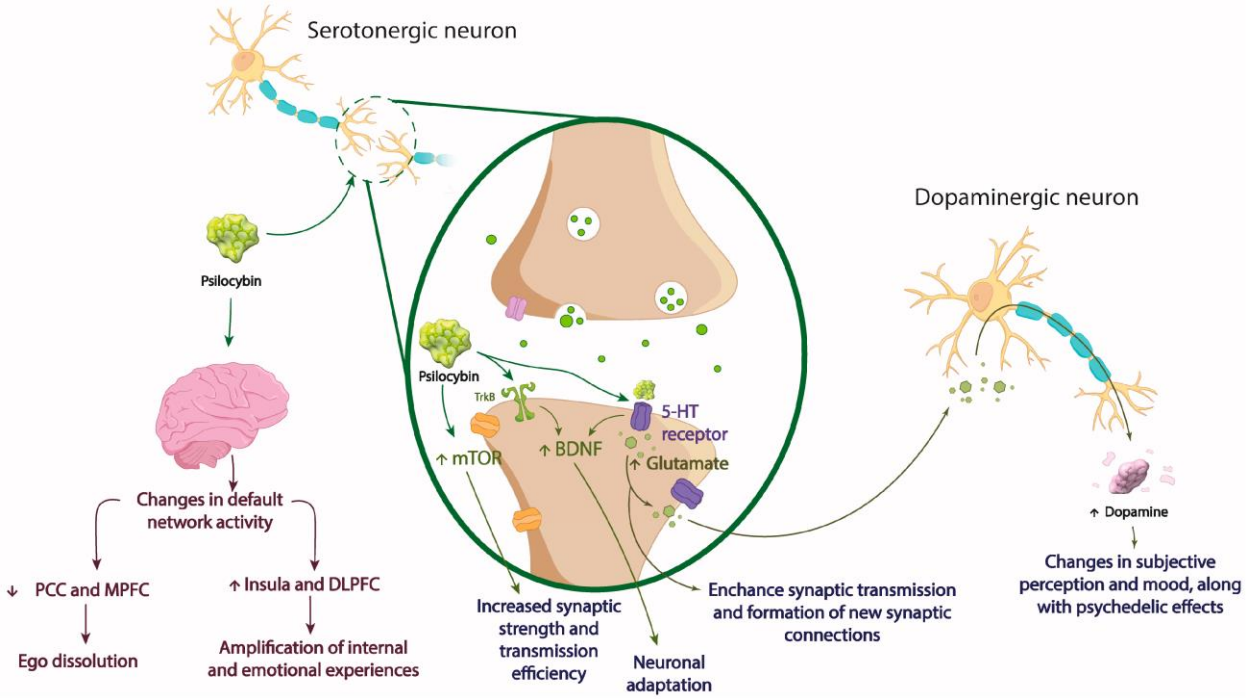

Fig 4.1. Neuronal synapse activity following psychedelic exposure. Reprinted with permission, MDPI 2024[109].

Epigenomics of psychedelics in mouse brains

Early studies have demonstrated correlative evidence that psychedelic drug administration is accompanied by significant epigenetic alterations. Recent epigenetic examinations of psychedelic exposure have centered on DNA methylation. Previously published work by Inserra et al revealed significantly modulated DNA methylation at over 600 CpG sites in mouse PFC following LSD treatment[110]. Further evaluation determined CpG sites differentially methylated after LSD exposure link to genes involved with neurotropic-, neurotrophic-, and neuroplasticity-related signaling[111]. Similar findings have also been shown with other psychedelics such as ayahuasca[112]. However, while methylation clearly plays a role in observed transcription, the overall low

quantity of significantly modulated sites and proteomic overlap suggests DNA methylation is likely not the predominant modification driving observed gene expression alterations. Early studies also suggested hyperacetylation at histone mark H3K27 may play an even larger epigenomic role in altering observed expression following psychedelic exposure[113]. This has been supported by previously published work from our lab using another psychedelic drug, 2,5-dimethoxy-4-iodoamphetamine (DOI), which demonstrated sustained active enhancer expression, as marked by sustained acetylation levels at this lysine residue[114]. Yet as psychedelics gain popularity for a variety of therapeutic purposes, there remains a gap in the literature on epigenomic studies involving psilocybin. Tangential studies implicate epigenetic mechanisms[115, 116], yet these studies did not directly examine histone modification or DNA methylation. Defining active epigenomic marks and downstream gene expression implications will not only uncover high-potential mechanistic targets for therapeutics but provide a full understanding of the lasting impact of psychedelics on neural circuitry. Detailed epigenomic profiling of cortex tissue utilizing our lab's low-input MOW-ChIP device will provide a first of its kind profile of a high-significance histone modification in mouse cortex following sustained psilocybin exposure.

Transcriptomics of psychedelics in mouse brains

Transcriptomic studies of psychedelic use have shown that gene expression may play a role in serotonin 5-HT2R receptor activation in brain regions critically related to behavioral effects, such as the cortex[114]. Additionally, psychedelics appear to trigger unique transcriptional responses in various cell types, notably in both excitatory and

inhibitory neurons[117]. However, when compared to epigenomics, gene expression changes brought about by psychedelics such as DOI are generally more muted, and the effects are not as lasting. Emerging evidence suggests this may not be the case with psilocybin, which induces significant cortical gene expression alterations following primary exposure. Recent microscopy studies show notable variance in targeted gene expression between brain regions following psilocybin exposure, with frontal cortex showing the largest difference in immediate early gene expression (c-Fos), a finding in contrast with other drugs such as DOI[118]. However, while indicative, microscopy studies are naturally limited in scope to low number of gene targets. Full transcriptomic sequencing of cortical neurons would help elucidate the role gene expression plays in dictating the therapeutic effects of psilocybin and help determine pathways and mechanisms involved.

Sex-related differences

Historically, neurological studies have been predominantly performed on male populations[119]. As more sex-related studies are performed, the degree of distinction between genders in results has exposed the clear need identifying existing gaps in the literature regarding female subjects, as well are avoiding creating future holes. This is indeed true for epigenomic and transcriptomic studies involving psychedelics, where even our own previous studies have excluded female cohorts despite significant findings[64]. As new epigenomic and transcriptomic studies including both sexes emerge, the large disparity in degree of effect of psychedelics between sexes becomes evident. One recent psychedelic study found 250% increase in differential gene expression in

the frontal cortex of female mice over males when given ibogaine[120]. Our own preliminary work (section 3.2.3) has indicated distinct differences in the magnitude of differential enhancer linked genes and associated pathways between male and female opioid addicts given a single dose of psilocybin. With woman displaying significantly higher incidence rates of several of the mental health disorders targeted by psychedelic treatments, it is imperative that we have a robust understanding of female specific effects to properly evaluate treatment decisions and align expectations[121, 122].

As psychedelic treatment in future mothers expands, it is critical that clinicians and patients have a full understanding of potential downstream impact on their offspring. To address this existing gap in the general body of knowledge, we will investigate potential multigenerational epigenomic alterations caused by sustained preconception-period maternal psychedelic drug use on their offspring. This will be accomplished using a multi-omic investigation using the offspring of the mothers who were treated with sustained psilocybin dosages before mating. Our multi-omic approach will include investigation of the two most common epigenomic modifications, histone modification and DNA methylation, to uncover the most likely mechanisms at hand, along with transcriptomic analysis of observed gene expression to discern genetic and epigenomic causes. This study will target the highly effected frontal cortex, resolving cell-types to identify changes in the neuronal populations which are most heavily impacted by psychedelic exposure.

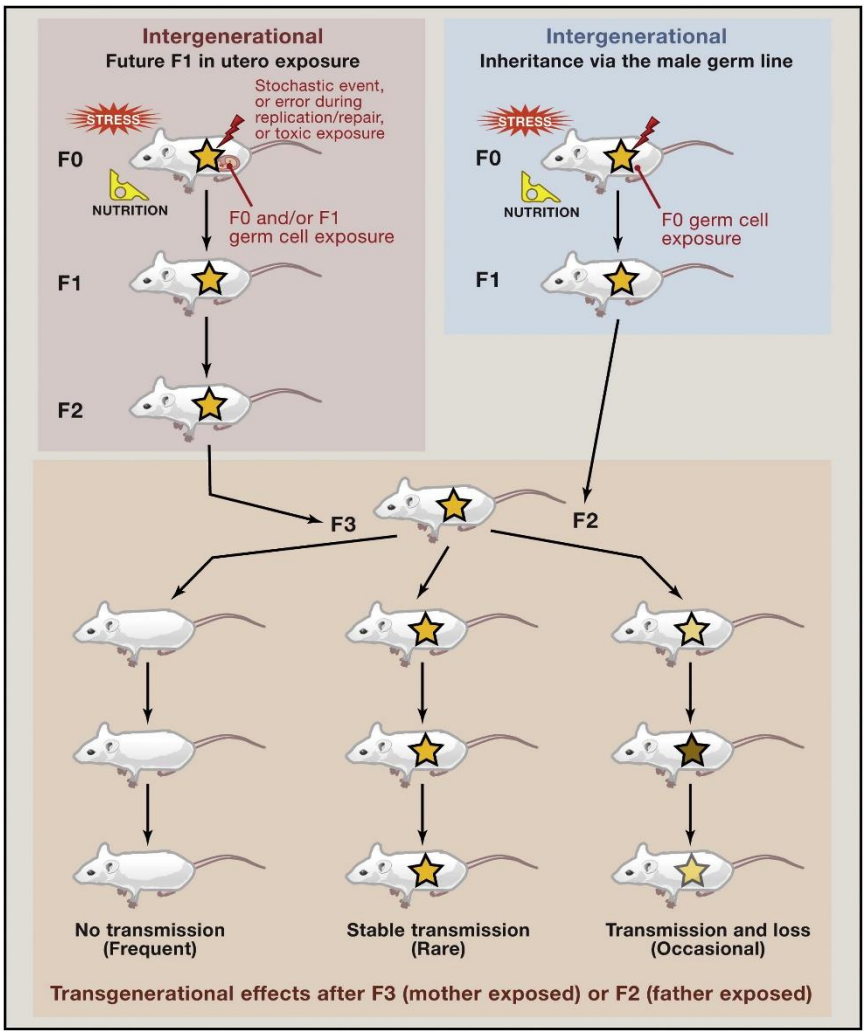

Fig. 4.2. Diagram of intergenerational and transgenerational effects. Reprinted with permission, Elsevier 2014[98].

4.3 Materials and Methods

Experimental model and subject details

All mice in this study were sourced by Dr. Javier Gonzalez-Maeso at Virginia Commonwealth University. In total, 6 female C57BL mice were injected with 1ml/mg psilocybin, while an additional 6 female C57BL mice were injected with saline solution.

All subjects were mated and offspring were weaned when appropriate (~21 days). At 10 weeks, offspring were sacrificed and frontal cortex tissue harvested and snap frozen. All procedures were conducted in accordance with NIH guidelines and were approved by the Virginia Commonwealth University Animal Care and Use Committee. Every reasonable effort was made to minimize animal suffering.

Nuclei isolation and sorting via FACS

Nuclei extraction was performed on frozen tissue using an adapted published protocol[89]. Samples were selected using a random number generator to limit downstream batch effects. Each step was performed on ice unless otherwise noted. A small (6-10mg) slice of frozen PFC tissue was excised from the original sample over dry ice, and immediately plunged into a tissue grinder ((D9063, Sigma-Aldrich) filled with 3ml of ice-cold nuclei extraction buffer (NEB) [[0.32M sucrose, 5mM CaCl2, 3mM Mg(Ac)2, 0.1mM EDTA, 10 mM tris-HCl, and 0.1%(v/v) Triton X-100] along with 30µl of freshly prepared protease inhibitor cocktail (PIC, Sigma-Aldrich), 3µl of 1 M dithiothreitol (DTT, Sigma-Aldrich), 3µl of 100mM phenylmethylsulfonyl fluoride (PMSF, Sigma-Aldrich) in isopropyl alcohol, and 4.5µl of recombinant ribonuclease (RNase) inhibitor (2313A,Takara Bio). The tissue was then homogenized by douncing 25 times with pestle A and 15 times with pestle B. A 40-µm cell strainer (22-363-547, ThermoFisher Scientific) was then used to filter the homogenate, which was then collected in a 15 ml centrifuge tube. Next, the cell suspension was centrifuged at 1000RCF for10 minutes. The supernatant was then discarded, and the collected pellet was resuspended in 0.5 ml of ice-cold NEB with freshly added 5µl of PIC, 0.5µl of PMSF, 0.5µl of DTT, and

0.75µl of RNase inhibitor. 500µl of this sample was mixed with 750µl of 50%(w/v) iodixanol (a mixture of 4ml of OptiPrep™ gradient (Sigma-Aldrich)) and 0.8 ml of diluent [150 mM KCl, 30 mM MgCl2, and 120 mM Tris-HCl]), along with 7.5µl of PIC, 0.75µl of PMSF, 0.75µl of DTT, and 1.125µl of RNase inhibitor. The mixture was centrifuged at 10,000 RCF at 4°C for 20 minutes. The supernatant was then removed and 300µl of 2% (w/v) normal goat serum (50062Z, Life technologies) in Dulbecco's PBS (DPBS, Life technologies), and 0.3µl of RNase inhibitor was gently added to the sample and incubated on ice for 10 minutes, before resuspended the nuclei pellet by pipetting. In order to label and separate NeuN+ and NeuN- fractions by FACS, 10µl of 2ng/ml anti-NeuN antibody conjugated with Alexa 488 (MAB377X, Millipore) in DPBS was added into the nuclei suspension. This suspension was mixed well by pipetting before being placed on a roto-mixer (Labnet) without light at 4°C for 1 hour. Next, NeuN+ and NeuN- fractions were sorted using a BD FACSAria™ cell sorter (BD Biosciences). Immediately following FACS sorting, 1µl of RNase inhibitor was mixed into 99µl of the sorted solution containing ~20,000 NeuN+ nuclei for RNA extraction. RNA extraction was performed using RNeasy mini kit (Qiagen) following manufacturers protocol. Following extraction, purified RNA was eluted into 24µl RNase-free H2O along with 1ul of RNase inhibitor and stored at -80°C overnight.

Chromatin fragmentation

Around 300 µl of the sorted solution containing ~60,000 NeuN+ nuclei was then added into 300µl of 2x lysis buffer [4% (vol/vol) Triton X-100, 100mM Tris-HCl (pH 8.0), 100mM NaCl and 30mM MgCl2 in Milli-Q water], along with 3µl of PIC and 3µl PMSF. This

solution was mixed well and incubated at room temperature for 10 minutes, before digestion by addition of 60µl of 100mM CaCl2 and 7.5µl 100U/ul MNase (88216, ThermoFIsher). The solution was mixed by vortex and incubated at room temperature for 10 minutes. Fragmentation was halted by adding 66µl of 0.5M UltraPure EDTA (15575-038, Life Technologies) and incubating on ice for 10 minutes. The resulting mixture was centrifuged at 16,100g for 10 minutes at 4°C and all of the supernatant was collected as sample. Of this, 250µl (~20,000 nuclei) was stored at 4°C for MOWChIP assays. The remaining supernatant, to be used for WGBS, was brought up to 500µl with low-EDTA TE buffer then mixed with 20µl proteinase K solution (1mg/ml) for protein digestion. This solution was mixed using a thermal shaker at 1000rpm and 65°C for 1 hour. Each sample was brought back to room temperature purified using SPRI beads (Beckman Coulter SPRIselect) following manufacturers protocol at a 1.2:1 bead to sample ratio for fragmented DNA purification. DNA was eluted into 40µl low-EDTA TE buffer and stored at 20°C for future bisulfite conversion.

MOWChIP

Our lab's in-house MOWChIP procedure consists of a number of steps and was performed as previously reported. First, each microfluidic channel was filled with IP buffer at 50µl/min with sporadic opening and closing of the control valve (30 PSI) to remove bubbles. Protein A beads, previously coated with anti-H3K27ac (39135, Active Motif) antibodies and incubated overnight, were injected by pipette into the device chamber. Beads were collected by magnetic bar, and the tubing assembly (~24" PFA tubing attached to 1ml syringe) containing ~120µl of chromatin collected from ~10,000

sorted NeuN+ nuclei was inserted into the device inlet. With the control valve closed, a magnet was applied to the bottom of the device to beads, and the sample was injected by syringe pump at 1.5µl/min into the device. After an initial 2 minutes of sample injection, the magnet was moved to the front of the control valve, and beads were packed against the valve for ~5 minutes. Once packed, the magnet was removed and the sample was fully pumped into the device, followed by 7.5µl IP buffer. PFA tubing assembly was then removed from the inlet and 2 short (3cm) C-flex tubing assemblies attached to the pressure control system and containing ~40µl low salt wash solution were inserted into the device inlet and outlet. Oscillatory washing was performed by the alternating pressurization of the inlets and outlets as controlled by LabVIEW program (300 seconds, 0.5 second cycle duration, 0.6 PSI). Upon completion, the tubing was replaced with C-flex tubing loaded with high salt wash solution and samples were washed a second time. The tubing was then removed, and the long PFA tubing was once again inserted into the device inlet and a 4” PFA tubing elbow was placed in the outlet. IP buffer was pumped into the device at 200µl/min, with beads deposited into a fresh 1.5ml microcentrifuge tube through the outlet tubing.

For DNA isolation, tubes were first placed on a magnetic rack for 1 minute to collect beads and supernatant discarded. Magnetic beads were resuspended in 100µl elution buffer (10mM Tris-HCL (pH 8.0), 50mM NaCl, 10mM EDTA, 1% (wt/vol) SDS, 0.1M $NaHCO_3$ in ultrapure H_2O) along with 4µl proteinase K (1mg/ml). Samples were mixed by pipetting and incubated for 8 hours at 65°C for protein digestion. DNA was extracted with phenol chloroform and purified by ethanol precipitation with the final DNA product reconstituted in 5µl low-EDTA TE buffer.

Construction of ChIP-seq libraries

ChIP-seq libraries were constructed from 10,000 nuclei replicates using ThruPLEX DNA-seq kit (Takara) following manufacturers recommended protocol with slight modification. Input libraries were generated from 1,000 nuclei for each condition. Briefly, 5µl of purified DNA was mixed with 1µl template preparation D buffer and 0.5µl template preparation D enzyme, before undergoing template preparation reaction on a thermocycler (22°C for 25 minutes, 55°C for 20 minutes, 4°C for up to 2 hours). Library synthesis was then performed by adding 0.5µl library synthesis D buffer and 0.5µl library synthesis D enzyme and incubating at 22°C for 40 minutes. An amplification mixture of 12.5µl library amplification buffer, 0.5µl enzyme, 0.75µl H2O, and 1.25µl Evagreen fluorescent dye was added to the library synthesis product and amplified for 10-12 cycles following recommended protocol for library amplification reaction on qPCR. The amplification product was purified and size-selected using SPRI beads (Beckman Coulter SPRIselect) following manufacturers protocol at a 0.75:1 bead to sample ratio, with final product eluted into 10µl low-EDTA TE buffer.

Construction of RNA-seq libraries

RNA-seq libraries we're constructed from RNA extracted from 10,000 sorted nuclei per replicate following Smartseq2 protocol with slight modifications[64, 67, 123]. Reverse transcription mix (5% PEG 8000, 0.5U/ul RNase inhibitor, 0.5µM OligodT30, 10mM dNTPs, 25mM Tris-HCl, 30mM NaCl, 2.5mM MgCl2, 1mM GTP, 8mM DTT, 2µM TSO,

2U/µl Maxima H-minus RT enzyme, 0.75µl nuclease free H2O) was added to 12.5µl purified RNA and incubated in a thermocycler at 42°C for 90 minutes, 10 cycles of 50°C for 2 minutes followed by 42°C for 2 minutes, then 85°C for 5 minutes. Following reverse transcription, 25µl HiFi Hotstart mix, 0.1µM ISPCR primers, 2.5µl 20X Evagreen dye, and 2µl nuclease free H2O were added to the sample mix cDNA preamplification was performed on qPCR following suggested amplification protocol. Replicates were bead cleaned at a 1:1 ratio with SPRI beads and eluted into 5µl low-EDTA TE buffer. Libraries were then generated from cDNA using in-house produced Tn5 and Nextera XT indices as previously described[124, 125]. Following final amplification step, libraries were bead cleaned at a 0.8:1 bead to sample ratio with SPRI beads and eluted into 10µl low-EDTA TE buffer for sequencing.

Construction of WGBS libraries

Bisulfite conversion was performed on DNA extracted from 20,000 nuclei per replicate (2 replicates per sample). Starting DNA input was quantified using nanodrop with pico-green dye and analyzed for fragment size quality using high sensitivity DNA analysis kit (5067-4626, Agilent) on a TapeStation system (2200, Agilent). For each replicate, λ-phage DNA (D1521, Promega) was added at 0.5 wt% for bisulfite conversion quality control. Bisulfite conversion was performed using EZ DNA Methylation Gold Kit (Zymo Research) following manufactures suggested protocol. Briefly, replicates were brought to 20µl with nuclease free H2O then mixed with 130µl of prepared CT conversion reagent solution. Each replicate was then incubated on a thermocycler at 98°C for 10 minutes, 64°C for 2.5 hours, then 4°C for up to 20 hours. Samples were then purified,

desulphonated, and washed on Zymo-Spin IC columns. Bisulfite converted samples were then eluted into microcentrifuge tubes in 10µl of M-Elution buffer. Sample were brought to 15µl with low-EDTA TE buffer then immediately underwent library preparation.

WGBS library preparation was performed using xGen Methyl-seq kit (IDT) using manufacturers recommended protocol. All libraries were amplified 10-12 cycles before purified with 0.85X SPRI bead ratio, then eluted into 10µl Low-EDTA TE buffer.

Library quality control and sequencing

Libraries were assessed for quality control purposes prior to sequencing. ChIP library enrichment was evaluated using qPCR (Bio-Rad CFX Connect) and region-specific primers (Polr2a positive, AFM negative) as previously described[61]. All ChIP, RNA, and WGBS library fragment sizes were then assessed using high sensitivity DNA analysis kit (5067-4626, Agilent) on a TapeStation system (2200, Agilent), and any primer-dimers were removed using SPRI-bead purification. Library concentrations were examined using a KAPA library quantification kit (KK4809, KapaBiosystems), and quantified libraries were pooled to 10nM. Novaseq X with paired-end 150-nt read was used to sequence each library, generating ~15 million reads for each RNA and ChIP sample, and at least 60 million reads for each WGBS sample.

ChIP-seq data processing

Read trimming was performed using default settings by Trim Galore! (Babraham Institute, https://www.bioinformatics.babraham. ac.uk/projects/trim_galore/). Raw ChIP-seq reads and input data were then mapped to mouse mm10 genome (GRCm38) using Bowtie2 (2.2.5), and peaks were called using MACS2 (2.2.7.1). In order to improve data quality, MM10 blacklisted regions (ENCODE) were removed using SAMtools[90]. Mapped reads were extended by 100bp on either side (200 bp total) using BEDtools, with reads then normalized using their respective input [91].

RNA-seq data processing

Paired-end RNA-seq data was first trimmed using trim galore!. Reads were then mapped to GRCm38 genome using hisat2 package. Read count tables for each replicate were then generated using featurecounts. Samples were then examined for quality by exon rate, alignment rate, and total gene coverage before being passed to differential analysis.

WGBS data processing

WGBS data was processed using Bismark suite with minor modifications. The paired end data was trimmed by 10 bp from both the 3' end of read 1 and 5' end of read 2 to eliminate calls from errant methylation artifacts inserted during library preparation using Trim Galore!. Reads were aligned to mm10 genome using Bismark package incorporating bowtie2. Alignment deduplication and methylation extraction were also performed in Bismark using default settings. Each replicate then underwent the identical Bismark pipeline again, this time aligning to the lambda DNA genome to assess bisulfite conversion. Only samples with >99% bisulfite conversion were kept for downstream analysis.

Differential analysis for ChIP-seq data

Each sample with two replicates passing QC standards (>10,000 peaks, FRIP > 5) were passed on to differential analysis. Male and female samples within the same condition (SAL vs. PSI) were pooled to generate the sample set passed as input into diffbind R package for generating a consensus peakset. Consensus peaksets were created using dba.peakset in diffbind (DBA_BLACKLIST_MM10, consensus = DBA_TREATMENT, minOverlap = 2). The read counts were extracted using the diffbind function dba.count (data_blacklist_remove, summits = FALSE, peaks = ConsensusPeaks, filter=1, bScaleControl = TRUE, minCount=1, score=DBA_SCORE_TMM_MINUS_FULL). Differential analysis was performed using diffbind function dba.contrast (data_analysis, categories=DBA_TREATMENT,minMembers = 2), with only differential peaks meeting

stringent criteria (FDR > 0.05, FC > 0.6) retained. Differential enhancers (defined as identified peaks that don't fall within ±2000bp of TSS) were annotated first by checking for regional overlap with published Hi-C data on mouse brain neurons[68]. Overlapped peak regions were annotated based on these existing Hi-C gene links, and the rest of the enhancers were associated with their nearest-neighbor gene using ChIPSeeker with default condition and TxDb.Mmusculus.UCSC.mm10.knownGene.

Differential analysis for RNA-seq data

RNA-seq differential analysis was performed using R package DSeq2. Input count tables were generated using featurecounts. Uniform Manifold Approximation and Projection (UMAP) was used to identify principal components with outsized impact. The following covariates were regressed out: demographic covariates (sex, condition) and technical covariates (align rate, unique reads, exon rate, the number of covered genes, and sequencing batch) by correlating the top 6 principal components (Pearson correlation > 0.7) with these covariates[94]. Differential analysis was then performed to identify differential peaks between the two conditions (FDR <0.05, log2FC < 0.6). The p-values were adjusted by performing a standard Bonferroni correction.

Differential analysis for WGBS data

Differential analysis was performed on WGBS data using R package DSS. For each sample, Bismark coverage files were merged to generate a single DSS input file per sample. A CpG site list was generated for all candidate sites with over 5x coverage in at

least half of the total samples. CpG sites in the top 0.1% of coverage were removed from the set as outliers. Samples were then pooled by group and per site average methylation percentages were extracted using getMeth function. Differentially methylated loci (DMLs) were then determined using DML fit function DMLtest.multiFactor(DMLfit, coef="groupP"). Significant DMLs were determined with an fdr < 0.05.

We then examined expanded differentially methylated regions with at least 3 differentially methylated CpGs with P < 0.05 with maximum spacing between sites of 300bp using the callDMR function. Per-site methylation of CpGs within each DMR region were aggregated for magnitude and directionality.

K-means clustering, UHC, GO term analysis, KEGG pathway analysis

K-means clustering and unsupervised hierarchical clustering (UHC) were performed on differential enhancers and DEGs. To normalize the counts based on the number of samples, each row of the count file was divided by $\sqrt{sum(X)^2}$. The optimal number of clusters for k-means was determined using the silhouette method[95]. UHC was performed using the default settings of the pheatmap library in R, with cluster position on the x-axis rotated about their dendrogram tentacle for continuity. Gene ontology (GO) biological enrichment analysis and KEGG pathway analysis was performed using shinyGO version 0.85.1[97]. Both GO and KEGG analysis were performed using FDR adjusted P values < 0.05 to ensure generation of high-confidence terms.

Taiji analysis pipeline

ChIP-seq and RNA-seq counts were processed through taiji pipeline to establish transcription factor influence across samples. Promotor regions were annotated and defined as +/- 2000bp from TSS. Significantly enriched peaks outside of these regions we're retained as putitative distal enhancers and linked to probable target genes using unsupervised learning method EpiTensor. Enhancer-promotor pairs were passed on if one interaction anchor overlapped a distal H3K27ac peak and its associated anchor overlapped a promotor region. Assumed TF binding motifs were sourced from CIS-BP database and compared with our list of active regulatory elements. TFs were established as containing at least one significant binding site within +/-75bp of a peak summit. A directed connection was assigned from a transcription factor to a gene whenever the factors motif overlapped either the gene's promotor region or an enhancer previously linked to that promotor. All downstream gene targets were identified for each TF. Only genes highly weighted (>0.5) in at least 5 samples were designated as key regulates. For sex specific comparisons, highly weighted genes (>0.5) in at least 3 samples were designated as key regulates. Highly weighted genes comprising differential TF networks were overlapped with DMR data to identify genes co-regulated by DNA methylation.

4.4 Results and Discussion

Epigenomic and transcriptomic alterations induced by maternal psilocybin exposure

Previous studies on rapid acting antidepressants (RAADs) such as ketamine, have shown that gestational exposure can induce epigenetic modifications[126, 127]. However, far less is known about whether psychedelic-induced epigenomic alterations can be transmitted through gametes[111]. Prolonged exposure to psychoactive substances such as opioids and alcohol have been associated with measurable changes in gametic DNA methylation[128-130]. Given our previous findings that psilocybin induces epigenomic alterations in mouse PFC neurons, which mirror and in some cases reverse changes elicited by other psychoactive compounds, and considering the durable nature of psychedelic-induced epigenomic states, we sought to evaluate the inherited effects of pre-gestational maternal psilocybin exposure on offspring epigenomes.

Adult female dams were treated with either psilocybin (PSI) or saline (VEH) 14 days prior to mating (Fig 4.3A). Once born, offspring were weaned at an appropriate time and caged by with their same-sex siblings. At 10 weeks, adult offspring were sacrificed and frontal cortex excised. NeuN+ neurons were isolated from sliced cortical tissue for epigenomic and transcriptomic evaluation. ChIP-seq assays targeting H3K27ac histone mark were performed using in-house MOWChIP platform allowing for sensitive evaluation from just 10,000 nuclei per assay. DNA methylation was probed by bisulfite conversion followed by whole genome bisulfite sequencing (WGBS) starting from 20,000 nuclei per assay. Gene expression within each mouse was evaluated using Smartseq3 protocol with 10,000 nuclei as input. Technical replicates were generated from each sample so that each assay was performed twice.

To ensure high quality histone peak and gene expression profiling, ChIP-seq and RNA-seq samples were sequenced to at least 15 million reads per replicate. WGBS samples

were sequenced to at least 60 million reads per replicate for high per-sample CpG coverage. For quality control, ChIP-seq samples with low FRIP, peaks, or unique reads were filtered out of the dataset, ensuring high quality peaks with minimal background noise (average FRIP = 12.8%). Similarly, samples with low gene, coverage, total reads, and exon rate were removed from the transcriptomic dataset, while replicates with incomplete bisulfite conversion (<99% by lambda DNA methylation) were removed from WGBS analysis. High quality datasets were visually inspected using integrative genomic viewer (Fig 4.3B).

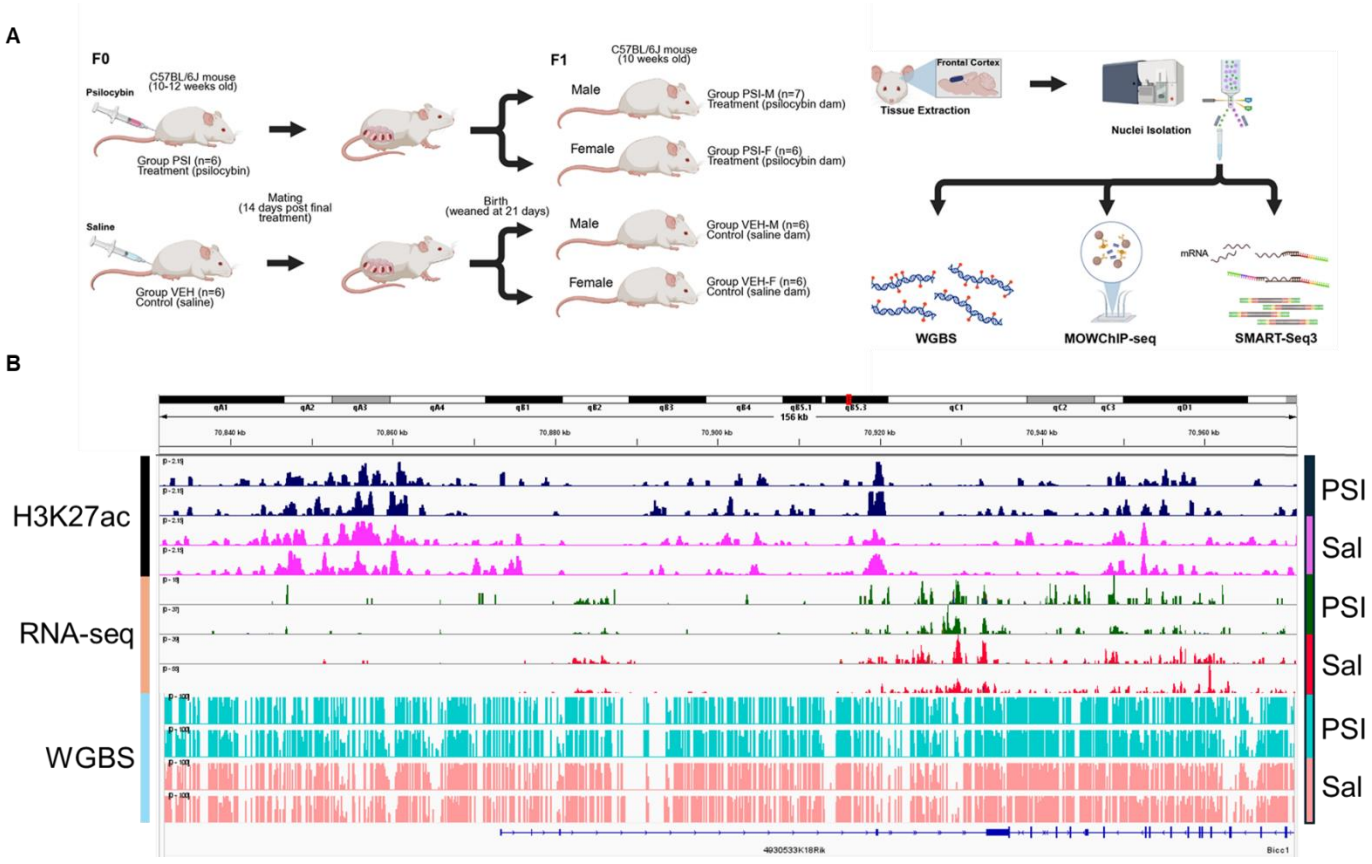

Fig 4.3. General overview of (A) Experimental design (B) data collected from sequencing, as represented by IGV, including H3K27ac ChIP peaks, DNA methylation status, and gene expression counts.

To examine H3K27ac histone binding alterations between maternal psilocybin exposure states, ChIP-seq samples between both cohorts were mapped and peaks overlapped to generate a consensus peak set. In total, the generated consensus peak set contained 130,146 peaks. Differential analysis was performed from the peakset using Diffbind R package. This analysis identified 735 high confidence significant differential peaks, with 619 of these peaks displaying increased binding activity in the cohort with psilocybin treated mothers (Fig 4.4A). Peaks were overlayed with previously published Hi-C chromatin conformation data on mouse cortical neurons to identify direct gene links for enhancer peaks. Peaks not connected through Hi-C data were annotated to nearest neighbor gene. Peaks within promotor regions (+/- 2000bp of TSS) were removed from the set to generate a final group of 659 differentially bound enhancer linked genes. GO term analysis of enhancer linked genes showed enrichment in neuron related biological processes in both increased and decrease activity sets, as expected from PFC neuron samples.

To examine how epigenomic alterations may impact downstream gene expression, RNA-seq count data from both cohorts was compiled and subjected to differential analysis using DESeq2 R pipeline. In total, 606 genes were found to be differentially expressed between PSI and VEH groups, with a larger share (451 genes) downregulated in offspring with psilocybin compared to controls (Fig 4.4B&C). As the H3K27ac enhancer mark is generally associated with open chromatin states, the trend

towards downregulation in our sample set from VEH to PSI suggests multiple modes of gene regulation at work. This is strengthened by examination of overlap between differential enhancer linked genes and differentially expressed gene sets (Fig 4.4D). In total, only 4 genes display altered H3K27ac binding activity and differential RNA expression. *Steap4* and *C730036E19Rik* display corresponding increases in expression trends in PSI compared to controls. *Vmn2r104* and *4933433F19Rik* show opposite trends, with the former displaying increased genes expression and decreased enhancer activity, while the later displays decreased expression and increased H3K27ac binding. As such, it is clear the effect of enhancer activity alterations on nearest neighbor genes are not the key driver of observed changes in expression.

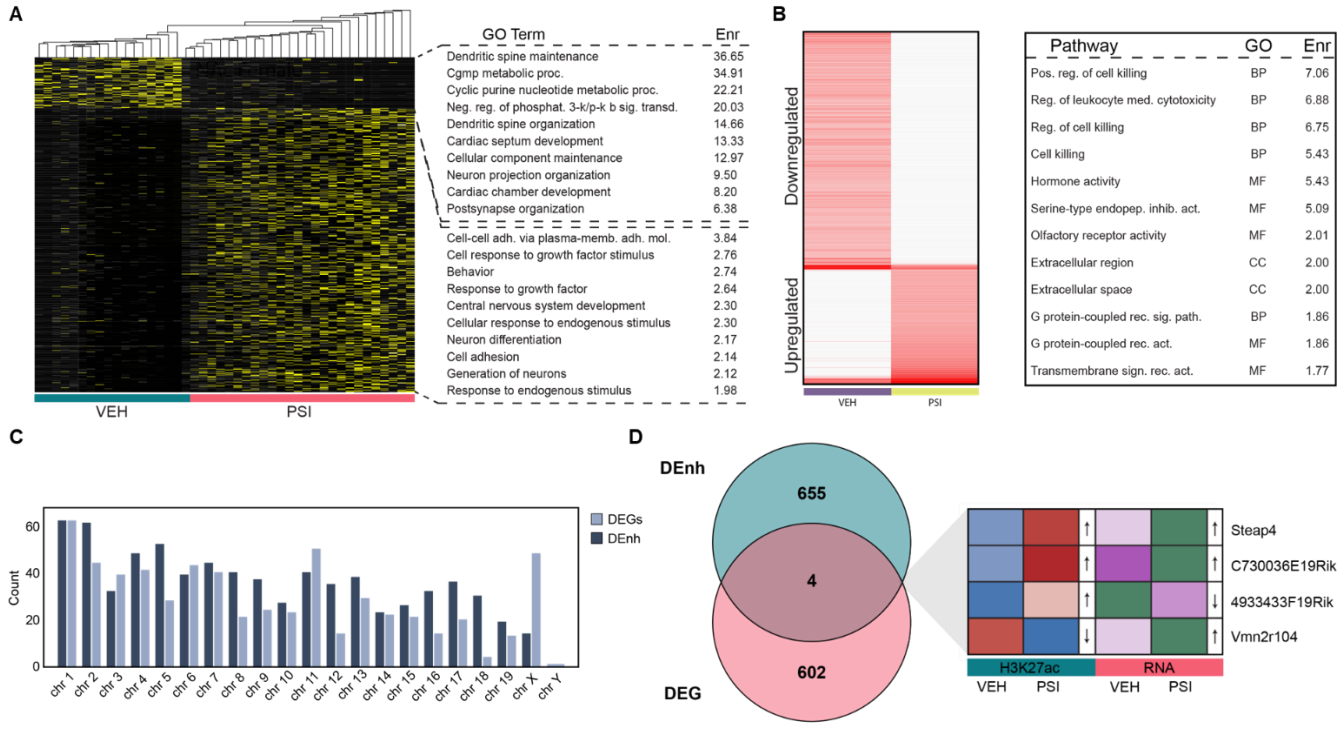

Fig 4.4. Epigenomic and transcriptomic alterations following maternal psilocybin exposure (A) Differential analysis of H3K27ac ChIP-seq peaks and biological

processes highly enriched in associated genes dependent directionality of enhancer expression change. (B) Differential gene expression and highly enriched pathways/processes, (C) total differential enhancer linked genes and differentially expressed genes by chromosome. (D) Overlap of differentially expressed enhancer associated genes and DEGs.

DNA methylation states were examined as an alternative epigenomic mechanism underpinning maternal psilocybin induced gene expression changes. Replicates were combined to generate a per-sample methylation set and differential analysis was performed using DSS package in R. In total, 435 significant differentially methylated regions were identified between PSI and VEH cohorts (Fig 4.5A). DMRs were tended to localize at intergenic and intronic regions and between 10-100kb from transcription start sites, with 9.7% found at gene promotor regions (Fig 4.4B). Overall, these regions were split evenly between hypermethylated (312 DMRs) and hypomethylated (123 DMRs) states with a slight preference for the former. This is in line with previously reported findings on direct methylation differences in PFC of psychedelic exposed mice which favored hypermethylated states as opposed to hypomethylation[110]. DMRs were annotated to their nearest neighbor gene before undergoing GO term analysis. Regions of significant increases in methylation were found at genes enriched in terms associated with neuronal regeneration and stem cell division (Fig 4.4D). Hypomethylated DMRs were enriched in genes associated with cellular and synaptic communication and its regulation, as well as in neuronal growth and creation (Fig 4.4E).

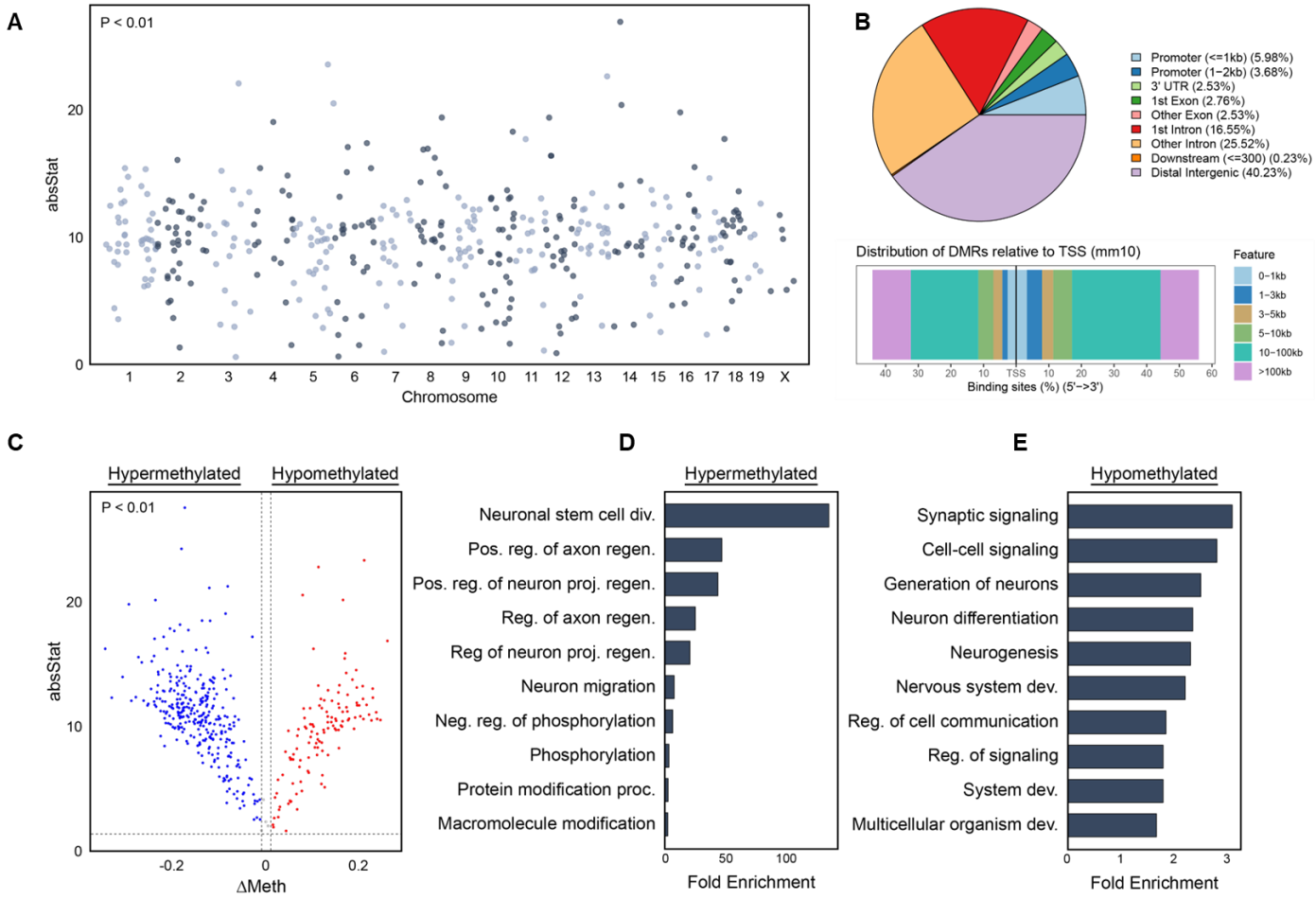

Fig 4.5. DNA methylation alterations between offspring with and without maternal psilocybin exposure (A) Chromosomal distribution of DMRs. (B) Distribution of genetic function and genomic positioning of significant DMRs. (C) Directionality of methylation status at significant DMRs. Top enriched biological process gene ontology (GO) terms associated with genes with (D) hypomethylated and (E) hypermethylated DMRs.

Mechanism of epigenomic inheritance

Due to the thorough resetting of epigenetic reprogramming in mammals during gametogenesis and post-fertilization, the prevalence of inherited epigenetic modifications has been hotly debated[131, 132]. As such, it is imperative to identify mechanisms of epigenomic signal transmission leading to observable inheritance. Replicative transmission through allele-specific DNA methylation at imprinted regions has proven to be amongst the most effective and well characterized methods of preserving parent-specific gene expression programs during development[133, 134].

We thus examined the role of replicative transmission within our observed DNA methylation data at previously reported imprinted gene regions (IGRs) to evaluate its impact[134]. Significant differentially methylated regions between PSI and VEH cohorts were found at 6 canonical imprinted mouse genes (Fig 4.6A). Of these, 5 IGRs exhibited significantly decreased methylation state in offspring of mothers exposed to psilocybin compared to controls. *Pde4d*, the only IGR displaying increasing methylation status, is critical to learning and long-term memory creation within the prefrontal cortex and synaptic health[135, 136]. Additionally, methylation at the gene promotor plays a key role in regulating Pde4d transcription[137]. Increased *Pde4d* promotor methylation is associated with decreased expression of the gene, with transcriptional inhibition leading to observed beneficial memory effects.

Of the 6 differentially methylated canonical imprinted genes *Mafb*, *Mirg*, and *Arid1b* are specifically linked to maternal alleles[138, 139]. All three of these genes are heavily linked to brain development with involvement in neural migration, dendritic spine development and synaptic plasticity. Disruption of expression of these genes, including in the frontal cortex, has been heavily linked to altered social behaviors in adults including autism

spectrum disorder and intellectual disability[140-142]. As these IGRs are associated strictly with maternal alleles, the observed aberrant DNA methylation patterns can be attributed with high confidence to the effects of psilocybin treatment.

When observed as a group, KEGG analysis of the 6 IGRs containing significantly differentially methylated regions demonstrates a single highly enriched pathway (Fig 4.6B). While no direct evidence of psychedelic exposure altering parathyroid hormone (PTH) function exists, the ability of PTH to cross the blood-brain barrier has implications on cognitive processes. For example, increased PTH activity (hyperparathyroidism) has direct involvement in impaired cognitive function and dementia[143]. Although the IGRs discussed here have clear phenotypic implications for mouse neurobiology, the limited overlap between our DMR set and canonical imprinting regions suggests additional mechanisms of epigenomic signal transmission are likely involved.

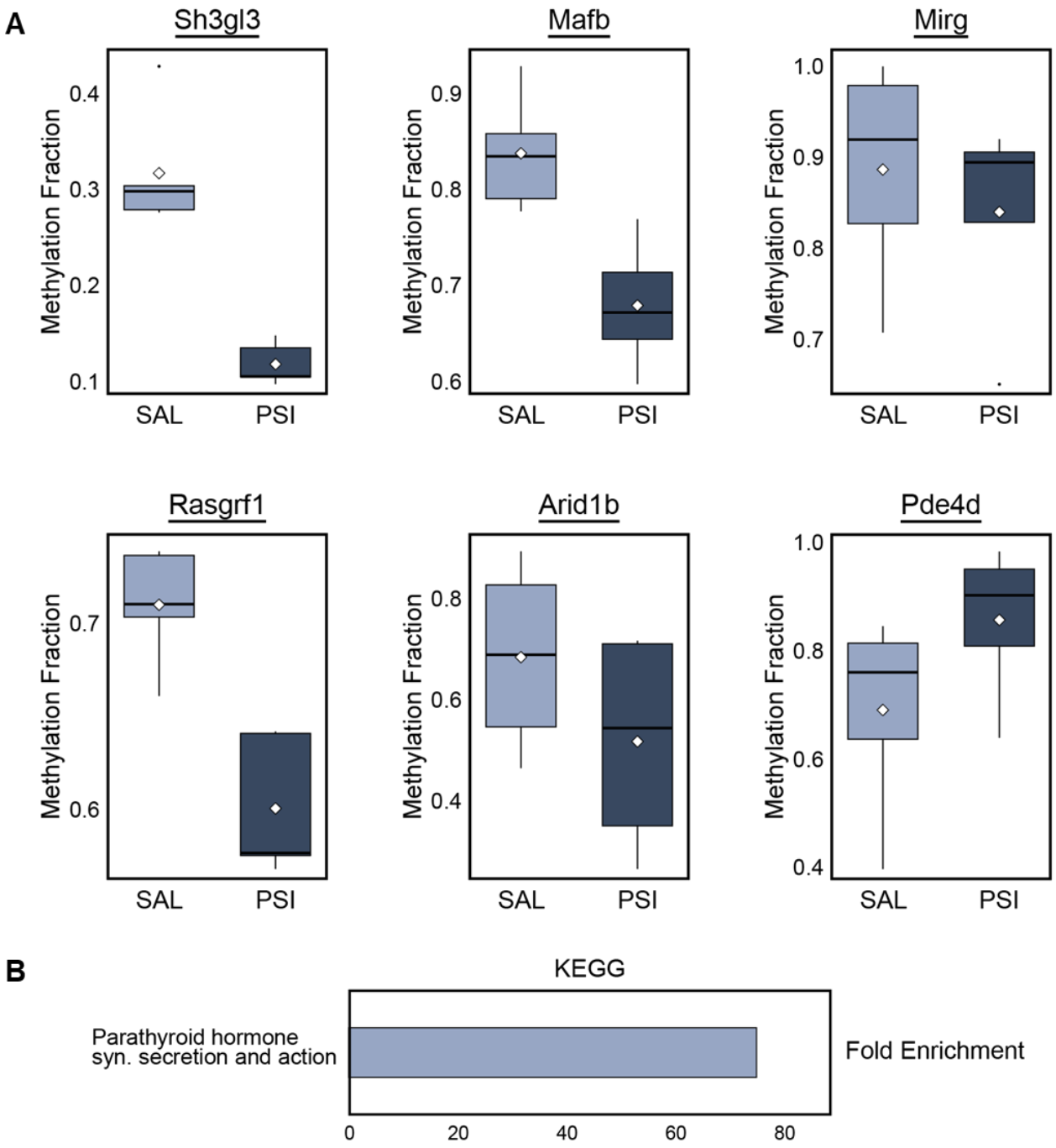

Fig 4.6. DNA methylation alterations at genes with canonically imprinted gene regions (IGRs) (A) Mean methylation status of samples at DMRs between offspring with and without maternal psilocybin exposure found at genes with known canonical

IGR. (B) Overrepresented pathway associated with significantly differentially methylated IGR associated genes.

Transcription factor alterations induced by maternal psilocybin exposure

To pinpoint regulators responsible for the transcriptional shifts in our dataset, we utilized Taiji, a systems-level method that couples epigenomic marks with expression readouts to rank transcription factors by their predicted regulatory reach. This enabled the generation of high resolution, sample-specific networks linking active regulatory elements to their downstream targets. Starting from matched ChIP-seq and RNA-seq data for each sample, active transcription factors and associated regulatory networks were predicted and ranked by Personalized PageRank (PPR) algorithm. Differentially expressed TFs (FDR > 0.05) between maternal PSI and SAL cohorts were determined by absolute PPR difference (Fig 4.7A). We identified 11 TFs with significant differentially expressed regulatory networks between the two cohorts. Included in this group is the *Egr2* early gene transcription factor, among the most heavily linked genes to psychedelic induced alterations in the brain. Here, we observe an overall decrease in *Egr2* regulatory impact in frontal cortex neurons of mice with psilocybin treated mothers compared to controls. Psychedelic exposure has previously been shown to alter *Egr2* transcription in neurons of this brain region, as well as long-term decreases in predicted *Egr2* motif-linked enhancer expression[64, 114]. Previous studies have also shown significant regulatory alterations of the plasticity-related TF gene *JunB* in mouse prefrontal cortex following psilocybin exposure in a dose-dependent manner[144]. Here, we see an overall decrease in JunB regulatory function in the PSI group compared to

control cohort. Additionally, we observe increases in the PSI group of *ZFP281*, *NFIC*, and *Hnf4g* TFs which play important roles in neuron differentiation, development, and function[145-147].

Differential TFs were parsed for their gene-regulatory networks for highly effected downstream expression targets (TF-target weight >0.5, present in at least 5 samples). The gene-regulatory network yielded 114 genes with significant changes in their regulation by the TFs in our differential set (Fig 4.7B), specifically derived from *EGR2* and *RREB1*. KEGG analysis showed affected genes were highly enriched in signaling pathways including CGMP-PKG, cAMP, and Ca^{2+}, along with long-term potentiation, dopaminergic synapse, and amphetamine and nicotine addiction related terms (Fig 4.7C). These signaling pathways are key components of neuronal transmission and plasticity in the brain and have a direct relation to reward-seeking behaviors observed in addiction[148-150]. Moreover, psilocybin is well documented to modulate genes that support long-term potentiation, promoting synaptic strengthening with consequent enhancements in learning and memory. Collectively, these results reveal an altered regulatory landscape in the frontal cortex, characterized by diminished transcription factor activity at genes associated with neurological processes known to be directly impacted by primary psychedelic exposure, and demonstrate that this state is transmitted to the F1 generation. Additionally, the overlap between altered signaling pathways and reward-related behavioral terms observed here and in our previous study involving psilocybin treatment for opioid addiction suggests that mechanisms implicated in the therapeutic effects of psilocybin in combating OUD are also perturbed in a downstream generation, albeit in an opposing direction.

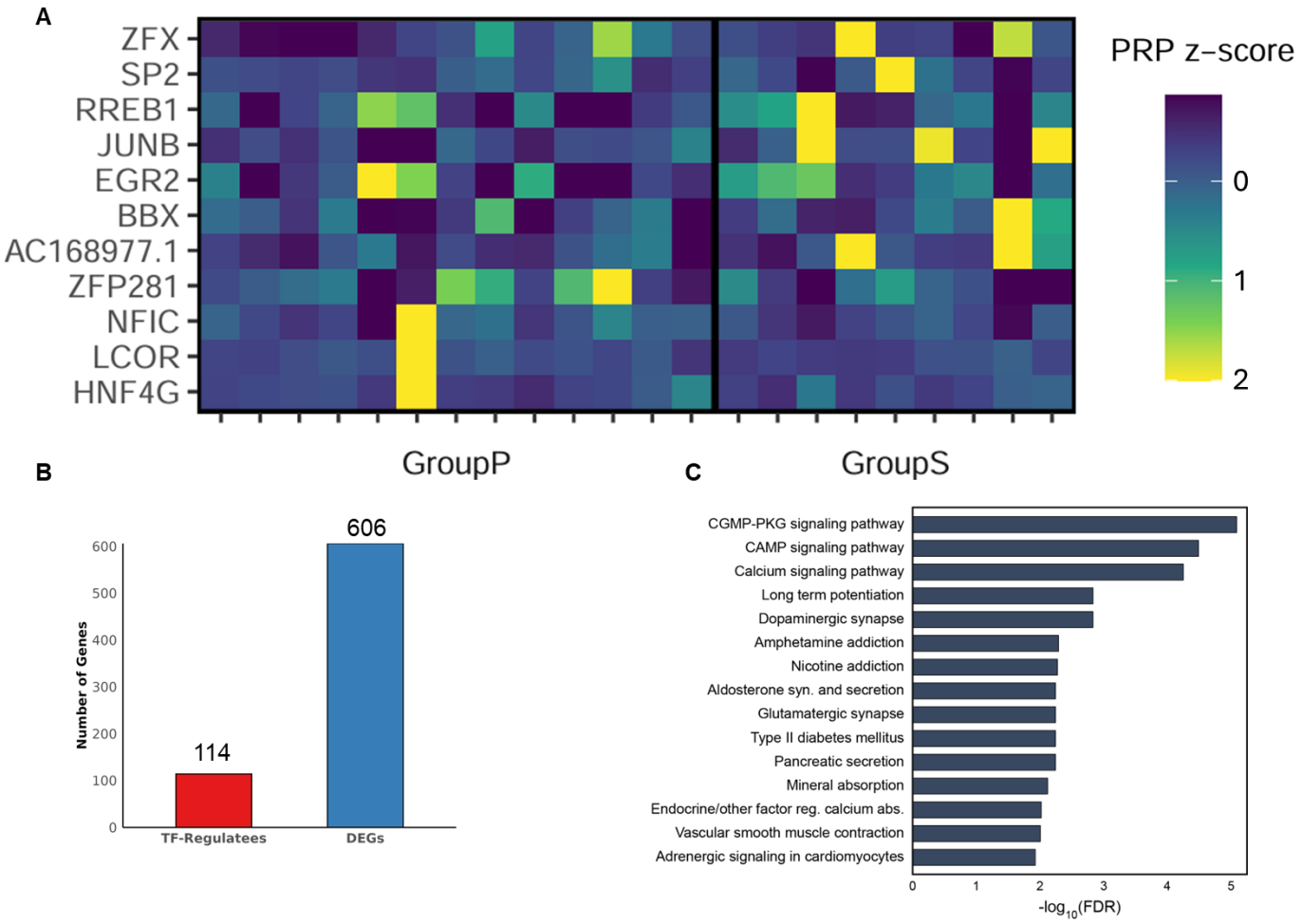

Fig 4.7. Transcriptional regulatory processes in the frontal cortex of offspring with and without maternal psilocybin exposure. (A) Heatmap of significant differential transcription factors between offspring with and without maternal psilocybin exposure. (B) total number of highly weighted TF-regulatee gene edges derived from differential transcription factors compared to overall differential gene expression. (C) Overrepresented pathways significantly enriched among the differential transcription factor linked regulatee genes.

Transcription factor binding is amongst the most common reconstructive mechanisms of epigenomic signal transmission and is believed to be more prevalent in mammals than

replicative methods. As an example, binding of TFs to differentially methylated loci can hinder methylation during global methylation and thus transmit epigenetic signatures through generations. Additionally, methylation status can alter binding of targeting TFs and generally co-regulate expression of downstream genes. As such, differential methylation status of genes in our identified TF-regulatee network was examined to identify instances of multimodal epigenetic control affecting transcription. Of the 114 genes in our differential TF regulatory network, 13 contained at least one significant differentially methylated region (Fig 4.8A&B). Of these, *ANK3*, *TCF4*, and *CELF2* genes are coregulated by *EGR2* and *RREB1* transcription factors while carrying DMRs. Specifically, the PSI cohort of mice experience decreased activity from both *EGR2* and *RREB1* transcription factors, while simultaneously displaying increased methylation at *ANK3* and *TCF4* genes. Both *ANK3* and *TCF4* genes are heavily involved in dictating neuronal polarity and development, with repression of expression leading to neurological disorders such as schizophrenia and bipolar disorder.[151, 152] Previous studies have also shown a direct relationship between increased DNA methylation at *ANK3* gene with decreased transcription at long time scales (>60 days) in rodent PFC tissue[153].

Taken as a whole, the set of 13 differentially methylated TF-regulated genes is highly enriched in the same top signaling pathways (CGMP-PKG, CAMP) as the set not accounting for methylation status. However, this differentially methylated set is uniquely enriched in genes relating to the oxytocin signaling pathway and morphine addiction. As mentioned previously, our prior work shows the effects of psilocybin on frontal cortex neuron epigenome, with oxytocin signaling pathways a posited route of observed

behavioral phenotype alterations. Interestingly, this gene set is also enriched in the KEGG term (parathyroid hormone synthesis secretion and action) as found amongst observed differentially methylated imprinted genes. This is partially due to the inclusion of the sole hypermethylated DMR containing imprinted gene, *Pde4d*, in this list. We observe here the co-regulation of this gene by decreased *EGR2* TF activity with increased DNA methylation, as well as the intersection of two potential modes of epigenomic signal transmission, replicative and reconstructive.

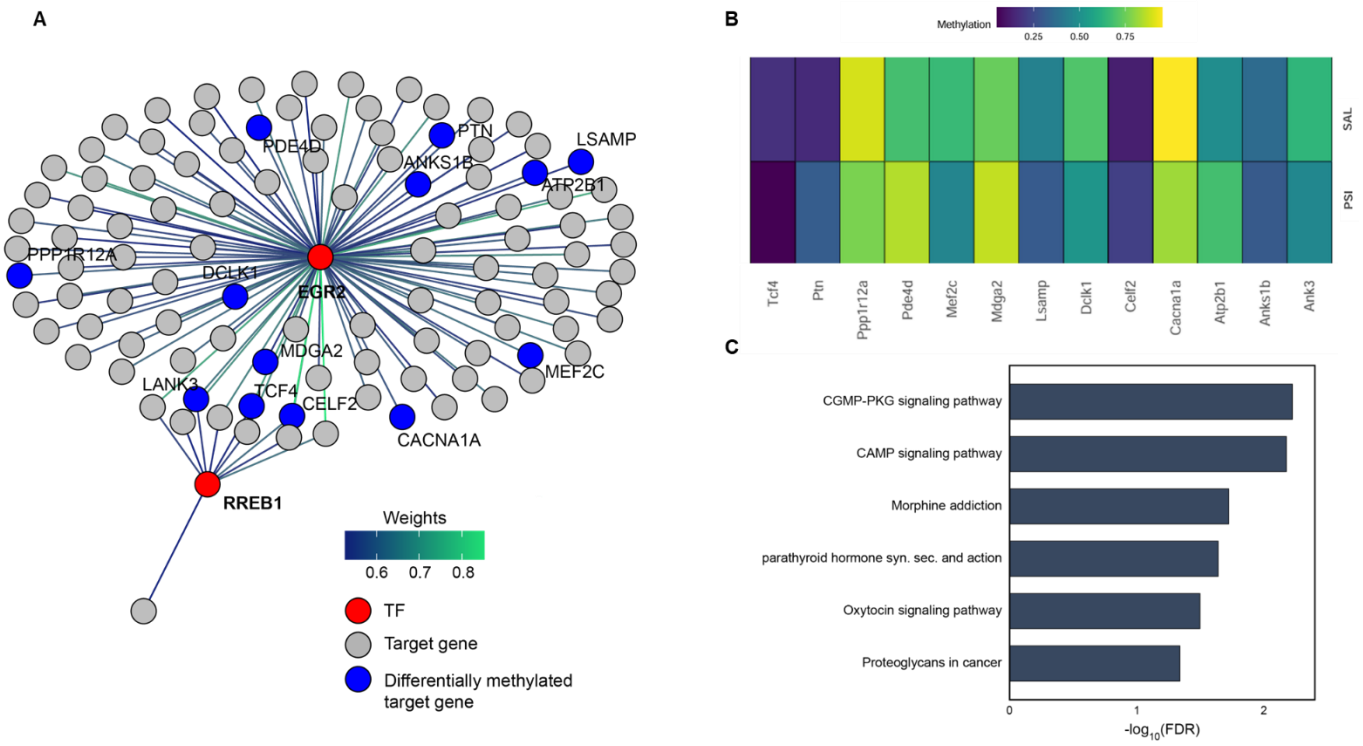

Fig 4.8. Downstream DNA methylation alterations at transcription factor regulated genes. (A) Gene regulatory activities between differential transcription factors and highly weighted target genes (B) Heatmap of methylation status at DMRs located at regulatee genes of identified differential transcription factors (C) Overrepresented

pathways significantly enriched among differential transcription factor linked regulatee genes containing DMRs.

Sex-specific inherited epigenomic alterations

The previously discussed study from our lab observed significant sex-specific psilocybin induced epigenomic alterations in the mouse frontal cortex. We thus sought to determine the extent to which these sex-specific alterations are inherited by the offspring of treated mothers. Differential analysis between sex-matched cohorts of offspring based on maternal psilocybin exposure status were performed using H3K27ac ChIP-seq, WGBS, and RNA-seq data (Fig 4.9). Differential epigenetic and transcriptomic mark dispersion was relatively even between the male and female genomes. Male cohorts experienced significantly increased DMR localization in chr 4, with a corresponding increase in differential gene expression. Additionally, chromosome 11 experienced the most significant variation in differential gene expression, with increased localization in male cohorts. This was associated with decreased DMR localization in males compared to females at sites along this chromosome.

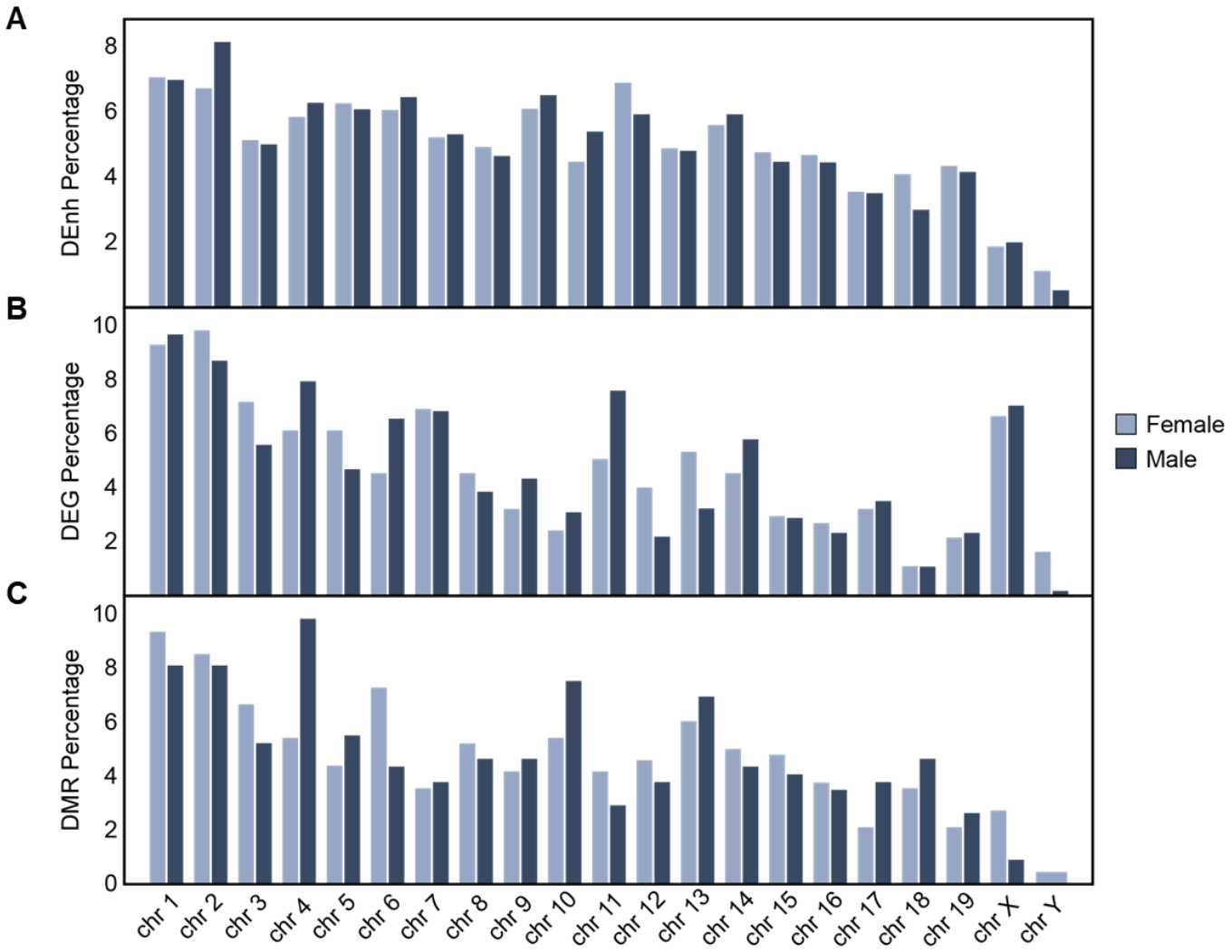

Fig 4.9. Sex-specific epigenomic and transcriptomic alterations in frontal cortex neurons of offspring with and without maternal psilocybin exposure. (A) Distribution of differential enhancer linked genes by chromosomal position between male and female subjects. (B) Distribution of differentially expressed genes by chromosomal position. (C) Distribution of differentially methylated regions by chromosomal position.

Transcription factor regulatory networks were built out to discern the epigenomic and transcriptomic relationship arising within male and female groups. In total, we identified

7 differentially page-ranked transcription factors (Fig 4.10A) between males and 8 differential TFs between females (Fig 4.10B). Both male and female cohorts displayed decreased *RREB1* regulatory function in cohorts where the mother received psilocybin compared to saline injections. The other shared differentially expressed TF within intragender comparisons, *Tcfap2a*, showed opposite activity trends between sexes, increasing in the maternal PSI cohort in males while decreasing in the maternal PSI cohort in females. The *Tcfap2* transcription factor is crucial for neuronal development and maturation throughout the nervous system[154, 155].

Unique differentially page-ranked transcription factors between male cohorts include *Zfx*, *Sp2*, *Nfic*, *Lcor*, and *Hnf4g*, all of which were altered in the non-sex discriminatory comparison. As these TFs are not differentially expressed within female cohorts, we can say their inclusion in the sexless comparison (Fig 4.7A) derives from a difference in male subjects. Similar investigation of differentially expressed transcription factors within the female cohorts identifies female specific contributions to the sexless comparison, specifically *Egr2*, *Bbx*, and *Zfp281*. Additionally, several TFs are differentially page-ranked between female cohorts that were absent from the sexless comparison. These include *Tcfap2d*, *Foxj3* and *Egr1*, all of which exhibited increased regulatory function in the maternal psilocybin group. *Egr1* belongs to the family of immediate early genes whose expression is significantly altered in cortical regions by psychedelic compounds such as psilocybin and LSD. The serotonin receptor 5-HT2a upon which psilocybin acts to illicit hallucinogenic effects is the primary mediator of Egr1 expression in the frontal cortex[156, 157]. Transcriptional regulation by *Egr1* TF plays an important role in brain development and synaptic plasticity, and its increased activity is posited to be a key

driver of observed therapeutic properties of psychedelic compounds[114, 158, 159]. The increased activity observed in PSI females, comparable to the effects reported in studies on directly treated subjects (F0), supports the idea that psilocybin exposure may alter downstream receptiveness to positive neurogenic outcomes in subsequent generations.

Differential transcription factor regulated gene networks were generated for both male and female cohorts. Differentially ranked transcription factors in females displayed highly weighted downstream regulator effects on 85 genes, while similar analysis on male mice revealed just 8 genes. KEGG analysis of downstream regulated genes show no sex-specific differentiation in highly enriched pathways associated with long-term potentiation, synapses (glutamatergic and dopaminergic), and addiction (Fig 4.10C). However, this analysis does reveal some significant enriched pathways unique to each sex. In females, highly weighted genes within the altered regulatory network were distinctly enriched in neurogenesis promoting CGMP-PKG and Ca^{2+} signaling pathways. The calcium signaling pathway in the frontal cortex plays an important role in maintaining synaptic plasticity and is a key regulator of genes that promote neuronal growth[160, 161]. Similarly, the cGMP-PKG signaling pathway is integral in neuroplasticity and synaptic function fundamental to executive function in the prefrontal cortex[162-164]. In contrast, genes affected by the altered TF regulatory activity in the male cohort were uniquely enriched in terms relating to neuron deterioration, including long-term depression, pathways of neurodegeneration, and the neurodegenerative disease amyotrophic lateral sclerosis (ALS). Overall, the contrasting pathway enrichments point to sex-specific inherited regulatory consequences of maternal psilocybin exposure, with

female networks aligning with neuroplasticity promoting processes and male networks aligning with signaling programs linked to compromised neuronal states.

Finally, we assessed sex-specific DNA methylation differences across maternal psilocybin treatment and control groups. Females exhibited greater methylation variation, with 476 identified DMRs compared to 345 in males (Fig S4.1). Differential TF regulated genes were then assessed for overlap with DMRs identify sex-specific multimodal regulation. Female cohorts exhibited DMRs in 8 network genes (9.4% of TF-regulated targets). In contrast, the male regulatory network showed no significant DMRs at any TF-regulated loci. While noteworthy, the relatively high number of DMRs observed in each sex, together with the small number of genes passing weight thresholds within the male transcription factor regulatory network, suggests that the difference in identifiable co-regulated genes is driven primarily by limited gene inclusion in the network rather than by a substantial sex-specific difference in DNA methylation.

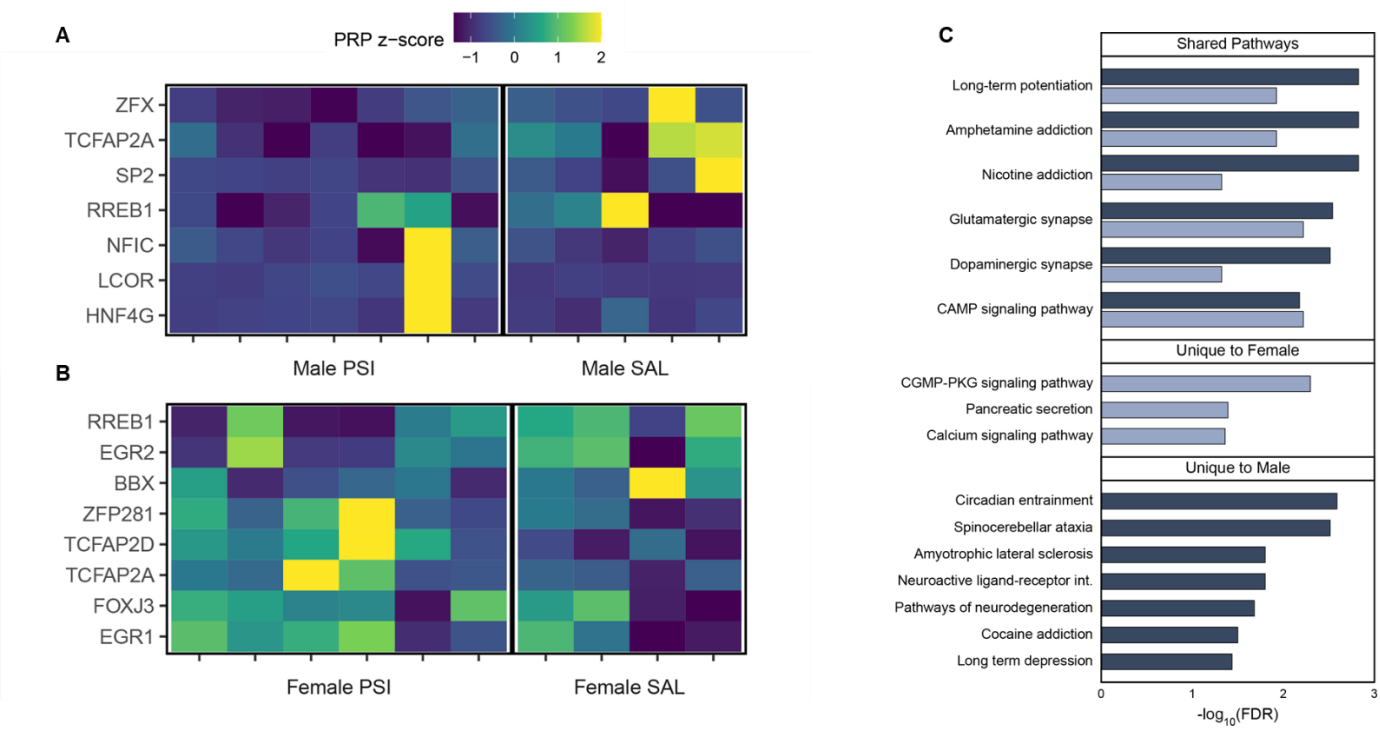

Fig 4.10. Role of sex in transcriptional regulatory processes altered by maternal psilocybin exposure. (A) Heatmap of significant differential transcription factors between male offspring with and without maternal psilocybin exposure. (B) Heatmap of significant differential transcription factors between female offspring with and without maternal psilocybin exposure. (C) Overrepresented pathways significantly enriched among the differential transcription factor-linked regulatee genes and their uniqueness to each sex cohort.

4.5 Conclusion

Although the scope of transmission mechanisms remain a topic of lively debate, evidence of observable inherited epigenomic marks has grown steadily across multiple model organisms and contexts. Yet while affects have been observed following acute

and prolonged administration of various drugs of abuse and rapid acting anti-depressants, we present here the first study identifying inherited alterations following psychedelic treatment. Our study demonstrated active enhancer and DNA methylation alterations induced by prolonged maternal psilocybin use. However, the direct transcriptional impact of these marks is difficult to elucidate with traditional differential analysis tools. Epigenomic alterations regulate genes in combinatorial mechanisms, and often far afield of neighboring sites due to chromosomal structuring. We thus utilized putative transcription factor motifs and downstream annotations in conjunction with our collected enhancer, DNA methylation, and gene expression data to uncover downstream multimodal regulatory changes.

Our analysis uncovered significant immediate early gene TF regulatory regression within offspring of maternal psilocybin users. Specifically, female offspring displayed diminished *EGR2* regulatory reach, affecting 113 genes. While psychedelic binding of the 5-HT_2A receptor is known to regulate only a limited number of genes in the brain, immediate early genes such as *Egr2* are amongst those repeatedly shown to sustain increased expression. As such, it is believed this gene plays an important role in the neuroplastic effects of psychedelic compounds in cortical structures. This diminished regulatory reach of *Egr2* TF may limit the susceptibility of maternally treated offspring to the positive neurogenic effects of primary psychedelic exposure. Further investigation of downstream responses following primary psilocybin exposure in maternally treated offspring will be essential to confirm the magnitude and functional relevance of this inherited modification.

4.6 Supplemental Figures

Cage	Maternal Ear tag #	Status	Sex (M/F)	Maternal Drug Treatment	DOB
126	41	Offspring	F	Psilocybin	9/26/2023
126	41	Offspring	F	Psilocybin	9/26/2023
126	41	Offspring	F	Psilocybin	9/26/2023
127	41	Offspring	M	Psilocybin	9/26/2023
127	41	Offspring	M	Psilocybin	9/26/2023
128	41	Offspring	F	Psilocybin	9/26/2023
128	41	Offspring	F	Psilocybin	9/26/2023
129	42	Offspring	M	Psilocybin	9/26/2023
129	42	Offspring	M	Psilocybin	9/26/2023
130	42	Offspring	M	Psilocybin	9/26/2023
130	42	Offspring	M	Psilocybin	9/26/2023
130	42	Offspring	M	Psilocybin	9/26/2023
130	42	Offspring	M	Psilocybin	9/26/2023
887	45	Offspring	F	Saline	9/26/2023
887	45	Offspring	F	Saline	9/26/2023
888	45	Offspring	M	Saline	9/26/2023
888	45	Offspring	M	Saline	9/26/2023
889	46	Offspring	F	Saline	9/27/2023
889	46	Offspring	F	Saline	9/27/2023

890	46	Offspring	M	Saline	9/27/2023
890	46	Offspring	M	Saline	9/27/2023
891	44	Offspring	F	Saline	9/27/2023
891	44	Offspring	F	Saline	9/27/2023
891	44	Offspring	F	Saline	9/27/2023
893	44	Offspring	M	Saline	9/27/2023
893	44	Offspring	M	Saline	9/27/2023
894	43	Offspring	F	Psilocybin	9/28/2023
894	43	Offspring	F	Psilocybin	9/28/2023
896	43	Offspring	M	Psilocybin	9/28/2023
896	43	Offspring	M	Psilocybin	9/28/2023

Table 4.1 Information on maternal psilocybin treated and non-treated offspring

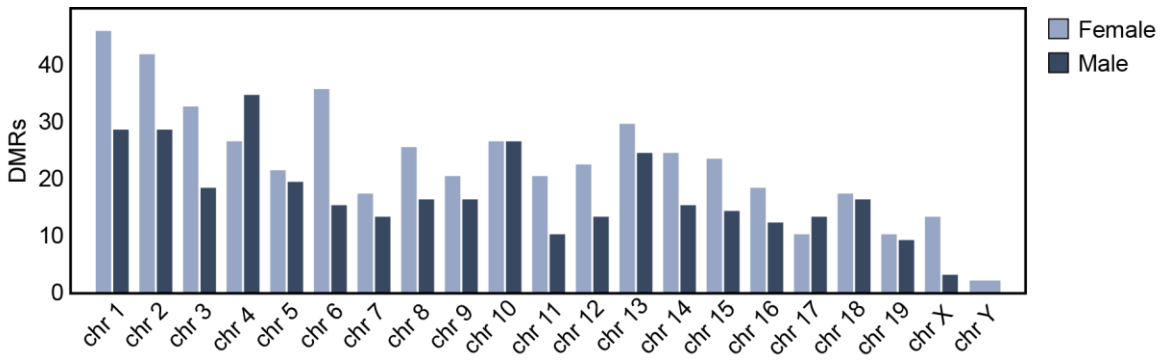

Fig 4.11. Sex-specific methylation alterations in frontal cortex neurons of offspring with and without maternal psilocybin exposure. (A) Distribution of

differentially methylated regions by chromosomal position between male and female subjects.

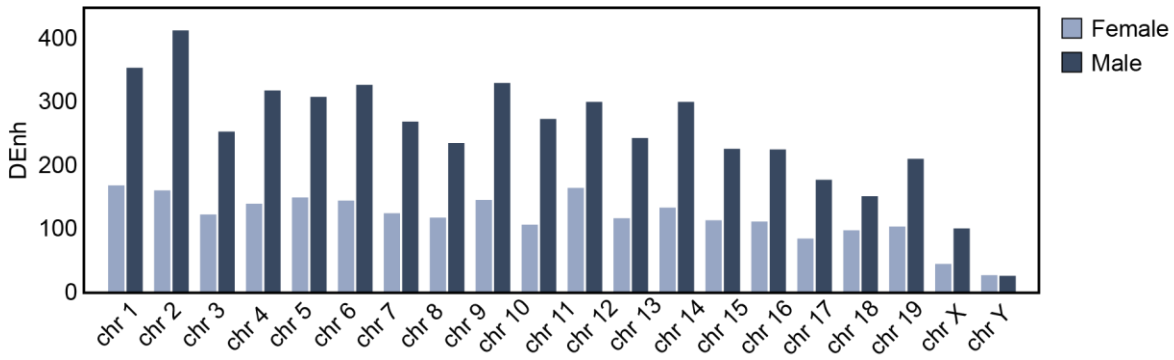

Fig 4.12. Sex-specific enhancer alterations in frontal cortex neurons of offspring with and without maternal psilocybin exposure. Distribution of differential enhancer linked gene counts by chromosomal position between male and female subjects.

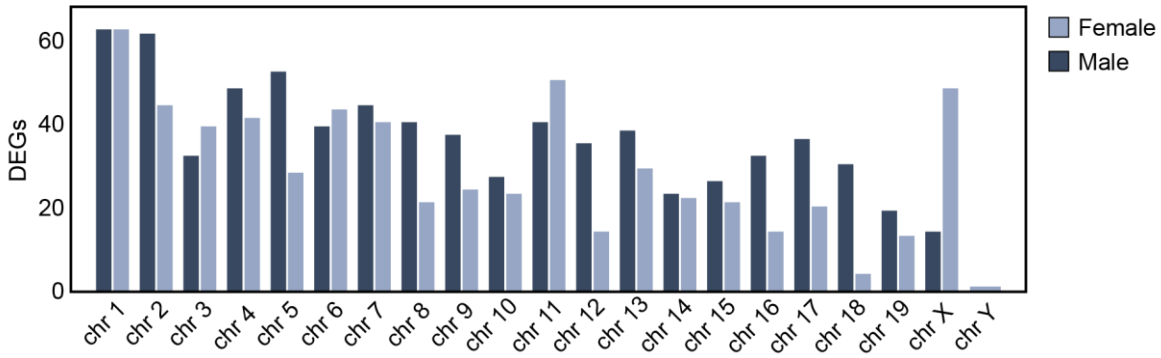

Fig 4.13. Sex-specific gene expression in frontal cortex neurons of offspring with and without maternal psilocybin exposure. Distribution of differentially expressed genes by chromosomal position between male and female subjects.

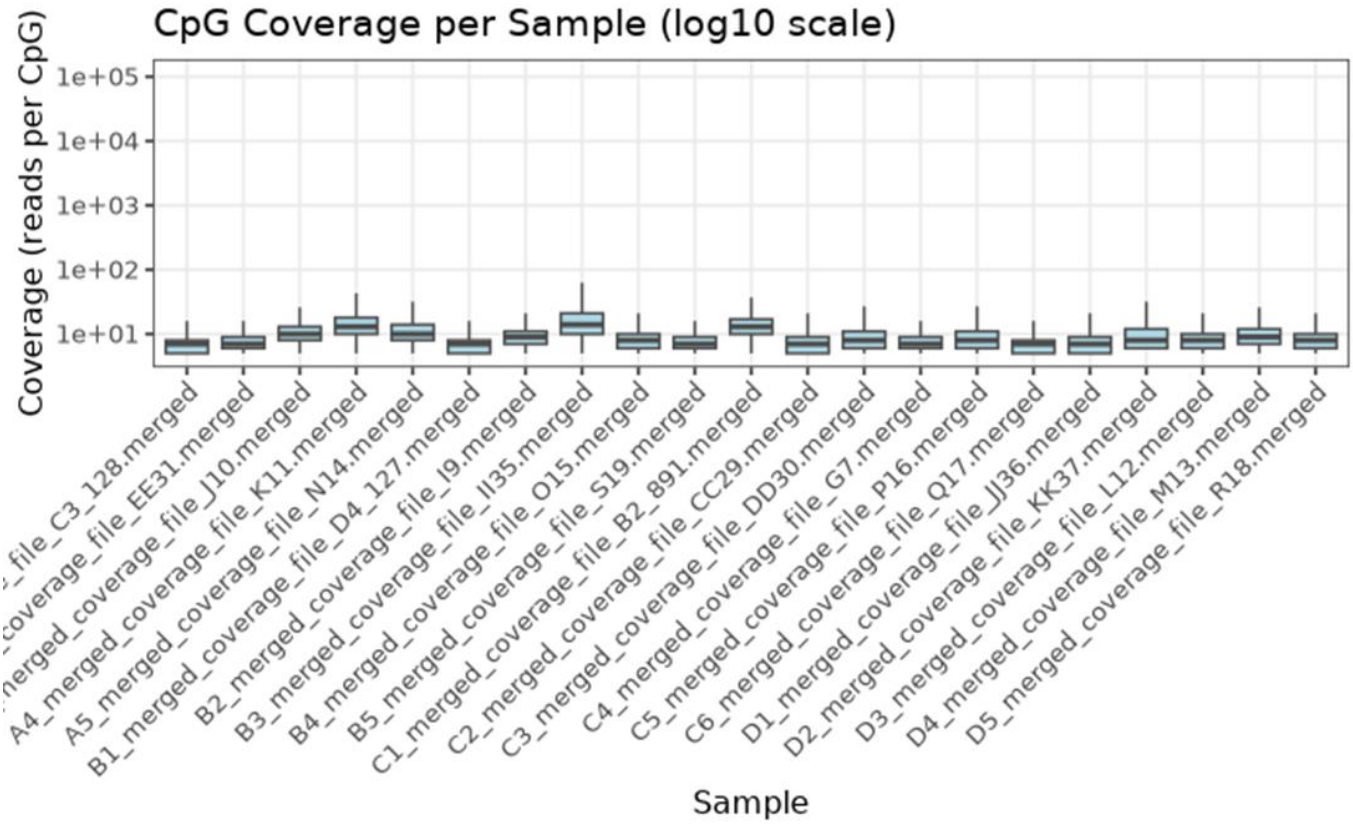

Fig 4.14. Whole genome bisulfite sequencing quality control. Average read coverage at each CpG site across two replicates for each sample.

4.7 Contributions

T.M.H, under the direction of Dr. Chang Lu, performed ChIP-seq, WGBS, and RNA-seq assays from mouse brains provided by Dr. Javier Gonzalez-Maeso and collected by Maya Gaines-Smith and Michael Fiorillo. T.M.H performed all bioinformatics discussed in this chapter.

5. Summary and outlook

As the cost and throughput of sequencing technology continues to improve, the feasibility and prevalence of *omic* investigations has expanded accordingly. However, with these technological advances, new bottlenecks arise across the experimental pipeline, from sample preparation to assay input requirements. Microfluidic systems have emerged as a powerful solution to several of these critical challenges, providing compact and reliable sample handling while miniaturizing assays to substantially reduce input requirements. In this thesis, I describe several applications of microfluidic platforms to address key procedural bottlenecks and perform sensitive biological investigations that would otherwise be limited by prohibitively high input demands.

Sequencing assays generally can be dissected into three distinct protocols: The extraction of nucleic acids from the acquired sample, the conversion of these nucleic acids to readable libraries, and sequencing of these libraries. For a sequencing platform to be truly considered “point-of-care”, all three of these components need to be equally portable. The invention of nanopore sequencers has addressed the last of these three components by providing sequencing devices that fit in the palm of one’s hand and powered by a laptop. Yet to this point, lack of innovation addressing the other two sample preparation steps have minimized the on-site impact of this truly remarkable technology. In chapter 2, I present a platform that addresses one of these two existing bottlenecks by successfully generating sequence ready libraries in an all-in-one amplification, purification, and library preparation system. This brings true point-of-care sequencing closer to reality than ever before. However, to truly create a field deployable sequencing system, future work will need to be performed to address the final existing

bottleneck of nucleic acid extraction. Some initial work has been presented probing this question and have demonstrated its feasibility in the future. Additionally, any such technology would need to be integrated into a system with our library preparation microreactor and nanopore sequencing device. I am certain that this bottleneck will be addressed soon and this fully integrated system will become a reality.

The remainder of this thesis focused on application of microfluidic platforms for epigenomic investigations of small input sample sets. Specifically, these investigations centered around the epigenomic impact of psychedelic exposure on cortical and sub-cortical neuron populations in mice. In chapter 3, I uncovered large recovery of epigenomic alterations induced by prolonged opioid used in the nucleus accumbens and prefrontal cortex of mice. These recoveries were more extensive in the nucleus accumbens, which is a key downstream modulator of reward seeking behavior underlying addiction. Amongst the two regions studied here, I found outsized epigenomic recovery in male cohorts compared to females. This finding is in agreement with accompanying behavioral studies which demonstrate pronounced reduction in drug seeking behavior and withdrawal symptoms in oxycodone addicted males following psilocybin treatment compared to female counterparts. Given the disproportionate burden of opioid use disorder among males, this finding points to psylocibin’s potential to inform therapeutic strategies aimed at one of the most vulnerable demographics. I look forward to future clinical investigation of psilocybin and other psychedelic compounds to modulate drug seeking behavior within opioid addicted human cohorts.

As evidence continues to support the efficacy of psychedelic compounds for treating various psychological disorders, their clinical adoption will likely expand. However,

before widespread implementation, it is imperative that physicians fully understand the downstream effects of these treatments. As psychedelics emerge as a potential therapeutic for conditions such as depression, of which highest incidence rates are found amongst young adult (18-25 years old) females, I sought to understand how prolonged psilocybin exposure of prenatal mothers may affect their offspring epigenome. While similar studies involving rapid acting antidepressants such as ketamine have demonstrated inherited epigenomic alterations, no such study had been performed examining traditional psychedelics. Our results of our large multi-omic investigation uncovered significant inherited alterations to transcription factor regulatory networks in the offspring of psychedelic treated mothers. Specifically, key transcription factors with expression increased within cortical neurons following primary psychedelic exposure, *Egr2* and *JunB*, were shown to have diminished regulatory reach within the maternal offspring cohort. This may limit downstream efficacy of primary exposure to *Egr2* and *JunB* excitatory compounds within these offspring. Additionally, the minimization of *Egr2* regulatory function was localized to the female offspring cohort, suggesting significant sex-specific epigenomic alterations inherited from maternal exposure.

Because epigenomic reprogramming in mammals thought to be robust, the marked disruptions to histone modifications and DNA methylation observed in the offspring are unexpectedly significant. In the future, I aim to expand the scope of this study to include the frontal cortex samples of the mothers of the examined offspring. This will involve a cohort of 6 prenatal dams following an extended daily psilocybin regimen, and 6 saline injected prenatal dams as a control. Frontal cortex tissue from these two groups will

undergo the same ChIP-seq, WGBS, and RNA-seq assays as the offspring cohorts. From this I can compare which specific alterations found within the offspring cohorts are replicated directly from the mothers and which are distinct. Additionally, I seek to understand the impact of the diminished regulatory network within the offspring cohorts by challenging them with direct psilocybin exposure. This would involve generating additional mice in each of the 4 cohorts outlined in chapter 5. These mice would undergo behavioral testing and gene expression profiling following primary psilocybin exposure. As psilocybin binding of 5-HT2A receptor in the cortical neurons mediates increased *Egr2* expression, I seek to discern the transcriptional impact of increased *Egr2* activity of highly weighted genes identified within the regulatory network. As KEGG analysis of down-weighted genes linked to *Egr2* are heavily localized around neurogenic stimulatory signaling pathways, I hypothesize that offspring will experience diminished cognitive returns compared to those elicited in offspring of non-treated mothers. I look forward to performing these studies in the near future to generate a full profile of the inherited epigenetic landscape of maternal psychedelic exposure to better inform future clinical decisions.

6. References

1. Maggioli, M.F., et al., *Adaptive Immune Responses following Senecavirus A Infection in Pigs.* J Virol, 2018. **92**(3).

2. Vannucci, F.A., et al., *Identification and Complete Genome of Seneca Valley Virus in Vesicular Fluid and Sera of Pigs Affected with Idiopathic Vesicular Disease, Brazil.* Transbound Emerg Dis, 2015. **62**(6): p. 589-93.

3. Knight-Jones, T.J. and J. Rushton, *The economic impacts of foot and mouth disease - what are they, how big are they and where do they occur?* Prev Vet Med, 2013. **112**(3-4): p. 161-73.

4. Grubman, M.J. and B. Baxt, *Foot-and-mouth disease.* Clin Microbiol Rev, 2004. **17**(2): p. 465-93.

5. Joshi, L.R., et al., *Pathogenesis of Senecavirus A infection in finishing pigs.* J Gen Virol, 2016. **97**(12): p. 3267-3279.

6. Leme, R.A., et al., *Clinical Manifestations of Senecavirus A Infection in Neonatal Pigs, Brazil, 2015.* Emerg Infect Dis, 2016. **22**(7): p. 1238-41.

7. Saeng-Chuto, K., et al., *The first detection of Senecavirus A in pigs in Thailand, 2016.* Transbound Emerg Dis, 2018. **65**(1): p. 285-288.

8. Vieira, M.V., et al., *The third wave of Seneca Valley virus outbreaks in pig herds in southern Brazil.* Braz J Microbiol, 2022. **53**(3): p. 1701-1706.

9. Hause, B.M., et al., *Senecavirus A in Pigs, United States, 2015.* Emerg Infect Dis, 2016. **22**(7): p. 1323-5.

10. Montiel, N., et al., *Vesicular Disease in 9-Week-Old Pigs Experimentally Infected with Senecavirus A.* Emerg Infect Dis, 2016. **22**(7): p. 1246-8.

11. Zhang, X., et al., *Review of Seneca Valley Virus: A Call for Increased Surveillance and Research.* Front Microbiol, 2018. **9**: p. 940.

12. Rahbari, R., N. Moradi, and M. Abdi, *rRT-PCR for SARS-CoV-2: Analytical considerations.* Clin Chim Acta, 2021. **516**: p. 1-7.

13. Bracht, A.J., et al., *Real-Time Reverse Transcription PCR Assay for Detection of Senecavirus A in Swine Vesicular Diagnostic Specimens.* PLoS One, 2016. **11**(1): p. e0146211.

14. Feronato, C., et al., *Development and evaluation of a nested-PCR assay for Senecavirus A diagnosis.* Trop Anim Health Prod, 2018. **50**(2): p. 337-344.

15. Fowler, V.L., et al., *Development of a novel real-time RT-PCR assay to detect Seneca Valley virus-1 associated with emerging cases of vesicular disease in pigs.* J Virol Methods, 2017. **239**: p. 34-37.

16. Zhang, J., et al., *Development and evaluation of a real-time RT-PCR and a field-deployable RT-insulated isothermal PCR for the detection of Seneca Valley virus.* BMC Vet Res, 2019. **15**(1): p. 168.

17. Dao Thi, V.L., et al., *A colorimetric RT-LAMP assay and LAMP-sequencing for detecting SARS-CoV-2 RNA in clinical samples.* Sci Transl Med, 2020. **12**(556).

18. Notomi, T., et al., *Loop-mediated isothermal amplification of DNA.* Nucleic Acids Res, 2000. **28**(12): p. E63.

19. Kashir, J. and A. Yaqinuddin, *Loop mediated isothermal amplification (LAMP) assays as a rapid diagnostic for COVID-19.* Med Hypotheses, 2020. **141**: p. 109786.

20. Mori, Y. and T. Notomi, *Loop-mediated isothermal amplification (LAMP): a rapid, accurate, and cost-effective diagnostic method for infectious diseases.* J Infect Chemother, 2009. **15**(2): p. 62-9.

21. Garg, N., F.J. Ahmad, and S. Kar, *Recent advances in loop-mediated isothermal amplification (LAMP) for rapid and efficient detection of pathogens.* Curr Res Microb Sci, 2022. **3**: p. 100120.

22. Rolando, J.C., et al., *Real-time kinetics and high-resolution melt curves in single-molecule digital LAMP to differentiate and study specific and non-specific amplification.* Nucleic Acids Res, 2020. **48**(7): p. e42.

23. Lin, B., J. Hui, and H. Mao, *Nanopore Technology and Its Applications in Gene Sequencing.* Biosensors (Basel), 2021. **11**(7).

24. Zheng, P., et al., *Nanopore sequencing technology and its applications.* MedComm (2020), 2023. **4**(4): p. e316.

25. Brinkmann, A., et al., *AmpliCoV: Rapid Whole-Genome Sequencing Using Multiplex PCR Amplification and Real-Time Oxford Nanopore MinION Sequencing Enables Rapid Variant Identification of SARS-CoV-2.* Front Microbiol, 2021. **12**: p. 651151.

26. Ciuffreda, L., H. Rodriguez-Perez, and C. Flores, *Nanopore sequencing and its application to the study of microbial communities.* Comput Struct Biotechnol J, 2021. **19**: p. 1497-1511.

27. Hess, J.F., et al., *Library preparation for next generation sequencing: A review of automation strategies.* Biotechnol Adv, 2020. **41**: p. 107537.

28. Berger, B. and Y.W. Yu, *Navigating bottlenecks and trade-offs in genomic data analysis.* Nat Rev Genet, 2023. **24**(4): p. 235-250.

29. Snider, A., et al., *A Microfluidics Workflow for Sample Preparation for Next-Generation DNA Sequencing.* SLAS Technol, 2019. **24**(2): p. 196-208.

30. Hoffmann, A., et al., *Automation of customizable library preparation for next-generation sequencing into an open microfluidic platform.* Sci Rep, 2024. **14**(1): p. 17150.

31. Murphy, T.W., et al., *Microfluidic Platform for Next-Generation Sequencing Library Preparation with Low-Input Samples.* Anal Chem, 2020. **92**(3): p. 2519-2526.

32. Kim, S., et al., *High-throughput automated microfluidic sample preparation for accurate microbial genomics.* Nat Commun, 2017. **8**: p. 13919.

33. Torezin Mendonca, G., et al., *A new RT-LAMP-on-a-Chip Instrument for SARS-CoV-2 diagnostics.* Microchem J, 2022. **180**: p. 107600.

34. Papadakis, G., et al., *Portable real-time colorimetric LAMP-device for rapid quantitative detection of nucleic acids in crude samples.* Sci Rep, 2022. **12**(1): p. 3775.

35. Song, M., S. Hong, and L.P. Lee, *Multiplexed Ultrasensitive Sample-to-Answer RT-LAMP Chip for the Identification of SARS-CoV-2 and Influenza Viruses.* Adv Mater, 2023. **35**(10): p. e2207138.

36. Donia, A., et al., *Integration of RT-LAMP and Microfluidic Technology for Detection of SARS-CoV-2 in Wastewater as an Advanced Point-of-Care Platform.* Food Environ Virol, 2022. **14**(4): p. 364-373.

37. Wan, L., et al., *LampPort: a handheld digital microfluidic device for loop-mediated isothermal amplification (LAMP).* Biomed Microdevices, 2019. **21**(1): p. 9.

38. Liu, F., et al., *Rescue of Senecavirus A to uncover mutation profiles of its progenies during 80 serial passages in vitro.* Vet Microbiol, 2021. **253**: p. 108969.

39. Li, C., et al., *Senecavirus-Specific Recombination Assays Reveal the Intimate Link between Polymerase Fidelity and RNA Recombination.* J Virol, 2019. **93**(13).

40. Sanjuan, R. and P. Domingo-Calap, *Mechanisms of viral mutation.* Cell Mol Life Sci, 2016. **73**(23): p. 4433-4448.

41. Xu, W., et al., *Genome wide analysis of the evolution of Senecavirus A from swine clinical material and assembly yard environmental samples.* PLoS One, 2017. **12**(5): p. e0176964.

42. Townsend, E.A., S.S. Negus, and M.L. Banks, *Medications Development for Treatment of Opioid Use Disorder.* Cold Spring Harb Perspect Med, 2021. **11**(1).

43. Bogenschutz, M.P., et al., *Psilocybin-assisted treatment for alcohol dependence: a proof-of-concept study.* J Psychopharmacol, 2015. **29**(3): p. 289-99.

44. Johnson, M.W., et al., *Pilot study of the 5-HT2AR agonist psilocybin in the treatment of tobacco addiction.* J Psychopharmacol, 2014. **28**(11): p. 983-92.

45. Jones, G., et al., *Associations between classic psychedelics and opioid use disorder in a nationally-representative U.S. adult sample.* Sci Rep, 2022. **12**(1): p. 4099.

46. Li, Y., *Modern epigenetics methods in biological research.* Methods, 2021. **187**: p. 104-113.

47. Bannister, A.J. and T. Kouzarides, *Regulation of chromatin by histone modifications.* Cell Res, 2011. **21**(3): p. 381-95.

48. Marzi, S.J., et al., *A histone acetylome-wide association study of Alzheimer's disease identifies disease-associated H3K27ac differences in the entorhinal cortex.* Nat Neurosci, 2018. **21**(11): p. 1618-1627.

49. Arif, M., et al., *Epigenetic modification: a regulatory mechanism in essential hypertension.* Hypertens Res, 2019. **42**(8): p. 1099-1113.

50. Nielsen, D.A., et al., *Epigenetics of drug abuse: predisposition or response.* Pharmacogenomics, 2012. **13**(10): p. 1149-60.

51. Sheng, J., et al., *Histone H3 phosphoacetylation is critical for heroin-induced place preference.* Neuroreport, 2011. **22**(12): p. 575-80.

52. Wang, Z., et al., *Epigenetic upregulation of PSD-95 contributes to the rewarding behavior by morphine conditioning.* Eur J Pharmacol, 2014. **732**: p. 123-9.

53. Egervari, G., et al., *Striatal H3K27 Acetylation Linked to Glutamatergic Gene Dysregulation in Human Heroin Abusers Holds Promise as Therapeutic Target.* Biol Psychiatry, 2017. **81**(7): p. 585-594.

54. Chen, W.S., et al., *Effects of histone deacetylase inhibitor sodium butyrate on heroin seeking behavior in the nucleus accumbens in rats.* Brain Res, 2016. **1652**: p. 151-157.

55. Browne, C.J., et al., *Epigenetic Mechanisms of Opioid Addiction.* Biol Psychiatry, 2020. **87**(1): p. 22-33.

56. Brind'Amour, J., et al., *An ultra-low-input native ChIP-seq protocol for genome-wide profiling of rare cell populations.* Nat Commun, 2015. **6**: p. 6033.

57. Schmidl, C., et al., *ChIPmentation: fast, robust, low-input ChIP-seq for histones and transcription factors.* Nat Methods, 2015. **12**(10): p. 963-965.

58. Zhu, D., et al., *A simple, robust, cost-effective, and low-input ChIP-seq method for profiling histone modifications and Pol II in plants.* New Phytol, 2024. **244**(4): p. 1658-1669.

59. Ma, S., et al., *Low-input and multiplexed microfluidic assay reveals epigenomic variation across cerebellum and prefrontal cortex.* Sci Adv, 2018. **4**(4): p. eaar8187.

60. Cao, Z., et al., *A microfluidic device for epigenomic profiling using 100 cells.* Nat Methods, 2015. **12**(10): p. 959-62.

61. Zhu, B., et al., *MOWChIP-seq for low-input and multiplexed profiling of genome-wide histone modifications.* Nat Protoc, 2019. **14**(12): p. 3366-3394.

62. Sun, H., et al., *Morphine epigenomically regulates behavior through alterations in histone H3 lysine 9 dimethylation in the nucleus accumbens.* J Neurosci, 2012. **32**(48): p. 17454-64.

63. Zhang, Z., et al., *MeCP2 repression of G9a in regulation of pain and morphine reward.* J Neurosci, 2014. **34**(27): p. 9076-87.

64. de la Fuente Revenga, M., et al., *Prolonged epigenomic and synaptic plasticity alterations following single exposure to a psychedelic in mice.* Cell Rep, 2021. **37**(3): p. 109836.

65. Kundakovic, M. and M. Tickerhoof, *Epigenetic mechanisms underlying sex differences in the brain and behavior.* Trends Neurosci, 2024. **47**(1): p. 18-35.

66. Becker, J.B. and M. Hu, *Sex differences in drug abuse.* Front Neuroendocrinol, 2008. **29**(1): p. 36-47.

67. Naler, L.B., et al., *Epigenomic and transcriptomic analyses reveal differences between low-grade inflammation and severe exhaustion in LPS-challenged murine monocytes.* Commun Biol, 2022. **5**(1): p. 102.

68. Bonev, B., et al., *Multiscale 3D Genome Rewiring during Mouse Neural Development.* Cell, 2017. **171**(3): p. 557-572 e24.

69. Zhong, X., et al., *Human cell adhesion molecules: annotated functional subtypes and overrepresentation of addiction-associated genes.* Ann N Y Acad Sci, 2015. **1349**(1): p. 83-95.

70. Ly, C., et al., *Psychedelics Promote Structural and Functional Neural Plasticity.* Cell Rep, 2018. **23**(11): p. 3170-3182.

71. Shao, L.X., et al., *Psilocybin induces rapid and persistent growth of dendritic spines in frontal cortex in vivo.* Neuron, 2021. **109**(16): p. 2535-2544 e4.

72. Gould, T.J., *Addiction and cognition.* Addict Sci Clin Pract, 2010. **5**(2): p. 4-14.

73. Nagai, T., et al., *Phosphoproteomics of the Dopamine Pathway Enables Discovery of Rap1 Activation as a Reward Signal In Vivo.* Neuron, 2016. **89**(3): p. 550-65.

74. Nestler, E.J., *Molecular neurobiology of addiction.* Am J Addict, 2001. **10**(3): p. 201-17.

75. Li, C.Y., X. Mao, and L. Wei, *Genes and (common) pathways underlying drug addiction.* PLoS Comput Biol, 2008. **4**(1): p. e2.

76. Liu, Z., et al., *Epigenomic tomography for probing spatially defined chromatin state in the brain.* Cell Rep Methods, 2024. **4**(3): p. 100738.

77. Volkow, N.D., M. Michaelides, and R. Baler, *The Neuroscience of Drug Reward and Addiction.* Physiol Rev, 2019. **99**(4): p. 2115-2140.

78. Koob, G.F. and N.D. Volkow, *Neurocircuitry of addiction.* Neuropsychopharmacology, 2010. **35**(1): p. 217-38.

79. Zanni, G., et al., *Female and male rats readily consume and prefer oxycodone to water in a chronic, continuous access, two-bottle oral voluntary paradigm.* Neuropharmacology, 2020. **167**: p. 107978.

80. Mavrikaki, M., et al., *Oxycodone self-administration in male and female rats.* Psychopharmacology (Berl), 2017. **234**(6): p. 977-987.

81. Radke, A.K., et al., *Contributions of nucleus accumbens dopamine to cognitive flexibility.* Eur J Neurosci, 2019. **50**(3): p. 2023-2035.

82. Klanker, M., M. Feenstra, and D. Denys, *Dopaminergic control of cognitive flexibility in humans and animals.* Front Neurosci, 2013. **7**: p. 201.

83. Conn, K., et al., *Psilocybin restrains activity-based anorexia in female rats by enhancing cognitive flexibility: contributions from 5-HT1A and 5-HT2A receptor mechanisms.* Mol Psychiatry, 2024.

84. Euston, D.R., A.J. Gruber, and B.L. McNaughton, *The role of medial prefrontal cortex in memory and decision making.* Neuron, 2012. **76**(6): p. 1057-70.

85. Meinhardt, M.W., et al., *Psilocybin targets a common molecular mechanism for cognitive impairment and increased craving in alcoholism.* Sci Adv, 2021. **7**(47): p. eabh2399.

86. Goldstein, R.Z. and N.D. Volkow, *Dysfunction of the prefrontal cortex in addiction: neuroimaging findings and clinical implications.* Nat Rev Neurosci, 2011. **12**(11): p. 652-69.

87. Peters, J., R.T. LaLumiere, and P.W. Kalivas, *Infralimbic prefrontal cortex is responsible for inhibiting cocaine seeking in extinguished rats.* J Neurosci, 2008. **28**(23): p. 6046-53.

88. Moorman, D.E., et al., *Differential roles of medial prefrontal subregions in the regulation of drug seeking.* Brain Res, 2015. **1628**(Pt A): p. 130-46.

89. Lake, B.B., et al., *Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain.* Science, 2016. **352**(6293): p. 1586-90.

90. Li, H., et al., *The Sequence Alignment/Map format and SAMtools.* Bioinformatics, 2009. **25**(16): p. 2078-9.

91. Quinlan, A.R. and I.M. Hall, *BEDTools: a flexible suite of utilities for comparing genomic features.* Bioinformatics, 2010. **26**(6): p. 841-2.

92. Saremi, B., et al., *A comparison of strategies for generating artificial replicates in RNA-seq experiments.* Sci Rep, 2022. **12**(1): p. 7170.

93. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biol, 2014. **15**(12): p. 550.

94. Hu, B., et al., *Neuronal and glial 3D chromatin architecture informs the cellular etiology of brain disorders.* Nat Commun, 2021. **12**(1): p. 3968.

95. Rousseeuw, P.J., *Silhouettes - a Graphical Aid to the Interpretation and Validation of Cluster-Analysis.* Journal of Computational and Applied Mathematics, 1987. **20**: p. 53-65.

96. Yu, G., et al., *clusterProfiler: an R package for comparing biological themes among gene clusters.* OMICS, 2012. **16**(5): p. 284-7.

97. Ge, S.X., D. Jung, and R. Yao, *ShinyGO: a graphical gene-set enrichment tool for animals and plants.* Bioinformatics, 2020. **36**(8): p. 2628-2629.

98. Heard, E. and R.A. Martienssen, *Transgenerational epigenetic inheritance: myths and mechanisms.* Cell, 2014. **157**(1): p. 95-109.

99. Tombari, R.J., et al., *Developmental Neurotoxicity Screen of Psychedelics and Other Drugs of Abuse in Larval Zebrafish (Danio rerio).* ACS Chem Neurosci, 2023. **14**(5): p. 875-884.

100. Goldberg, L.R. and T.J. Gould, *Multigenerational and transgenerational effects of paternal exposure to drugs of abuse on behavioral and neural function.* Eur J Neurosci, 2019. **50**(3): p. 2453-2466.

101. Finegersh, A. and G.E. Homanics, *Paternal alcohol exposure reduces alcohol drinking and increases behavioral sensitivity to alcohol selectively in male offspring.* PLoS One, 2014. **9**(6): p. e99078.

102. Vassoler, F.M., et al., *Epigenetic inheritance of a cocaine-resistance phenotype.* Nat Neurosci, 2013. **16**(1): p. 42-7.

103. Vargas, M.V., et al., *Psychedelics promote neuroplasticity through the activation of intracellular 5-HT2A receptors.* Science, 2023. **379**(6633): p. 700-706.

104. Hendricks, P.S., M.W. Johnson, and R.R. Griffiths, *Psilocybin, psychological distress, and suicidality.* J Psychopharmacol, 2015. **29**(9): p. 1041-3.

105. Grob, C.S., et al., *Pilot study of psilocybin treatment for anxiety in patients with advanced-stage cancer.* Arch Gen Psychiatry, 2011. **68**(1): p. 71-8.

106. Wilcox, J.A., *Psilocybin and Obsessive Compulsive Disorder.* J Psychoactive Drugs, 2014. **46**(5): p. 393-5.

107. Nardou, R., et al., *Psychedelics reopen the social reward learning critical period.* Nature, 2023. **618**(7966): p. 790-798.

108. Martin, D.A., et al., *Chronic LSD alters gene expression profiles in the mPFC relevant to schizophrenia.* Neuropharmacology, 2014. **83**: p. 1-8.

109. Larrea, A., et al., *Mitochondrial Metabolism in Major Depressive Disorder: From Early Diagnosis to Emerging Treatment Options.* J Clin Med, 2024. **13**(6).

110. Inserra, A., et al., *Modulation of DNA methylation and protein expression in the prefrontal cortex by repeated administration of D-lysergic acid diethylamide (LSD): Impact on neurotropic, neurotrophic, and neuroplasticity signaling.* Prog Neuropsychopharmacol Biol Psychiatry, 2022. **119**: p. 110594.

111. Inserra, A., et al., *Epigenetic mechanisms of rapid-acting antidepressants.* Transl Psychiatry, 2024. **14**(1): p. 359.

112. Ruffell, S.G.D., et al., *Ceremonial Ayahuasca in Amazonian Retreats-Mental Health and Epigenetic Outcomes From a Six-Month Naturalistic Study.* Front Psychiatry, 2021. **12**: p. 687615.

113. Brown, I.R. and C.C. Liew, *Lysergic acid diethylamide: effect on histone acetylation in rabbit brain.* Science, 1975. **188**(4193): p. 1122-3.

114. Gonzalez-Maeso, J., et al., *Hallucinogens recruit specific cortical 5-HT(2A) receptor-mediated signaling pathways to affect behavior.* Neuron, 2007. **53**(3): p. 439-52.

115. Wojtas, A., et al., *Effect of Psilocybin and Ketamine on Brain Neurotransmitters, Glutamate Receptors, DNA and Rat Behavior.* Int J Mol Sci, 2022. **23**(12).

116. Germann, C.B., *The Psilocybin-Telomere Hypothesis: An empirically falsifiable prediction concerning the beneficial neuropsychopharmacological effects of psilocybin on genetic aging.* Med Hypotheses, 2020. **134**: p. 109406.

117. Martin, D.A. and C.D. Nichols, *Psychedelics Recruit Multiple Cellular Types and Produce Complex Transcriptional Responses Within the Brain.* EBioMedicine, 2016. **11**: p. 262-277.

118. Davoudian, P.A., L.X. Shao, and A.C. Kwan, *Shared and Distinct Brain Regions Targeted for Immediate Early Gene Expression by Ketamine and Psilocybin.* ACS Chem Neurosci, 2023. **14**(3): p. 468-480.

119. Moores, G., et al., *Sex differences in neurology: a scoping review.* BMJ Open, 2023. **13**(4): p. e071200.

120. Biosca-Brull, J., et al., *A transcriptomic analysis in mice following a single dose of ibogaine identifies new potential therapeutic targets.* Transl Psychiatry, 2024. **14**(1): p. 41.

121. Albert, P.R., *Why is depression more prevalent in women?* J Psychiatry Neurosci, 2015. **40**(4): p. 219-21.

122. Walters, K., et al., *Recent trends in the incidence of anxiety diagnoses and symptoms in primary care.* PLoS One, 2012. **7**(8): p. e41670.

123. Picelli, S., et al., *Full-length RNA-seq from single cells using Smart-seq2.* Nat Protoc, 2014. **9**(1): p. 171-81.

124. Amini, S., et al., *Haplotype-resolved whole-genome sequencing by contiguity-preserving transposition and combinatorial indexing.* Nat Genet, 2014. **46**(12): p. 1343-9.

125. Udayasuryan, B., et al., *Fusobacterium nucleatum infection modulates the transcriptome and epigenome of HCT116 colorectal cancer cells in an oxygen-dependent manner.* Commun Biol, 2024. **7**(1): p. 551.

126. Yu, Y., et al., *Effects of ketamine-induced H3K9 hypoacetylation during pregnancy on cardiogenesis of mouse offspring.* Birth Defects Res, 2023. **115**(7): p. 770-781.

127. Cheung, H.M. and D.T.W. Yew, *Effects of Perinatal Exposure to Ketamine on the Developing Brain.* Front Neurosci, 2019. **13**: p. 138.

128. Chorbov, V.M., et al., *Elevated levels of DNA methylation at the OPRM1 promoter in blood and sperm from male opioid addicts.* J Opioid Manag, 2011. **7**(4): p. 258-64.

129. Ouko, L.A., et al., *Effect of alcohol consumption on CpG methylation in the differentially methylated regions of H19 and IG-DMR in male gametes: implications for fetal alcohol spectrum disorders.* Alcohol Clin Exp Res, 2009. **33**(9): p. 1615-27.

130. Byrnes, J.J., et al., *Adolescent opioid exposure in female rats: transgenerational effects on morphine analgesia and anxiety-like behavior in adult offspring.* Behav Brain Res, 2011. **218**(1): p. 200-5.

131. Birney, E., G.D. Smith, and J.M. Greally, *Epigenome-wide Association Studies and the Interpretation of Disease -Omics.* PLoS Genet, 2016. **12**(6): p. e1006105.

132. Lappalainen, T. and J.M. Greally, *Associating cellular epigenetic models with human phenotypes.* Nat Rev Genet, 2017. **18**(7): p. 441-451.

133. Fitz-James, M.H. and G. Cavalli, *Molecular mechanisms of transgenerational epigenetic inheritance.* Nat Rev Genet, 2022. **23**(6): p. 325-341.

134. Kobayashi, H., *Canonical and Non-canonical Genomic Imprinting in Rodents.* Front Cell Dev Biol, 2021. **9**: p. 713878.

135. Shi, Y., et al., *Phosphodiesterase-4D Knockdown in the Prefrontal Cortex Alleviates Memory Deficits and Synaptic Failure in Mouse Model of Alzheimer's Disease.* Front Aging Neurosci, 2021. **13**: p. 722580.

136. Zhang, C., et al., *Memory enhancing effects of BPN14770, an allosteric inhibitor of phosphodiesterase-4D, in wild-type and humanized mice.* Neuropsychopharmacology, 2018. **43**(11): p. 2299-2309.

137. Huang, Z., et al., *Dynamic expression pattern of Pde4d and its relationship with CpG methylation in the promoter during mouse embryo development.* Biochem Biophys Res Commun, 2013. **441**(4): p. 982-7.

138. Andergassen, D., et al., *Mapping the mouse Allelome reveals tissue-specific regulation of allelic expression.* Elife, 2017. **6**.

139. Morison, I.M., C.J. Paton, and S.D. Cleverley, *The imprinted gene and parent-of-origin effect database.* Nucleic Acids Res, 2001. **29**(1): p. 275-6.

140. Kim, H. and E. Kim, *Genetic background determines synaptic phenotypes in Arid1b-mutant mice.* Front Psychiatry, 2023. **14**: p. 1341348.

141. Maimaiti, S., et al., *Neuron-specific Mafb knockout causes growth retardation accompanied by an impaired growth hormone/insulin-like growth factor I axis.* Exp Anim, 2019. **68**(4): p. 435-442.

142. Whipple, A.J., et al., *Imprinted Maternally Expressed microRNAs Antagonize Paternally Driven Gene Programs in Neurons.* Mol Cell, 2020. **78**(1): p. 85-95 e8.

143. Lourida, I., et al., *Parathyroid hormone, cognitive function and dementia: a systematic review.* PLoS One, 2015. **10**(5): p. e0127574.

144. Jefsen, O.H., et al., *Transcriptional regulation in the rat prefrontal cortex and hippocampus after a single administration of psilocybin.* J Psychopharmacol, 2021. **35**(4): p. 483-493.

145. Heng, Y.H., et al., *NFIX regulates neural progenitor cell differentiation during hippocampal morphogenesis.* Cereb Cortex, 2014. **24**(1): p. 261-79.

146. Driller, K., et al., *Nuclear factor I X deficiency causes brain malformation and severe skeletal defects.* Mol Cell Biol, 2007. **27**(10): p. 3855-3867.

147. Pieraccioli, M., et al., *ZNF281 inhibits neuronal differentiation and is a prognostic marker for neuroblastoma.* Proc Natl Acad Sci U S A, 2018. **115**(28): p. 7356-7361.

148. Jouvert, P., et al., *Activation of the cGMP pathway in dopaminergic structures reduces cocaine-induced EGR-1 expression and locomotor activity.* J Neurosci, 2004. **24**(47): p. 10716-25.

149. Misra, K. and S.C. Pandey, *The decreased cyclic-AMP dependent-protein kinase A function in the nucleus accumbens: a role in alcohol drinking but not in anxiety-like behaviors in rats.* Neuropsychopharmacology, 2006. **31**(7): p. 1406-19.

150. Bading, H., *Nuclear calcium signalling in the regulation of brain function.* Nat Rev Neurosci, 2013. **14**(9): p. 593-608.

151. Sozanska, N., et al., *The molecular properties of the bHLH TCF4 protein as an intrinsically disordered hub transcription factor.* Cell Commun Signal, 2025. **23**(1): p. 154.

152. Logue, M.W., et al., *The ankyrin-3 gene is associated with posttraumatic stress disorder and externalizing comorbidity.* Psychoneuroendocrinology, 2013. **38**(10): p. 2249-57.

153. Luoni, A., et al., *Ankyrin-3 as a molecular marker of early-life stress and vulnerability to psychiatric disorders.* Transl Psychiatry, 2016. **6**(11): p. e943.

154. Kantarci, H., et al., *Tfap2a promotes specification and maturation of neurons in the inner ear through modulation of Bmp, Fgf and notch signaling.* PLoS Genet, 2015. **11**(3): p. e1005037.

155. Zainolabidin, N., et al., *Distinct Activities of Tfap2A and Tfap2B in the Specification of GABAergic Interneurons in the Developing Cerebellum.* Front Mol Neurosci, 2017. **10**: p. 281.

156. Liu, J., et al., *Acute psilocybin increased cortical activities in rats.* Front Neurosci, 2023. **17**: p. 1168911.

157. Lerer, E., et al., *Effects of psilocybin, psychedelic mushroom extract and 5-hydroxytryptophan on brain immediate early gene expression: Interaction with serotonergic receptor modulators.* Front Pharmacol, 2024. **15**: p. 1391412.

158. Duclot, F. and M. Kabbaj, *The Role of Early Growth Response 1 (EGR1) in Brain Plasticity and Neuropsychiatric Disorders.* Front Behav Neurosci, 2017. **11**: p. 35.

159. Minatohara, K., M. Akiyoshi, and H. Okuno, *Role of Immediate-Early Genes in Synaptic Plasticity and Neuronal Ensembles Underlying the Memory Trace.* Front Mol Neurosci, 2015. **8**: p. 78.

160. Berridge, M.J., *Dysregulation of neural calcium signaling in Alzheimer disease, bipolar disorder and schizophrenia.* Prion, 2013. **7**(1): p. 2-13.

161. Arnsten, A.F.T., D. Datta, and M. Wang, *The genie in the bottle-magnified calcium signaling in dorsolateral prefrontal cortex.* Mol Psychiatry, 2021. **26**(8): p. 3684-3700.

162. Bollen, E., et al., *Improved long-term memory via enhancing cGMP-PKG signaling requires cAMP-PKA signaling.* Neuropsychopharmacology, 2014. **39**(11): p. 2497-505.

163. Li, D.Y., et al., *Targeting the nitric oxide/cGMP signaling pathway to treat chronic pain.* Neural Regen Res, 2023. **18**(5): p. 996-1003.

164. Gallo, E.F. and C. Iadecola, *Neuronal nitric oxide contributes to neuroplasticity-associated protein expression through cGMP, protein kinase G, and extracellular signal-regulated kinase.* J Neurosci, 2011. **31**(19): p. 6947-55.

7. Publications

1. **Hadlock, T. M.,** Niece, J. D., Li, S., Mor, S., Lu, C. A microreactor system for point-of-care viral genome sequencing. *Biosensors and Bioelectronics* 297, (2026) 118369.

2. Jaster, A. M., **Hadlock, T. M.,** Buzzi. B., Maltman, J. L., Silva, G. M., Saha, S., Koseli, E., Pondelick, A. M., Thakur, N., Zhang, X., Li, G., Ledesma-Corvi, S., Moore, K. N., Peterson, H. R., Fujita, B., Zylki, A. L., Lewsi, M. R., Poklis, J. L., Halquist, M. S., Wolstenho Lu, C., Damaj, M. I., Gonzalez-Maeso, J. Sex-specific role for the 5-HT2A receptor in psilocybin-induced extinction of opioid rewarding properties. *Nature Communications* 16 (2025) 10206.

3. **Hadlock, T. M.,** Jaster, A. M., Fiorillo, M., Li, G., Gonzalez-Maeso, J., Lu, C. Multi-omic investigation of transgenerational effects of psychoactive therapeutics on mouse reward circuitry. *In preparation*.

4. **Hadlock, T. M.,** Fiorillo, M., Zhang, X., Li, G., Gonzalez-Maeso, J., Lu, C. Multi-omic investigation of sex-specific effects of sub-therapeutic LSD dosages on mouse frontal cortex. *In preparation*.

5. Zhu, B., Li, G., Saunders, J. M., Naler, L**., Hadlock, T. M.,** Wang, C., Garcia-Sastre, A., Gonzalez-Maeso, J., Lu, C. Prenatal and postnatal effects of maternal immune activation on synaptic and neurodevelopmental pathways via epigenetic mechanisms. *Translational Psychiatry* (Accepted, 2026).

6. Cao, C., Pal, T., Zhang, X., Liu, Z., **Hadlock, T. M.,** Neice, J. D., Sun, W., Vinatzer, B., Li, S., Chen, J. Microneedle-assisted CRISPR biosensing platform for early detection of plant diseases. *Small* (Under peer review, 2026)

8. Custom scripts used in this thesis

8.1 Chip-seq script

```
#!/bin/bash

Group=GROUPGOESHERE

subgroup=SUBgroupGOESHERE

Histone=HISTONEGOESHERE

cd /projects/lu_lab/Tom/LSD_Op/ChIP_seq/$Group/$Histone/$subgroup

cd Raw_Data/

gunzip *.gz

trim_galore *.fastq

cd ..

FILES=$PWD/Raw_Data/*.fq

FQ=.fq

SAM=.sam

LOG=.log

BAM=.bam

SORT=_sort.bam

UNI=_unique.bam

PREBED=_pre.bed

BED=.bed

EXT=_extend.bed

EXPWIN=_extend_promoter_win.bed
```

```
EXG4000=_extend_geno_4000.bed

EXPCOR=_extend_promoter_nor.bed

EXGCOR=_extend_geno_4000_nor.bed

G100=_geno_win_100.bed

SORTG100=_geno_win_100_sort.bed

BedG=.bedGraph

NEBw=.bw

mkdir Aligned_BAM

mkdir Aligned_SAM

mkdir SORT_BAM

mkdir UNIQUE_BAM

mkdir BED

mkdir EXTEND_BED

mkdir EXT_PROMOTER_WIN

mkdir EXT_GENO_WIN

mkdir Correlation

mkdir MACS

mkdir GenoWin100Sort

mkdir BedGraph

mkdir Nor_Ext_BW

input_length=$(wc -l < /ChIP_seq/Input/$Group/$Histone/$subgroup/input.bed )

for fn in $FILES

do
```

```
echo `basename "$fn"`

f=`basename "${fn%.*}"`

echo $f

bowtie2 -p 16 -x /ChIP_seq/Data/mm10/Sequence/Bowtie2Index/mm10 -U
$PWD/Raw_Data/$f$FQ -S $PWD/Aligned_SAM/$f$SAM
2>$PWD/Aligned_SAM/$f$LOG

samtools view -bq 10 $PWD/Aligned_SAM/$f$SAM > $PWD/Aligned_BAM/$f$BAM

samtools sort $PWD/Aligned_BAM/$f$BAM -o $PWD/SORT_BAM/$f$SORT

samtools index $PWD/SORT_BAM/$f$SORT

samtools rmdup -s $PWD/SORT_BAM/$f$SORT $PWD/UNIQUE_BAM/$f$UNI

bedtools bamtobed -i $PWD/UNIQUE_BAM/$f$UNI > $PWD/BED/$f$PREBED

bedtools subtract -a $PWD/BED/$f$PREBED -b
/ChIP_seq/Data/mm10/mm10.blacklist.bed > $PWD/BED/$f$BED

bedToBam -i $PWD/BED/$f$BED -g
/ChIP_seq/Data/bed_winbed/mm10/mm10genome.bed > $PWD/BED/$f$BAM

bedtools slop -i $PWD/BED/$f$BED -g
/ChIP_seq/Data/bed_winbed/mm10/mm10genome.bed -b 100 >
$PWD/EXTEND_BED/$f$EXT

bedtools coverage -counts -b $PWD/EXTEND_BED/$f$EXT -a
/ChIP_seq/Data/Promoter/mm10/mm10Promoter.bed >
$PWD/EXT_PROMOTER_WIN/$f$EXPWIN

bedtools coverage -counts -b $PWD/EXTEND_BED/$f$EXT -a
/ChIP_seq/Data/bed_winbed/mm10/mm10genome_win_4000.bed >
$PWD/EXT_GENO_WIN/$f$EXG4000
```

```
bedtools coverage -counts -b $PWD/EXTEND_BED/$f$EXT -a
/ChIP_seq/Data/bed_winbed/mm10/mm10genome_win_100.bed >
$PWD/EXT_GENO_WIN/$f$G100

sort -k1,1 -k2,2g -u -o $PWD/GenoWin100Sort/$f$SORTG100
$PWD/EXT_GENO_WIN/$f$G100

ChIP_length=$(wc -l < $PWD/BED/$f$BED)

paste $PWD/EXT_PROMOTER_WIN/$f$EXPWIN
/ChIP_seq/Input/$Group/$Histone/$subgroup/input_extend_promoter_win.bed | awk -v
OFS="\t" '{print $4/'$ChIP_length'*1000000-$8/'$input_length'*1000000}' >
$PWD/Correlation/$f$EXPCOR

paste $PWD/EXT_GENO_WIN/$f$EXG4000
/ChIP_seq/Input/$Group/$Histone/$subgroup/input_extend_genome_4000win.bed |
awk -v OFS="\t" '{print $4/'$ChIP_length'*1000000-$8/'$input_length'*1000000}' >
$PWD/Correlation/$f$EXGCOR

macs2 callpeak -t $PWD/BED/$f$BED -c
/ChIP_seq/Input/$Group/$Histone/$subgroup/input.bed -f BED -g mm -n $f -q 0.05 --
outdir $PWD/MACS

paste $PWD/EXT_GENO_WIN/$f$G100
/LSD_Op/ChIP_seq/Input/$Group/$Histone/$subgroup/input_extend_genome_100win.b
ed | awk -v OFS="\t" '{print $1,$2,$3,$4/'$ChIP_length'*1000000-
$8/'$input_length'*1000000}' > $PWD/BedGraph/$f$BedG

bedSort $PWD/BedGraph/$f$BedG $PWD/BedGraph/$f$BedG

bedGraphToBigWig $PWD/BedGraph/$f$BedG
/ChIP_seq/Data/bed_winbed/mm10/mm10genome.bed $PWD/Nor_Ext_BW/$f$NEBw

done
```

8.2 Diffbind counts

```
####setup
library(DiffBind)
library(dplyr)
library(GenomicRanges)
library(ChIPpeakAnno)
library(ChIPseeker)
# # See DiffBind documentation for how to put together sample sheet
setwd("C:/ChIP_seq")
infiles = "C:/ChIP_seq/infiles"
outfiles = "C:/ChIP_seq/outfiles"
gene_link = "C:/ChIP_seq/mm10Promoter.bed"
#groups=c('V/V_F','V/P_F','O/V_F','O/P_F', 'V/V_M','V/P_M','O/V_M','O/P_M')
fdr=0.05
run_fcs=c(1) #if you want to check multiple fold-changes, you can add them here


data <- dba(sampleSheet =
"C:/ChIP_seq/PSI_Op_DiffAn_Samplesheet_MaleFemaleNAC_Dup95.csv")
data
data_blacklist_remove = dba.blacklist(data, blacklist = DBA_BLACKLIST_MM10,
greylist = TRUE)
```

```
gl = read.table(gene_link, header = FALSE, stringsAsFactors = FALSE, sep="\t",
col.names=c("chr","start","stop"))

gl_gr = makeGRangesFromDataFrame(gl, keep.extra.columns = TRUE)

# Consensus should be the column that has the groupings that you are comparing

rep_consensus_set <- dba.peakset(data_blacklist_remove, consensus =
DBA_TREATMENT, minOverlap = 2)

rep_consensus_set

rep_consensus <- dba(rep_consensus_set,
mask=rep_consensus_set$masks$Consensus, minOverlap=1)

rep_consensus

ConsensusPeaks <- dba.peakset(rep_consensus, bRetrieve=TRUE)

ConsensusPeaks

#Generate counts matrix

data_analysis <- dba.count(data_blacklist_remove,summits = FALSE, peaks =
ConsensusPeaks,filter=1, bScaleControl = TMM_MINUS_FULL)

counts <- dba.peakset(data_analysis, bRetrieve=T, DataType=DBA_DATA_FRAME)

write.csv(counts, 'counts_NAC_MF.csv')
```

8.3 Dseq2 differential analysis

```
############preparation##############
library(dplyr)
library(DESeq2)
library(umap)
library(ggplot2)

rm(list = ls())
options(stringsAsFactors = F)
setwd("C:/ChIP_seq")
factors = read.csv('cofactorsNACMF.csv')
row.names(factors) = factors[,1]
factors = factors[,-1]
factors$Peaks = round(factors$Peaks,2)
factors$NSC = round(factors$NSC,2)
factors$sequencing_depth = round(factors$sequencing_depth,2)
factors$unique_reads = round(factors$unique_reads,2)
raw_def = read.csv('counts_NAC_MF.csv')
rownames(raw_def) = raw_def[,4]
def = raw_def[,5:ncol(raw_def)]
def = round(def)
peak_info = raw_def[,1:4]
```

```
####strong correlation: cor>0.7################

#################################B_1 vs.
A_1#################################################################

setwd("C:/ChIP_seq")

meta = bind_rows(factors[11:18,1:7],factors[1:10,1:7])

# might need to change below so the second numbers are equivalent

meta$condition_num = c(rep(1,8),rep(2,10))

group_B1A1 = data.frame(def[,11:18],def[,1:10])

miss_1 <- c()

for (i in 1:nrow(group_B1A1)) {

  if(length(which(group_B1A1[i,1:10]<10)) > 0.5*ncol(group_B1A1[,1:10])) miss_1 <-
append(miss_1,i)

}

miss_2 <- c()

for (i in 1:nrow(group_B1A1)) {

  if(length(which(group_B1A1[i,11:18]<10)) > 0.5*ncol(group_B1A1[,11:18])) miss_2 <-
append(miss_2,i)

}

miss = union(miss_1,miss_2)

group_B1A1 = group_B1A1[-miss,]

t_group_B1A1 = t(group_B1A1)

umap_results <- umap(t_group_B1A1,n_neighbors=8,min_dist=0.1,n_components = 6)

umaps = data.frame(umap_results[["layout"]])
```

```
umaps$condition = as.factor(meta$condition)
ggplot(umaps,mapping = aes(x=X1,y=X2,col=condition)) +
  geom_point(aes(colour = condition))


cor_con_batch = cor(meta$condition_num,meta$batch,method = 'pearson')
cor_con_peaks = cor(meta$condition_num,meta$Peaks,method = 'pearson')
cor_con_NSC = cor(meta$condition_num,meta$NSC,method = 'pearson')
cor_con_FRiP = cor(meta$condition_num,meta$FRiP,method = 'pearson')
cor_con_sequencing_depth =
cor(meta$condition_num,meta$sequencing_depth,method = 'pearson')
cor_con_uni_reads = cor(meta$condition_num,meta$unique_reads,method = 'pearson')


corr_uni_reads_1 = cor(umaps$X1,meta$unique_reads,method = 'pearson')
corr_uni_reads_2 = cor(umaps$X2,meta$unique_reads,method = 'pearson')
corr_uni_reads_3 = cor(umaps$X3,meta$unique_reads,method = 'pearson')
corr_uni_reads_4 = cor(umaps$X4,meta$unique_reads,method = 'pearson')
corr_uni_reads_5 = cor(umaps$X5,meta$unique_reads,method = 'pearson')
corr_uni_reads_6 = cor(umaps$X6,meta$unique_reads,method = 'pearson')
#
corr_peaks_1 = cor(umaps$X1,meta$Peaks,method = 'pearson')
corr_peaks_2 = cor(umaps$X2,meta$Peaks,method = 'pearson')
corr_peaks_3 = cor(umaps$X3,meta$Peaks,method = 'pearson')
corr_peaks_4 = cor(umaps$X4,meta$Peaks,method = 'pearson')
```

```
corr_peaks_5 = cor(umaps$X5,meta$Peaks,method = 'pearson')

corr_peaks_6 = cor(umaps$X6,meta$Peaks,method = 'pearson')

#

corr_NSC_1 = cor(umaps$X1,meta$NSC,method = 'pearson')

corr_NSC_2 = cor(umaps$X2,meta$NSC,method = 'pearson')

corr_NSC_3 = cor(umaps$X3,meta$NSC,method = 'pearson')

corr_NSC_4 = cor(umaps$X4,meta$NSC,method = 'pearson')

corr_NSC_5 = cor(umaps$X5,meta$NSC,method = 'pearson')

corr_NSC_6 = cor(umaps$X6,meta$NSC,method = 'pearson')

#

corr_FRiP_1 = cor(umaps$X1,meta$FRiP,method = 'pearson')

corr_FRiP_2 = cor(umaps$X2,meta$FRiP,method = 'pearson')

corr_FRiP_3 = cor(umaps$X3,meta$FRiP,method = 'pearson')

corr_FRiP_4 = cor(umaps$X4,meta$FRiP,method = 'pearson')

corr_FRiP_5 = cor(umaps$X5,meta$FRiP,method = 'pearson')

corr_FRiP_6 = cor(umaps$X6,meta$FRiP,method = 'pearson')

#

corr_sequencing_depth_1 = cor(umaps$X1,meta$sequencing_depth,method =
'pearson')

corr_sequencing_depth_2 = cor(umaps$X2,meta$sequencing_depth,method =
'pearson')

corr_sequencing_depth_3 = cor(umaps$X3,meta$sequencing_depth,method =
'pearson')
```

```
corr_sequencing_depth_4 = cor(umaps$X4,meta$sequencing_depth,method =
'pearson')

corr_sequencing_depth_5 = cor(umaps$X5,meta$sequencing_depth,method =
'pearson')

corr_sequencing_depth_6 = cor(umaps$X6,meta$sequencing_depth,method =
'pearson')

#

corr_batch_1 = cor(umaps$X1,meta$batch,method = 'pearson')

corr_batch_2 = cor(umaps$X2,meta$batch,method = 'pearson')

corr_batch_3 = cor(umaps$X3,meta$batch,method = 'pearson')

corr_batch_4 = cor(umaps$X4,meta$batch,method = 'pearson')

corr_batch_5 = cor(umaps$X5,meta$batch,method = 'pearson')

corr_batch_6 = cor(umaps$X6,meta$batch,method = 'pearson')

#

group_list1 = meta$condition

group_list2 = meta$unique_reads

group_list3 = meta$Peaks

group_list4 = meta$NSC


group_list5 = meta$FRiP

group_list6 = meta$sequencing_depth

group_list7 = meta$batch

#
```

```
colData=data.frame(row.names = colnames(group_B1A1),
                   condition=group_list1,
                   reads=group_list2,
                   peaks=group_list3,
                   NSC=group_list4,
                   FRiP = group_list5,
                   sequencing_depth = group_list6,
                   batch = group_list7)
colData$condition = factor(colData$condition,levels = c('B','A'))
colData$reads =as.numeric(colData$reads)
colData$peaks=as.numeric(colData$peaks)
colData$NSC=as.numeric(colData$NSC)
colData$FRiP=as.numeric(colData$FRiP)
colData$sequencing_depth=as.numeric(colData$sequencing_depth)
colData$batch=factor(colData$batch,)
#
dds=DESeqDataSetFromMatrix(countData = group_B1A1,
                           colData = colData,
                           design = ~ condition)
# #tmp = model.matrix(~group_list + gender + Age + PMI, colData)


dds=DESeq(dds)
#
```

```
res_NACM=results(dds,
                contrast = c("condition","B","A"))
resOrdered_NACM=res_NACM[order(res_NACM$padj),]
DEG_NACM=as.data.frame(resOrdered_NACM)
DEG_NACM=na.omit(DEG_NACM)
nrDEG_NACM=DEG_NACM[which(DEG_NACM$padj<0.05 &
abs(DEG_NACM$log2FoldChange) > 0.6),]
choose_gene_neuron=rownames(nrDEG_NACM)
choose_matrix=group_B1A1[choose_gene_neuron,]
DEG_matrix = data.frame(nrDEG_NACM,choose_matrix)
up = nrDEG_NACM[which(nrDEG_NACM$log2FoldChange > 0.6),]
down = nrDEG_NACM[which(nrDEG_NACM$log2FoldChange < -0.6),]
diff_peak = peak_info[rownames(DEG_matrix),]
new_DEG_matrix = bind_cols(diff_peak,DEG_matrix)
setwd("C:/data/LSD_Op/ChIP_seq")
all_NACM = group_B1A1[rownames(DEG_NACM),]
all_matrix = data.frame(DEG_NACM,all_NACM)
write.csv(all_matrix,'All_B1A1_NACMF.csv')
write.csv(new_DEG_matrix,'Differential_B1A1_NACMF.csv')
dat <- read.csv("Differential_B1A1_NACMF.csv")
dat1 <- dat[,2:4]
write.table(dat1,'Differential_B1A1_NACMF_out.bed',
          row.names = FALSE, col.names = FALSE, sep="\t", quote = FALSE)
```

8.4 Differential enhancer matrix:

```
############preparation##############

library(ggplot2)

library(dplyr)

library(stats)

library(gplots)

library(tidyverse)

library(pheatmap)

setwd("C:/data/LSD_Op/ChIP_seq")

dataframe_A <- read.csv("counts_NAC_MF.csv")

bed_file1 <- read.table("C1A1_NAC_combined_Annotated.bed", header = FALSE,
stringsAsFactors = FALSE)

bed_file2 <- read.table("D1A1_NAC_combined_Annotated.bed", header = FALSE,
stringsAsFactors = FALSE)

bed_file3 <- read.table("C1D1_NAC_combined_Annotated.bed", header = FALSE,
stringsAsFactors = FALSE)

bed_file4 <- read.table("G1E1_NAC_combined_Annotated.bed", header = FALSE,
stringsAsFactors = FALSE)

bed_file5 <- read.table("H1E1_NAC_combined_Annotated.bed", header = FALSE,
stringsAsFactors = FALSE)

bed_file6 <- read.table("G1H1_NAC_combined_Annotated.bed", header = FALSE,
stringsAsFactors = FALSE)

# Combine the six data frames into one
```

```
dataframe_B <- rbind(bed_file1, bed_file2, bed_file3, bed_file4, bed_file5, bed_file6)

# Get unique values from column 2 of dataframe B

values_to_match <- unique(dataframe_B[, 2])

# Filter rows from dataframe A where column 2 matches any value from dataframe B

matching_rows <- dataframe_A[dataframe_A[, 2] %in% values_to_match, ]

write.csv(matching_rows, "NACMF_3comp_DP_Count.csv", row.names = FALSE)

#check for duplicate start values

# Check for duplicates in "START" column

duplicate_rows <- matching_rows[duplicated(matching_rows$START) |
duplicated(matching_rows$START, fromLast = TRUE), ]

# Print rows with duplicates in "START" column

if (nrow(duplicate_rows) > 0) {

  cat("Duplicate values in START column:\n")

  print(duplicate_rows)

} else {

  cat("No duplicate values found in START column.\n")

}

##May have to remove duplicates manually, raise start position by 1 then save and
read-in file

matching_rows <- read.csv("NACMF_3comp_DP_Count.csv")

# Create a new dataframe with columns from the 5th column onwards

new_dataframe <- matching_rows %>% select(5:ncol(matching_rows))

# Normalize the data using L2 norm
```

```
normalize_l2 <- function(x) {
  norm <- sqrt(sum(x^2))
  return(x / norm)
}
# Apply the normalization function to each row
normalized_rows <- t(apply(new_dataframe, 1, normalize_l2))
# Convert the result back to a dataframe
normalized_dataframe <- as.data.frame(normalized_rows)
# Extract columns containing Group.A, Group.B, Group.C, and Group.D
GroupA_columns <- normalized_dataframe %>%
  select(matches("A."))
GroupC_columns <- normalized_dataframe %>%
  select(matches("C."))
GroupD_columns <- normalized_dataframe %>%
  select(matches("D."))
GroupE_columns <- normalized_dataframe %>%
  select(matches("E."))
GroupG_columns <- normalized_dataframe %>%
  select(matches("G."))
GroupH_columns <- normalized_dataframe %>%
  select(matches("H."))
# Calculate the mean for each Group
GroupA_mean <- rowMeans(GroupA_columns, na.rm = TRUE)
```

```
GroupC_mean <- rowMeans(GroupC_columns, na.rm = TRUE)

GroupD_mean <- rowMeans(GroupD_columns, na.rm = TRUE)

GroupE_mean <- rowMeans(GroupE_columns, na.rm = TRUE)

GroupG_mean <- rowMeans(GroupG_columns, na.rm = TRUE)

GroupH_mean <- rowMeans(GroupH_columns, na.rm = TRUE)

row_means <- rowMeans(cbind(GroupA_mean, GroupC_mean, GroupD_mean,
GroupE_mean, GroupG_mean, GroupH_mean), na.rm = TRUE)

# Create the new dataframe

input_1 <- data.frame(

  start = matching_rows[,2],

  GroupA = GroupA_mean,

  GroupC = GroupC_mean,

  GroupD = GroupD_mean,

  GroupE = GroupE_mean,

  GroupG = GroupG_mean,

  GroupH = GroupH_mean,

  Average = row_means # Column name changed to "Average"

)

input_2 <- data.frame(

  CHR = matching_rows[,1],

  START = matching_rows[,2],

  END = matching_rows[,3],

  GroupA = GroupA_mean,
```

```
  GroupC = GroupC_mean,
  GroupD = GroupD_mean,
  GroupE = GroupE_mean,
  GroupG = GroupG_mean,
  GroupH = GroupH_mean,
  Average = row_means # Column name changed to "Average"
)
input_3 <- data.frame(START = matching_rows[,2], Average = row_means,
normalized_dataframe)
input_4 <- data.frame(matching_rows[,1:3], Average = row_means,
normalized_dataframe)
# Write the new dataframe to a CSV file
write.csv(input_1, "NACMF_3comp_6col_Grouped_input_norm.csv", row.names =
FALSE)
write.csv(input_2, "NACMF_3comp_6col_Grouped_input_norm_2.csv", row.names =
FALSE)
write.csv(input_3, "NACMF_3comp_6col_input_norm.csv", row.names = FALSE)
write.csv(input_4, "NACMF_3comp_6col_input_norm_2.csv", row.names = FALSE)
```

8.5 Clustered Pearson correlation coefficient

```
############Pearson Correlation ###############

my_data2 <- read.csv("combined_NACMF_3comp_6col_Grouped_input.csv")

colnames(my_data2) <- c("start", "VEH-VEH, Male","OXY-VEH, Male", "OXY-PSI, Male", "VEH-VEH, Female","OXY-VEH, Female", "OXY-PSI, Female", "Average", "Cluster")

custom_order <- c(1, 2, 3, 4)

# Use match function to create a temporary ordering column based on custom_order

my_data2$order <- match(my_data2$Cluster, custom_order)


# Sort my_data2 by the new order column

my_data2 <- my_data2[order(my_data2$order), ]

# Remove the temporary order column

my_data2$order <- NULL

cluster1_data <- my_data2[my_data2$Cluster == 1, ]

cluster2_data <- my_data2[my_data2$Cluster == 2, ]

cluster3_data <- my_data2[my_data2$Cluster == 3, ]

cluster4_data <- my_data2[my_data2$Cluster == 4, ]

C1 <- cluster1_data[,2:7]

C2 <- cluster2_data[,2:7]

C3 <- cluster3_data[,2:7]

C4 <- cluster4_data[,2:7]
```

```
calculate_row_correlations <- function(df) {
  # Initialize an empty list to store correlations
  correlations <- numeric(nrow(df))
  # Loop through each row
  for (i in 1:nrow(df)) {
    # Extract the vectors for the current row
    group1 <- c(df[i, 1], df[i, 2], df[i, 3])
    group2 <- c(df[i, 4], df[i, 5], df[i, 6])
    # Compute the Pearson correlation coefficient
    correlation <- cor(group1, group2, method = "pearson")
    # Store the result in the list
    correlations[i] <- correlation
  }
  # Return the list of correlations
  return(correlations)
}
# Call the function with the dataframe
correlations <- calculate_row_correlations(C1)
correlations_df <- data.frame(Correlation = correlations)
# Append the new column to the original dataframe
C1_with_correlations <- cbind(C1, Correlation = correlations_df$Correlation)
# Print the result
median_correlation <- median(C1_with_correlations$Correlation, na.rm = TRUE)
```

```
average_correlation <- mean(C1_with_correlations$Correlation, na.rm = TRUE)
sorted_df <- C1_with_correlations %>%
  arrange(Correlation)
# Select the top 100 rows with the lowest values
lowest_100_rows <- head(sorted_df, 100)
# Create a new dataframe with these rows
new_dataframe <- lowest_100_rows
calculate_row_correlations2 <- function(df) {
  # Initialize an empty list to store correlations
  correlations <- numeric(nrow(df))
  # Loop through each row
  for (i in 1:nrow(df)) {
    # Extract the vectors for the current row
    group1 <- c(df[i, 1], df[i, 2], df[i, 3])
    group2 <- c(df[i, 4], df[i, 5], df[i, 6])
    # Compute the Pearson correlation coefficient
    correlation <- cor(group1, group2, method = "pearson")
    # Store the result in the list
    correlations[i] <- correlation
  }
  # Return the list of correlations
  return(correlations)
}
```

```
# Call the function with the dataframe
correlations2 <- calculate_row_correlations2(C2)
correlations_df2 <- data.frame(Correlation = correlations2)
# Append the new column to the original dataframe
C2_with_correlations <- cbind(C2, Correlation = correlations_df2$Correlation)
# Print the result
median_correlation2 <- median(C2_with_correlations$Correlation, na.rm = TRUE)
average_correlation2 <- mean(C2_with_correlations$Correlation, na.rm = TRUE)
calculate_row_correlations3 <- function(df) {
  # Initialize an empty list to store correlations
  correlations <- numeric(nrow(df))
  # Loop through each row
  for (i in 1:nrow(df)) {
    # Extract the vectors for the current row
    group1 <- c(df[i, 1], df[i, 2], df[i, 3])
    group2 <- c(df[i, 4], df[i, 5], df[i, 6])
    # Compute the Pearson correlation coefficient
    correlation <- cor(group1, group2, method = "pearson")
    # Store the result in the list
    correlations[i] <- correlation
  }
  # Return the list of correlations
  return(correlations)
```

```
}
# Call the function with the dataframe
correlations3 <- calculate_row_correlations3(C3)
correlations_df3 <- data.frame(Correlation = correlations3)
# Example dataframe C1 (Replace this with your actual dataframe)
# Append the new column to the original dataframe
C3_with_correlations <- cbind(C3, Correlation = correlations_df3$Correlation)
# Print the result
median_correlation3 <- median(C3_with_correlations$Correlation, na.rm = TRUE)
average_correlation3 <- mean(C3_with_correlations$Correlation, na.rm = TRUE)
calculate_row_correlations4 <- function(df) {
  # Initialize an empty list to store correlations
  correlations <- numeric(nrow(df))
  # Loop through each row
  for (i in 1:nrow(df)) {
    # Extract the vectors for the current row
    group1 <- c(df[i, 1], df[i, 2], df[i, 3])
    group2 <- c(df[i, 4], df[i, 5], df[i, 6])
    # Compute the Pearson correlation coefficient
    correlation <- cor(group1, group2, method = "pearson")
    # Store the result in the list
    correlations[i] <- correlation
  }
```

```
  # Return the list of correlations
  return(correlations)
}
# Call the function with the dataframe
correlations4 <- calculate_row_correlations4(C4)
correlations_df4 <- data.frame(Correlation = correlations4)
# Example dataframe C1 (Replace this with your actual dataframe)
# Append the new column to the original dataframe
C4_with_correlations <- cbind(C4, Correlation = correlations_df4$Correlation)
# Print the result
median_correlation4 <- median(C4_with_correlations$Correlation, na.rm = TRUE)
average_correlation4 <- mean(C4_with_correlations$Correlation, na.rm = TRUE)
C1_hold <-C1_with_correlations[,7]
C2_hold <-C2_with_correlations[,7]
C3_hold <-C3_with_correlations[,7]
C4_hold <-C4_with_correlations[,7]
data_list <- list(Cluster1 = C1_hold, Cluster2 = C2_hold, Cluster3 = C3_hold, Cluster4 =
C4_hold)
# Create the box plot
boxplot(data_list,
        main = "NAc Pairwise Pearson Correlation Coefficient",  # Title of the plot
        #xlab = "Cluster",                    # X-axis label
        #ylab = "PPCC",                      # Y-axis label
```

```
col = c("#87CEEB", "#99B2FF", "#6699FF", "#000080"),  # Colors for each box
border = "black",                    # Border color of the boxes
notch = TRUE,                         # Add notches to the boxes
outline = FALSE)                      # Do not show outliers
```

8.6 Hi-C linkage

```
############preparation##############
library(dplyr)
library(ggplot2)
setwd("C:/ChIP_seq")
# Read the two input files
file1 <- read.table("Differential_B1A1_NACMF_out.bed", header = FALSE)
file2 <- read.csv("Cortex_HiC.csv", header = FALSE)
# Create an empty data frame to store matching rows
matched_rows <- data.frame()
# Create an empty data frame to store non-matching rows from file1
non_matched_rows <- data.frame()
# Iterate through each row in file1
for (i in 1:nrow(file1)) {
  # Get the current row from file1
  row1 <- file1[i, ]
  # Find matching rows in file2 based on column 1
  matching_rows_file2 <- file2[file2$V1 == row1$V1, ]
  # Initialize a flag to determine if the row1 satisfies the conditions
  satisfied_conditions <- FALSE
  # Iterate through matching rows in file2
  for (j in 1:nrow(matching_rows_file2)) {
```

```
    # Get the current matching row from file2
    row2 <- matching_rows_file2[j, ]
    # Check conditions
    if (row1$V2 <= row2$V3 && row1$V3 >= row2$V2) {
      # Set the flag to TRUE if conditions are met
      satisfied_conditions <- TRUE
      # Add the row from file1 to the data frame if conditions are met
      # Also add column 4 of file 2 to the new data frame
      matched_row <- c(row1, row2[4])  # Combine columns from both rows
      matched_rows <- rbind(matched_rows, matched_row)
      break  # Break out of the inner loop once a match is found
    }
  }
  # If conditions are not met, add the row to non_matched_rows
  if (!satisfied_conditions) {
    non_matched_rows <- rbind(non_matched_rows, row1)
  }
}
# Write the matched_rows data frame to a BED file
write.table(matched_rows, "B1A1_NAC_matched_rows.bed", quote = FALSE,
row.names = FALSE, col.names = FALSE, sep = "\t")
# Write the non_matched_rows data frame to a BED file
```

```
write.table(non_matched_rows, "B1A1_NAC_non_matched_rows.bed", quote = FALSE,
row.names = FALSE, col.names = FALSE, sep = "\t")
```

8.7 Peak annotation

```
setwd("C:/ChIP_seq")
library(ChIPseeker)
library(TxDb.Mmusculus.UCSC.mm10.knownGene)
library(clusterProfiler)
library(stringr)
library(reshape2)
library(ggplot2)
library(ggpubr)
rm(list = ls())
options(stringsAsFactors = F)
txdb <- TxDb.Mmusculus.UCSC.mm10.knownGene
setwd("C:/ChIP_seq")
bedPeaksFile = "B1A1_NAC_non_matched_rows.bed"
peak <- readPeakFile(bedPeaksFile)
keepChr=!grepl('_',seqlevels(peak))
seqlevels(peak, pruning.mode="coarse") <- seqlevels(peak)[keepChr]
cat(paste0('There are ', length(peak), 'peaks for this data'))
peakAnno <- annotatePeak(peak,
                         tssRegion = c(-2000, 2000),
                         TxDb = txdb,
                         annoDb = "org.Mm.eg.db")
```

```
peakAnno_df = as.data.frame(peakAnno)

## Remove Promoters

# Define a function to filter rows based on column 14

remove_outliers <- function(data) {

  # Filter rows where column 14 is not within the range (-2000, 2000)

  filtered_data <- data[!(data[, 14] > -2000 & data[, 14] < 2000), ]

  return(filtered_data)

}

# Example usage:

# Assuming df is your data frame, call the function to remove outliers

# Replace df with the name of your actual data frame

new_peakAnno_df <- remove_outliers(peakAnno_df)

# Now, filtered_df contains the data frame with rows removed where the value of
column 14 is within the range (-2000, 2000)

#### Only use below if you are cutting post HiC comp non-matched peaks

new_peakAnno_df = new_peakAnno_df[,-17]

new_peakAnno_df = new_peakAnno_df[,-4:-15]

####

##Combine the HiC & nearest neighbor enhancer annotations

HiC_Match <- read.table("B1A1_NAC_matched_rows.bed", header = FALSE)

colnames(HiC_Match)=c('seqnames', 'start','end', 'SYMBOL')

combined_bed <- rbind(new_peakAnno_df, HiC_Match)
```

```
write.table(combined_bed, "B1A1_NAC_combined_Annotated.bed", quote = FALSE,
row.names = FALSE, col.names = FALSE, sep = "\t")
```

8.8 K-means clustering

```
############preparation##############
library(ggplot2)
library(dplyr)
library(stats)
library(gplots)
library(tidyverse)
library(pheatmap)
my_data1 <- read.csv("NACMF_male_3comp_6col_Grouped_input_norm.csv")
rownames(my_data1) <- my_data1$start
# Select the column you want to use for clustering
my_data <- my_data1[, 2:4]  # Assuming you want to use the average column for clustering
# Perform k-means clustering, number of clusters determined by silhouette
k <- 4
set.seed(123)
kmeans_result <- kmeans(my_data, centers = k, nstart = 25)
# Get cluster assignments
cluster_labels <- kmeans_result$cluster
# Initialize list to store genes in each cluster
cluster_gene_lists <- vector("list", length = k)
# Fill the list with genes in each cluster
```

```
for (cluster in 1:k) {
  genes_in_cluster <- rownames(my_data1)[cluster_labels == cluster]
  cluster_gene_lists[[cluster]] <- genes_in_cluster
}
# Export gene lists to CSV
for (cluster in 1:k) {
  cluster_filename <- paste("cluster_", cluster,
"_genes_PFCMF_6col_grouped_norm4.csv", sep = "")
  write.csv(cluster_gene_lists[[cluster]], file = cluster_filename, row.names = FALSE)
}
my_data2 <- read.csv("NACMF_male_3comp_6col_Grouped_input_norm_2.csv")
my_data2$Cluster <- cluster_labels
MDT <- my_data2[,-3:-7]
MDT1 <- MDT[,-1]
write.csv(MDT1, "NACMF_male_3comp_6col_Grouped_Clusters.csv", row.names =
FALSE)
###Reorder clusters to 1, 2, 3, 4
data <- read_csv("NACMF_male_3comp_6col_Grouped_Clusters.csv", col_names =
TRUE)
# Map values in the second column according to your transformation
data$Cluster <- case_when(
  data$Cluster == 2 ~ 1,
  data$Cluster == 1 ~ 2,
```

```
  data$Cluster == 4 ~ 3,
  data$Cluster == 3 ~ 4,
  TRUE ~ data$Cluster  # handle any other values if necessary
)
# Write the modified data back to a CSV file
write_csv(data, "modified_PFCMF_male_3comp_6col_Grouped_Clusters.csv",
col_names = FALSE)
# Load necessary libraries
library(dplyr)
# File paths of your two CSV files
file1_path <- "modified_PFCMF_male_3comp_6col_Grouped_Clusters.csv"
file2_path <- "modified_PFCMF_female_3comp_6col_Grouped_Clusters.csv"
# Read the first CSV file
data1 <- read.csv(file1_path, header = FALSE, stringsAsFactors = FALSE)
colnames(data1) <- c("Column1", "Column2")
# Read the second CSV file
data2 <- read.csv(file2_path, header = FALSE, stringsAsFactors = FALSE)
colnames(data2) <- c("Column1", "Column2")
# Combine the two datasets
combined_data <- bind_rows(data1, data2)
# Remove duplicates based on Column1
combined_data_unique <- combined_data %>%
  distinct(Column1, .keep_all = TRUE)
```

```
# Write the combined and de-duplicated data to a new CSV file

write.csv(combined_data_unique,
"combined_PFCMF_3comp_6col_Grouped_Cluster_Numbers.csv", row.names =
FALSE)

# Load necessary libraries (step6)

library(dplyr)

# Read file 1 (assuming 8 columns)

file1 <- read.csv("PFCMF_3comp_6col_Grouped_input_norm.csv", header = TRUE)

# Read file 2 (assuming 2 columns)

file2 <- read.csv("combined_PFCMF_3comp_6col_Grouped_Cluster_Numbers.csv",
header = TRUE)

# Rename columns for clarity

colnames(file1) <- paste0("V", 1:8)

colnames(file2) <- c("Column1", "Column2")

# Perform left join to add Column2 values from file2 to file1

combined_data <- left_join(file1, file2, by = c("V1" = "Column1"))

# Mutate to update V9 based on matching values

combined_data <- combined_data %>%

  mutate(V9 = ifelse(!is.na(Column2), Column2, V9)) %>%

  select(-Column1, -Column2)

# Write the combined data with updated column 9 to a new CSV file

write.csv(combined_data, "combined_PFCMF_3comp_6col_Grouped_input.csv",
row.names = FALSE)

my_data2 <- read.csv("combined_PFCMF_3comp_6col_Grouped_input.csv")
```

```
colnames(my_data2) <- c("start", "VEH-VEH, Male","OXY-VEH, Male", "OXY-PSI, Male", "VEH-VEH, Female","OXY-VEH, Female", "OXY-PSI, Female", "Average", "Cluster")

custom_order <- c(1, 2, 3, 4)


# Use match function to create a temporary ordering column based on custom_order

my_data2$order <- match(my_data2$Cluster, custom_order)

# Sort my_data2 by the new order column

my_data2 <- my_data2[order(my_data2$order), ]

# Remove the temporary order column

my_data2$order <- NULL

# rownames(my_data2) <- my_data2$START

mydata3 <- my_data2[, 2:7]

# Convert to matrix for heatmap plotting

mydata3 <- as.matrix(mydata3)

library(pheatmap)

phtmap <- pheatmap::pheatmap(mydata3)

unique_labels <- unique(my_data2$Cluster)

labels_row <- rep("", nrow(mydata3))

for (i in 1:length(unique_labels)) {

  indices <- which(my_data2$Cluster == unique_labels[i])

  labels_row[indices[1]] <- paste("Cluster", as.character(unique_labels[i]))

}
```

```
breaks <- seq(0, 0.3, by = 0.001172)

# Plot heatmap with custom column clustering

figure <- pheatmap(mydata3,

                   scale = "none",

                   breaks = breaks,

                   cluster_rows = FALSE,       # Turn off row clustering

                   cluster_cols = FALSE,

                   #labels_row = labels_row,

                   border_color = NA,          # No border

                   fontsize_row = 8,           # Adjust row label font size

                   fontsize = 8,               # Adjust font size for cell values

                   color = colorRampPalette(c("white", "#ADD8E6" ,"#000080" ))(256))  #
Color palette
```

8.9 GO terms

```
###Prep

library(ggplot2)

library(ChIPseeker)

library(clusterProfiler)

###Gene Ontology

dat5 <- read.table("C1D1_PFC_combined_Annotated.bed")

dat6 <- dat5[,4]

Go_result_up <- enrichGO(dat6, 'org.Mm.eg.db', ont = "BP", keyType="SYMBOL",
pvalueCutoff = 0.05, pAdjustMethod = "BH")

ggo <- Go_result_up

ggo@result$logP = -log10(ggo@result$p.adjust)

print(max(ggo@result$logP))

ggo@result$Description <- factor(ggo@result$Description, levels =
rev(ggo@result$Description), ordered=TRUE)

sig_ggo <- ggo@result[ggo@result$p.adjust < 0.1,]

fig<-ggplot(sig_ggo[1:min(nrow(sig_ggo),15),], aes(x=logP,y=Description)) +

  xlab("-logP") +

  ylab("Biological Processes") +

  coord_cartesian(xlim=c(0,10)) +

  theme_classic()

print(fig)

dotplot(ggo)
```