Published in final edited form as:

J Neurosci Methods. 2018 October 01; 308: 62–73. doi:10.1016/j.jneumeth.2018.07.006.

Single Sample Sequencing (S3EQ) of epigenome and transcriptome in nucleus accumbens

SJ Xu and

Department of Systems Pharmacology and Translational Therapeutics, Perelman School of Medicine, The University of Pennsylvania, Philadelphia, PA, USA

EA Heller

Department of Systems Pharmacology and Translational Therapeutics and Penn Epigenetics Institute, Perelman School of Medicine, The University of Pennsylvania, Philadelphia, PA, USA
eheller@pennmedicine.upenn.edu

Abstract

BACKGROUND: High-throughput sequencing has been widely applied to uncover the molecular mechanisms underlying neurological and psychiatric disorders. The large body of data support the role of epigenetic mechanisms in neurological function of both human and animals. Yet, the existing data is limited by the fact that epigenetic and transcriptomic changes have only been measured in separate cohorts. This has limited precise correlation of epigenetic changes in gene expression.

NEW METHOD: Single Sample Sequencing (S3EQ) is an innovative approach to analyze both epigenetic and transcriptomic regulation within a single neuronal sample. Using this method, we analyzed chromatin immunoprecipitation (ChIP)- and RNA-sequencing data from the nucleus accumbens (NAc) of the same animal.

RESULTS: ChIP-S3EQ of neuronal nuclei reliably identified hPTM enrichment in the adult mouse NAc with high precision. Comparing cellular compartments, we found that the spliceosome of whole cell RNA-seq was more closely recapitulated by cytosolic RNA-S3EQ than nuclear RNA-seq. Finally, S3EQ showed increased sensitivity for correlating chromatin modifications with gene expression, especially for lowly expressed transcripts.

COMPARISON WITH EXISTING METHODS: S3EQ accurately generates both RNA- and ChIP-seq from a single sample, providing a clear advantage over existing methods which require two samples. ChIP-S3EQ performance was comparable to ChIP-seq, while RNA-S3EQ generated an almost identical expression profile to nuclear-enriched and whole cell RNA-seq. Finally, we directly compared RNA-seq by cellular compartments, addressing a limitation of RNA-seq studies limited to neuronal nuclei.

Corresponding author EA Heller eheller@pennmedicine.upenn.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

CONCLUSION: The S3EQ method can be applied to improve the correlative power of transcriptomic and epigenomic studies in neuronal tissue.

Keywords

High-throughput sequencing; ChIP-seq; RNA-seq; S3EQ

1 INTRODUCTION

The extensive use of high-throughput sequencing in neuroscience has led to a wealth of discoveries on the molecular mechanisms underlying neuronal plasticity, as well as neurological and psychiatric disorders in myriad brain regions. This large body of data supports the role of global gene regulation in a broad range of experiences in both human and animal tissues and has led to identification of the epigenetic regulation of specific target gene expression (Cahill et al. 2016; Parikshak et al. 2016; Sanders et al. 2017; Damez-Werno et al. 2016; Beyeler et al. 2017; Mews et al. 2017; Taniguchi et al. 2012; Taniguchi M, Carreira MB, Cooper YA, Bobadilla AC, Heinsbroek JA, Koike N, Larson EB, Balmuth EA, Hughes BW, Penrod RD, Kumar J, Smith LN, Guzman D, Takahashi JS, Kim TK, Kalivas PW, Self DW, Lin Y 2017; Zovkic et al. 2014; Day et al. 2013; Labonté et al. 2017; Jiang et al. 2015; Nott et al. 2016). It is now clear from preclinical literature that various behavioral paradigms, models of aging and neurodegeneration, and models of intellectual disability cause global alterations of histone post-translational modifications (hPTMs), which correlate with changes in mRNA expression. Yet, interpretation of these data is limited by the fact that the chromatin and gene expression changes have thus far been measured in separate cohorts of animals, making it difficult to deduce the direct causal relevance of a given hPTM to specific gene expression. To address this limitation, we developed Single Sample Sequencing (S3EQ), which allows a within-sample examination of both global chromatin modifications and mRNA expression. This method seeks both improve the correlative power and reproducibility of observed changes and to minimize the number of subjects required for a given high-throughput sequencing experiment.

A number of landmark next generation sequencing studies have utilized bulk tissue dissection to examine changes in global hPTMs and mRNA expression. Such an approach is commonly applied to neuropsychiatric disease, particularly that of affective disorders. One key brain region that plays significant role in addiction and depression is the nucleus accumbens (NAc), which mediates cognitive processing of motivation, aversion, reward, and reinforcement learning (Calipari et al. 2016; Saddoris et al. 2015; Day & Carelli 2007). RNA-seq of human brain reward regions in control and depressed patients, or in of mice following chronic stress reveal conserved and sexually-dimorphic stress-regulated gene expression (Bagot et al. 2016; Labonté et al. 2017). Another comprehensive study analyzed mouse NAc using chromatin immunoprecipitation (ChIP)-sequencing (seq) and RNA-seq following cocaine administration (J Feng et al. 2014) to correlate combinatorial patterns of hPTM regulation with cocaine-regulated gene expression. Beyond addiction and depression, next generation sequencing studies have been applied to neurodevelopmental diseases, such as Rett's Syndrome, in order to examine global enrichment of histone deacetylase 3 (HDAC3), methyl CpG binding protein (MeCP2) and histone H4 lysine 12 acetylation

(H3K12ac), histone H3 lysine 9 acetylation (H3K9ac), transcription factor FOXO (Nott et al. 2016). However, to date all studies have analyzed changes in gene expression and hPTMs in separate tissue samples and, in the case of rodent studies, have pooled a large number of subjects for each sample.

In addition to the limitations arising from correlating chromatin and gene expression changes measured in separate samples, the extreme diversity of intermingled cell types in neural tissue limits understanding of the locus of effect of the observed changes. In the studies cited above, bulk tissue dissection isolates a particular brain region but cannot distinguish gene regulatory events between cell types (Karpagam Srinivasan et al. 2016). One approach applied to combat this limitation is fluorescence-activated cell sorting (FACS), which separates neuronal from non-neuronal nuclei (von Schimmelfmann et al. 2016; Lobo et al. 2006; Liu et al. 2017; Jiang et al. 2015). While FACS addresses the issue of cellular heterogeneity, it is unclear if RNA-seq of nuclei alone can capture the full extent of gene expression in bulk tissue samples that include nuclei, soma and processes. A second major limitation of FACS is the relatively low recovery of nuclei, which therefore requires pooling of a large number of animal or human samples (Finegersh & Homanics 2016; von Schimmelfmann et al. 2016; Jiang et al. 2015). Acquisition of large sample cohorts is not always feasible especially with human tissue, complex animal behavioral studies or low throughput animal disease models, including those cited above (Nott et al. 2016; Labonté et al. 2017). Limited tissue yield following FACS contributes to the fact that all publications to date report RNA- and ChIP-seq analyses from separate brain tissue samples, further impeding reliable correlation between gene expression and epigenetic modifications.

The S3EQ method directly addresses these limitations by measuring within-sample transcriptomic and epigenomic profiles. Using this approach, we found that histone H3 lysine 36 trimethylation (H3K36me3) enrichment in adult mouse nucleus accumbens (NAc) obtained by ChIP-S3EQ (nuclear-enriched) was identical to that obtained by traditional (whole-cell) ChIP-seq. In addition, we directly compared the (polyA+)RNA content of nuclear-enriched, cytosolic and whole cell samples. We found that while there were negligible differences in overall gene expression profiles between nuclear-enriched and whole cell RNA-seq, the splicing profile of cytosolic RNA obtained by ChIP-S3EQ more closely resembled that of total RNA. Together, our data provide a valuable resource to detect neuronal gene expression by compartment and demonstrate that S3EQ enables high precision, next-generation ChIP- and RNA-sequencing from a single sample.

2 MATERIALS AND METHODS

2.1 Animals

Adult, male, 8-week old, C57BL/6J mice (Jackson) were used in this study. Mice were housed five per cage on a 12-h light-dark cycle at constant temperature (23°C) with access to food and water *ad libitum*. Mice were maintained under standard conditions consistent with National Institute of Health guidelines and approved by the Institutional Animal Care and Use Committee of the University of Pennsylvania. Food and water were available *ad libitum*.

2.2 Tissue collection

Mice were rapidly euthanized by cervical dislocation and decapitated, brains were extracted and placed into an ice cold adult mouse brain slicer matrix with 1 mm coronal section slice intervals (Zivic Instruments). Nac was identified based on the location of anterior commissure with the approximate anterior-posterior coordinate of +1.6 mm from Bregma. NAc was then dissected using a 2 mm aluminum micro-punch (Harris). Tissue was frozen on dry ice and stored at -80°C until downstream analysis.

2.3 S3EQ chromatin and RNA preparation

The following reagents were made fresh before the start of each experiment: Cell lysis buffer: 10mM Tris-HCl buffer, 10mM NaCl, 3mM MgCl₂ and 0.5% NP-40 in H₂O. Nuclear lysis buffer: 50mM Tris-HCl, 5mM EDTA and 1% SDS in H₂O. ChIP dilution buffer: 16.7mM Tris-HCl, 1.1% Triton-X 100, 0.01% SDS, 167mM NaCl and 1.2mM EDTA in H₂O. IP elution buffer: 0.1M NaHCO₃ and 1% SDS in H₂O. Blocking buffer: 0.5% BSA in PBS. Low salt buffer: 10% SDS, 1% Triton-100, 0.5M EDTA, 1M Tris, 5M NaCl in H₂O. High salt buffer: 10% SDS, 1% Triton-100, 0.5M EDTA, 1M Tris, 5M NaCl

Bilateral 2mm NAc punches were first lysed in 500ul cell lysis buffer. The tissue was fully homogenized with a pellet mixer (VWR 47747-370) until no intact tissue was observed. The cell nuclei were immediately pelleted by centrifuging the homogenate at 1500xg for 5 minutes at 4°C . The protocol was optimized by examining the cell suspension using light microscopy before and after centrifugation. After homogenization and before centrifugation we observed free nuclei and many small particulates. After centrifugation, small particles were removed and nuclei were pelleted. The supernatant was transferred to a fresh tube for cytoplasmic RNA extraction, and the pelleted nuclei were frozen on dry ice for later fixation. Cytoplasmic RNA was extracted according to the Qiagen RNeasy (74004) Supplementary Protocol. Specifically, 600ul buffer RLT and 430ul 100% ethanol was added to the supernatant and mixed by pipetting. Up to 700ul of the mixture were transferred to an RNeasy micro spin column (74004) placed in a 2ml collection tube and centrifuged at 8000xg for 20s. After flow-through was discarded, this step was repeated with the remaining sample and the sample spin column. 700ul buffer RW1 and 500ul buffer RPE were then added to the spin column and centrifuged at 8000xg for 15s to wash the membrane. RNeasy spin column was then placed in a new collection tube and centrifuged at full speed for 1 minute. RNeasy spin columns were then transferred to a new tube, and RNA was eluted in 30ul H₂O by centrifuging at 12000xg for 1 minute.

Cell nuclei pellets were re-suspended by pipetting in 450ul PBS buffer (Fisher BP399) with Protease Inhibitor (PI) Cocktail Tablets (Sigma 4693159001). Samples were cross-linked in a final concentration of 1 % formaldehyde by addition of 50ul 11 % formaldehyde (Thermo Scientific 28906) to each tube. Nuclei suspension was rotated in 500ul 1% formaldehyde for 6 minutes at room temperature, and the cross-linking reaction was terminated by addition of 100ul of 1M glycine. The samples were rotated at room temperature for an additional 5 minutes, then centrifuged at 5500xg for 5 minutes at 4°C . The supernatants were carefully removed from each sample and pellets were resuspended in 200ul nuclear lysis buffer with PI. Samples were then transferred to 1.5 ml Bioruptor® Plus TPX microtubes

(C30010010-50), and sheared using a Diagenode Bioruptor XL at high sonication intensity for 30 minutes (30 s on/30 s off).

To examine chromatin fragment sizes and DNA concentration, 10µl chromatin from each sample was removed, and resuspended in 180µl IP elution buffer, 8µl 5M NaCl, 4µl 0.5M EDTA and 1 µl Proteinase K (10mg/ml) (Fisher BP170050,) were added for reversecrosslinking and protein degradation, and incubate at 65C overnight. DNA was then extracted using MinElute Reaction Cleanup Kit (Qiagen 28206), following the manufacturer's protocol. Specifically, 1ml buffer PB was mixed with each sample before binding to the column, which was subsequently washed twice with 750µl buffer PE before drying and elution in 30µl H2O. Fragment size was verified at 150–300bp with an Agilent 2100 bioanalyzer using DNA High-sensitivity chip (Agilent 5067–4626). DNA purity was verified using nanodrop to ensure 260/280 ratio is ~1.8.

2.4 ChIP

Immunoprecipitation was carried out on the remaining sample using magnetic beads (Dynabeads M-280, Life Technologies). Beads were washed 3x in 500µl blocking buffer and re-suspended in ChIP dilution buffer. Antibody to H3K36me3 (3 µg/IP) (Abcam ab9050) was conjugated to beads by incubation for 6 – 16 hours at 4°C with constant rotation. Beads were washed and re-suspended in 500ul ChIP dilution buffer.

Chromatin quantity per sample was calculated to determine the volume of 1 µg and added to the antibody-conjugated beads. 10% of chromatin was used as input control, as described above. After incubating overnight, each sample was briefly centrifuged, separated from the beads (DynaMag) and discarded. The beads were then washed in order with low-salt buffer, high-salt buffer and TE buffer. 200µl ChIP elution buffer was added after last wash, and samples were eluted in a shaker (Eppendorf ThermoMixer® C), at 1000rpm at 65C for 30 minutes. 8ul 5M NaCl, 4µl 0.5M EDTA and 1µl (10mg/ml) Proteinase K were added for reverse-crosslinking and proteinase treatment. MinElute Reaction Cleanup Kit (Qiagen 28206) system was used to extract DNA as described above. DNA purity was verified using nanodrop to ensure 260/280 ratio is ~1.8.

ChIP-sequencing libraries were prepared using the NEBNext® ChIP-Seq Library Prep Master Mix Set for Illumina® (E6240L) using an adaptor oligo dilution of 1:20 for all samples. Prepared samples were pooled and sequenced on Illumina Hiseq 4000 platform to achieve ~30 million reads per sample. Raw sequencing data was processed to generate fastq files of 50 bp single-end reads for further processing.

2.5 Whole cell chromatin extraction

ChIP was performed on bilateral NAc 2 mm punches collected as described above. Tissue was first cross-linked in 500µl 1% Formaldehyde for 10 minutes at room temperature, and cross-linking was stopped by adding 100µl of 1M glycine. Samples were washed 3x in 500µl ice-cold PBS with protease inhibitor. PBS was then removed and tissue was re-suspended in 1000µl ChIP lysis buffer with PI. The NAc was fully homogenized with a pellet mixer (VWR 47747–370). Homogenate was centrifuged at 4°C and 5500xg for 5 minutes. 200µl ChIP nuclear lysis buffer with PI was added to each sample, and mixed by pipetting until

fully dissolved. Each sample was then transferred to a TPX tube and sheared using a Diagenode Bioruptor XL at high sonication intensity for 30 minutes (30s on/30s off). ChIP was carried out as described above (2.4).

2.6 RNA extraction from cytosolic or nuclear-enriched fractions

Nuclear-enriched and cytosolic RNA preparation was adapted from published protocol (Rio et al. 2010). To prepare cytosolic RNA, bilateral 2mm NAc punches were first lysed in 500µl cell lysis buffer. The NAc was fully homogenized with a pellet mixer (VWR 47747-370). The cell nuclei were immediately pelleted by centrifuging the homogenate at 4°C and 1500xg for 5 minutes. The supernatant was then transferred to a fresh tube without disturbing the nuclear-enriched pellet. The QIAGEN (74004) supplementary protocol was then used to isolate cytoplasmic RNA from the supernatant as described above (2.3).

For nuclear-enriched RNA extraction the pellet was homogenized in 700µl TRIzol (Invitrogen 15596026) using VWR pellet mixer. RNA was precipitated by addition of 140µl chloroform and vigorously mixing for 1 minute. Sample was centrifuged at 12000xg for 15 minutes at 4°C. After centrifugation, the aqueous layer (≈350µl) was removed from the sample and mixed with 600µl buffer RLT and 430µl 100% ethanol. The QIAGEN (74004) supplementary protocol was then used to isolate cytoplasmic RNA from the supernatant as described above (2.3).

2.7 ChIP-seq analysis

Sequences were aligned using Bowtie v2.1.0, allowing for 2 mismatches to the reference in 50bp. Uniquely aligned reads were confirmed for H3K36me3 ChIP-seq and used for subsequent analysis. Peak calling was performed using SICER with FDR < 0.05 as cutoff. Peak calling of IP data was performed normalized to input samples (Xu et al. 2014), whereas DiffBind (Ross-Innes et al. 2012) was used to identify the differential histone modification sites between the two different methods. Integrated profiles were created by averaging the observed signal for each bin for the selected set of relevant genes using deepTools computeMatrix (-a 5000 -b 5000); values displayed indicate fragments per 50-bp bin per million mapped reads.

Enrichment of islands over genome was assessed using HOMER software. By default, HOMER used the DESeq2 rlog transform to create normalized log2 read counts, thereafter generating log2 transformed enrichment relative to a specific genomic category (Fleinz et al. 2010). Gene Ontology was analyzed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) Bioinformatics Resource specifying for biological processes. IGV was used to visualize sequencing coverage.

2.8 RNA-seq analysis

For RNA-seq data, sequences were aligned using STAR-2.5.2a (Dobin et al. 2013). Raw gene read counts with default parameters were used to compute fold changes and significance of expression differences using DESeq (Simon 2010). The log expression values for each gene were averaged over each treatment group, and the log2 fold change was computed. Fleet

maps of raw reads were generated using R's pheatmap function with default parameters. Unsupervised clustering was performed by pheatmap in default settings.

Alternative splicing was quantified using the percent spliced in (PSI) metric using Multivariate Analysis of Transcript Splicing (rMATS, v3.2.4). For each event, rMATS reports counts supporting the inclusion (I) or splicing (S) of an event. To reduce spurious events due to low counts, we required all of the samples to have $I \geq 10$ and $S \geq 10$. For these events, the PSI is calculated as $PSI = I/(I + S)$. IGV was used to visualize sequencing coverage.

3 RESULTS

3.1 Single Sample Sequencing (S3EQ) work-flow

To maximize data obtained from a single sample and more accurately connect histone modifications to gene expression within a brain region of interest, we developed a one-day protocol for performing ChIP using cell nuclei, while also collecting cytosolic RNA of the same samples (Fig. 1). Each S3EQ sample was prepared from the NAc of a single mouse, which was micro-dissected as 1 mm thick $\times$ 2 mm diameter NAc punches. We first homogenized the tissue using cell lysis buffer to break down cell membrane and release nuclei. Low speed centrifugation then separated the cytosol (supernatant) from the nuclei (pellet). Nuclei were then mildly fixed with 1 % formaldehyde for 6 minutes while cytosolic RNA was extracted, eluted on column and prepared for RNA-seq, as detailed below. Similar to the whole cell ChIP method, glycine terminated fixation of the nuclear-enriched fraction, which was then sheared in lysis buffer to fractionate chromatin. Following this, samples were reverse crosslinked, DNA was eluted, purified on column and tested for purity. DNA was then prepared for ChIP-sequencing, as detailed below.

3.2 ChIP-S3EQ performs accurately compared to ChIP-seq

To systematically characterize S3EQ's performance in quantifying global histone modifications in brain, we carried out ChIP-seq of the NAc, a reward-related region in the mouse brain. We performed whole cell ChIP-s3Eq and nuclear-enriched CMIP-S3EQ in parallel on two biological replicates each using an antibody against H3K36me3. The fixation time was optimized to obtain chromatin fragments sheared to the same size (~200bps) from each method (Fig. 2A). We obtained 21.59 Gb in total of ChIP-seq sequence (50bp single-end reads, ~30 million reads per sample) that uniquely mapped to the mouse genome (NCBI37/mm9 assembly). In order to confirm efficacy of both protocols, we compared our H3K36me3 CMIP-S3EQ and ChIP-seq data to a publicly available dataset (Feng *et al.*, 2014, GSE42811) matched by organism, tissue type, developmental stage, and method (antibody and manufacturer, bead and fragment size). Our data appeared indistinguishable from this independent source when examined using the IGV browser (Robinson *et al.* 2011; Thorvaldsdóttir *et al.* 2013). Finally, we found that H3K36me3 enrichment was highly correlated between CMIP-S3EQ and ChIP-seq samples using Spearman's rank-order method with a bin size of 50kb (Fig. 2B).

We next examined H3K36me3 enrichment along annotated gene body regions, extending 5 kb up- and down-stream from the transcription start site (TSS) (Fig. 2C). Our results confirmed previous findings that H3K36me3 is mostly enriched at the gene-body (Sati et al. 2012). Notably, both ChIP-seq and CMIP-S3EQ methods yielded highly similar H3K36me3 signal density along protein coding regions, as well as up-stream of the TSS and down-stream of the transcription end site (TES) (Fig. 2C, D, E). Further, genomic enrichment analysis of H3K36me3 ChIP-seq and CMIP-S3EQ peaks finds genome-wide enrichment over transcribed regions (Fig. 2F). To more closely examine the quality and coverage of H3K36me3 CMIP-S3EQ, we analyzed a control gene, β -actin. As expected, H3K36me3 signal was highly similar between the two methods and enriched primarily at the exon regions (Fig. 2E). Thus, both global and gene-targeted analysis indicated that CMIP-S3EQ reliably identified H3K36me3 enrichment in the adult mouse NAc with high precision.

3.3 Different subcellular compartments exhibit highly similar expression profiles.

Because the S3EQ method involves extraction of both chromatin and RNA from the same tissue, nuclear RNA cannot be recovered. This is an especially prominent issue given the high number of transcriptomic studies of FACS sorted neuronal nuclei (Guo et al. 2014; Cai et al. 2014; Finegersh & Flomanics 2015; Lacar et al. 2016; Jiang et al. 2015). To examine the potential relevance of nuclear or cytosolic RNA loss, we carried out RNA-seq on RNA isolated from either cytosol alone, nuclear-enriched fraction alone, or whole cell. We applied poly(A)-selected RNA-sequencing, as it is the most commonly used sequencing approach for analyzing gene expression and splicing in neurons (Flu et al. 2017; Stefanelli et al. 2018; Aprea et al. 2013; Parikshak et al. 2016; Labonté et al. 2017).

Each RNA sample was prepared from bilateral NAc of 1 animal, with 3 biological replicates per condition. RNA yield and quality were quantified using the Qubit and Agilent Bioanalyzer. RNA recovery averaged 3.37 μ g per whole cell RNA sample, 1.44 μ g per S3EQ cytosolic RNA sample, and 2.18 μ g per nuclear-enriched RNA sample. All RNA integrity numbers (RIN) were above 8.5. After RNA was collected from each cellular compartment, we performed (polyA+)RNA-seq or -S3EQ using the Illumina Hi-seq 4000, which yielded ~30 million reads per sample. To first assess the global transcriptomic differences among the three purification methods, we analyzed sample correlation using Spearman's rank-order method with a bin size of 50 kb. (PolyA+)RNA extracted from different cellular compartments were highly correlated with respect to total RNA; the correlation score ranges from 0.83 to 1 (Fig. 3A). Another clustering method, principal component analysis (PCA) was used to confirm this finding (Fig. 3B). PCA of RNA-seq and -S3EQ data centered around (0, 3.3) with the cumulative variability of first principal component around 0.95 and second principle component around 0.05, indicating only one factor primarily contributes to the differences observed between the three samples. Notably, two clusters were observed: one of the three RNA-S3EQ (cytosolic) biological replicates and one consisted of RNA-seq replicates from nuclear-enriched and whole cell lysates. This suggests nuclear-enriched and whole cell total transcriptome are more closely related than that of the cytosol alone.

To complement our findings on the correlation between cellular compartment transcriptomes, we next analyzed the differences between cytosolic, nuclear-enriched, and

whole cell (polyA+)RNA at individual gene resolution using Deseq from Bioconductor (Simon 2010) (Fig. 3C,D,E). We found that there were no differentially expressed genes (DEG) between nuclear-enriched fraction and whole cell (0 out of 37681 genes examined) and negligible differences (0.058%) between cytosol and whole cell (9 out of 37681 genes examined) (Fig. 3C,D). Only two of the nine genes that are differentially expressed between cytosol and whole cell compartments are protein coding genes, while seven are small RNAs that have no annotated function (Supplementary Table 1). Interestingly, when comparing nuclear-enriched to cytosolic (polyA+)RNA (Fig. 3E), we identified 79 DEG (out of 12,894 genes examined) ($P < 0.05$, fold change > 2) (Supplementary Table 2) that were annotated by gene ontology (GO) analysis to function in focal adhesion, activities of protein kinases, glycoproteins and nuclei (Fig. 3E).

3.4 Splicing complexity of cytosolic RNA more closely resembles that of whole cell RNA than nuclear-enriched RNA

We next examined the splicing profiles of nuclear-enriched, cytosolic, and whole cell (polyA+)RNA using rMATS (v3.2.5) (Shen et al. 2014), which detects and analyzes alternative splicing events grouped by major alternative splicing types (Fig. 4A, FDR < 0.1). Interestingly, both cytosolic and nuclear-enriched (polyA+)RNA have considerable amounts of differential alternative splicing events relative to whole cell (polyA+)RNA (FDR < 0.05), with skipped exon (SE) as the most common splicing type. To accurately identify differentially spliced isoforms in cytosolic and nuclear-enriched mRNA compared to whole cell, we applied a stringent filter on rMATS output (FDR < 0.05 , delta PSI > 0.1 , inclusion/exclusion events > 10). Similarly to previous reports, we chose to apply this filter to both identify bona fide alternative splicing events, and avoid splicing bias caused by expression levels and low-depth reads. (Parikshak et al. 2016). We found 48 alternative splicing events comparing cytosolic to whole cell, and 36 splicing events comparing nuclear-enriched to whole cell RNA (Fig. 4B,C, Supplementary Table 3). Of note, these alternative splicing events represent only 1-3% of the total splicing profile (Fig.4C). Functional analysis on differentially spliced isoforms revealed genes that function in ATP binding and kinase activity (Fig. 4D).

To further understand the source of the differential isoforms, we annotated all mapped (polyA+)RNA reads from nuclear-enriched, cytosolic and whole cell fractions. This helps to elucidate if transcription of specific genomic regions is causing the difference observed in splicing profile. Notably, we observed a highly similar genomic region enrichment profile (identified by relative abundance to a specific genomic feature) only between cytosolic and whole cell (polyA+)RNA, suggesting that the observed differences in isoform complexity are due to isoforms excluded from the nuclear-enriched (polyA+)RNA (Fig. 4E). In fact, overall exon enrichment decreased, while increases in intergenic and microRNA enrichment was observed in nuclear-enriched (polyA+)RNA relative to cytosolic and whole cell (Fig. 4F).

3.5 H3K36me3 ChIP-S3EQ enrichment more closely correlates to gene expression than ChIP-seq

We next examined H3K36me3 signal distribution on both expressed genes (Fig. 5A) and identified alternative splicing regions (spliced exon and 500bps up- down-stream of spliced exon) (Fig. 5D). To determine the concordance of H3K36me3 coverage between whole cell and S3EQ methods (Fig. 5B), we selected a representative gene, lymphocyte antigen 6 complex (*Sca-1*), that was the most differentially expressed between cytosol and nucleus (FC=14.54, $P < 5.27 \times 10^{-6}$). This gene (*Sca-1*), codes for is a T-cell activating protein that is known to be expressed in neurons (Gennarino et al. 2018). Consistent with our previous observations (Fig. 2C,D,E), H3K36me3 was enriched at the gene body, particularly at the exons, identified by both S3EQ and ChIP-seq.

To further assess how global ChIP-seq signal from each method correlated to RNA expression, we plotted H3K36me3 ChIP-seq signals on the top and bottom 5% of expressed transcripts from corresponding RNA-S3EQ (cytosolic) and RNA-seq (whole cell and nuclear-enriched) data (Fig. 5C). We found that for the top 5% of expressed transcripts, both RNA-S3EQ and RNA-seq methods showed a similar H3K36me3 density distribution, while for the bottom 5%, the methods differed in H3K36me3 enrichment profile. (Fig. 5C). Specifically, H3K36me3 showed greater enrichment at exon start (ES) and exon end (EE) sites in the bottom 5% of expressed transcripts from RNA-S3EQ compared to RNA-seq (Fig.5C).

3.6 S3EQ accurately maps H3K36me3 to alternatively and differentially spliced transcripts

We next examined the H3K36me3 signal distribution on alternatively spliced transcripts, specifically skipped exons, which was the most common type in our datasets (Fig. 6A). Across the genome, both ChIP-seq and CMIP-S3EQ methods revealed H3K36me3 enrichment at skipped exons with depletion at skipped exon start (SES) and end (SEE) regions (~150bp up- and down-stream of SES and SEE) (Fig. 6B). Next, we examined H3K36me3 enrichment at Transformer-2 protein homolog alpha (*Tra2a*) an alternatively spliced gene that differed in expression between cytosolic and nuclear-enriched transcriptome, but not between cytosolic and whole cell (Fig. 6C). We found that H3K36me3 was similarly enriched at *Tra2a*, identified by both CMIP-S3EQ and ChIP-seq approaches. Finally, we plotted H3K36me3 density on differentially expressed skipped exons identified by rMATS comparing both cytosolic and nuclear-enriched fractions to whole cell (Fig. 6D,E). In concordance with our previous finding (Hu et al. 2017), H3K36me3 enrichment was greatest around differentially spliced exon start and end sites (Fig.6D), further supporting the notion that H3K36me3 is strongly associated with alternative splicing in tissue from reward-associated brain regions. In alternatively spliced genes, cytosolic CMIP-S3EQ showed an overall higher density than whole cell ChIP-seq in the alternatively spliced exons. There was no difference in overall signal distribution, indicating that CMIP-S3EQ performed similarly to ChIP-seq in detecting hPTM abundance at alternatively spliced exons (Fig.6E). Together, our data show that within-sample, ChIP-S3EQ has higher sensitivity to RNA-S3EQ than separate ChIP-seq to RNA-seq when analyzing both differential expression and differential alternative splicing.

4 DISCUSSION

We described a novel protocol, S3EQ, for identifying and analyzing both transcription and hPTM enrichment within the same sample. S3EQ is a simple yet effective improvement to whole cell ChIP and RNA-seq and allowed us to efficiently obtain both ChIP- and RNA-seq data from the same sample. Using this approach, we found that the RNA-S3EQ, cytosolic transcriptome better resembles whole cell RNA transcriptome than nuclear-enriched alone. Moreover, S3EQ data showed a stronger signal to noise distribution of hPTM enrichment at lowly expressed transcripts. Overall, S3EQ measurement of global epigenome and transcriptome with high precision using fewer animals.

A major distinction between RNA-S3EQ and published methods is that it recovers only cytosolic, but not nuclear, RNA, as the latter will be heavily degraded by formaldehyde (David et al. 2011). This approach is considerably distinct from both whole-cell RNA extraction (Wang et al. 2009; J Feng et al. 2014; Nott et al. 2016; Labonté et al. 2017) and nuclear RNA-seq from FACS sorted neuronal nuclei (Krishnaswami et al. 2016; von Schimmelmänn et al. 2016; Jiang et al. 2015). This latter approach was recently applied to examine hPTMs in mouse cerebral cortex, as well as human prefrontal and anterior cingulate cortex, using the neuronal nuclei marker, NeuN (Jiang et al. 2015). In this case, NeuN+ (neuronal) and NeuN- (non-neuronal) nuclei were pelleted post-FACS for native ChIP- and RNA-seq to measure the extent of histone H3 lysine 9 trimethylation (H3K9me3), histone H3 lysine 27 acetylation (H3K27ac), as well as enrichment of chromatin readers (Jiang et al. 2015), underscoring the potential utility of both sorted neuronal populations and native ChIP for quantification of hPTM enrichment. Analysis of epigenetic gene expression during development has also been accomplished using FACS. For example, one study used FACS to isolate cortical cells from embryonic day 15.5 that were subsequently analyzed by ChIP-seq of various forms of phosphorylated and total RNA polymerase II (PolII) as well as H3K4me3 and H3K27me3 (Liu et al. 2017), to gain insights into the developmental regulation of RNA elongation. In particular, in the case that abundant neuronal nuclei from the individual animal can be sorted, 4-6 million nuclei are required for ChIP-seq (Zhao et al. 2014). Thus, combining cell sorting with S3EQ is a potentially powerful way to correlate changes in hPTM and expression yet is currently limited to profiling of nuclear RNA.

Despite the large number of studies correlating nuclear-enriched RNA expression to changes in chromatin modifications, the difference between RNA extracted from different cellular compartments has not been systematically analyzed for brain tissue (González-Porta et al. 2013). We profiled RNA expression in different cellular compartments, providing information on another widely used approach, RNA-sequencing of NeuN+ nuclei isolated by FACS. The greatest difference we observed between cellular compartment transcriptomes were 98 DEG (0.76% of genes measured) that differed between nucleus and cytosol. The ontological function of these genes in focal adhesion, activities of protein kinases, glycoproteins and nucleus suggested that this differential expression relates mainly to cellular structure, consistent with the expected differences in RNA profile of each cellular compartment.

While differential gene expression between compartments was negligible, our analysis of RNA splicing complexity between compartments revealed that cytosolic splicing complexity more closely resembles that of whole cell than nucleus. In fact, exon enrichment was decreased, while intergenic and microRNA enrichment was increased in nuclear-enriched RNA relative to cytosolic and whole cell. Consistent with a previous study on human tissue (Zaghlool et al. 2013), our findings suggest that cytosolic RNA serves as a better representative of the whole cell splicing transcriptome when applying the PolyA+ sequencing method. We reason that since splicing analysis by rMATS relies on mature spliced transcripts to compare inclusion/exclusion levels (Shen et al. 2014), cytosolic RNA provides the same sensitivity as whole cell measures in splice junction detection, while RNA isolated from nuclei alone may lack some mature spliced transcripts.

With respect to obtaining a chromatin modification profile, many published studies use fixed, bulk neuronal tissue, sonicated for chromatin fragments (Milne et al. 2009; Ferguson et al. 2015; Heller et al. 2014; Maze et al. 2011; Labonté et al. 2017). In contrast, using ChIP-S3EQ, we extracted and fixed nuclei from whole cell first, then sheared nuclear-enriched lysates for chromatin extraction. We generated H3K36me3 ChIP-seq and -S3EQ enrichment profiles in adult mouse NAc and found that H3K36me3 ChIP-S3EQ (nuclear-enriched) enrichment profiles were identical to those obtained by traditional (whole-cell) ChIP method.

Given the within-sample nature of the S3EQ approach, we hypothesized greater precision in correlating expression and hPTM enrichment than that of independent samples. Indeed, we found that while for the top 5% of expressed transcripts, both S3EQ and classic methods showed a similar H3K36me3 density distribution, the methods differed in H3K36me3 distribution for the bottom 5% of expressed transcripts. Specifically, H3K36me3 was more enriched at exon start and end sites in the bottom 5% of expressed transcripts prepared by S3EQ compared to whole cell and nuclear-enriched RNA-seq. This finding suggests that the within-sample S3EQ approach allows a more sensitive correlation between ChIP- and RNA-seq signal than ChIP- and RNA-seq generated from different samples, especially on lowly expressed transcripts.

In recent years, the emergence of single-cell RNA sequencing has pushed the boundary of new discovery by identifying cell subtypes and their corresponding gene expression profiles (Rompala et al. 2016; Chevee et al. 2018; Bakken et al. 2017; Hrvatin et al. 2018). However, much of this work is still limited to neural stem cells (NSC) and their immediate progeny (Shin et al. 2015; Schwartzenruber et al. 2018). Specifically, in NSCs, when using FACS-based approaches, astrocytes cannot be separated from other sub-neuronal cell types (Dulken et al. 2017; Shin et al. 2015), which creates bias against the isolation of the cell population (Artegiani et al. 2017). Another challenge in single-cell RNA sequencing is low transcript-capture efficiency (Lacar et al. 2016) and shallow depth of sequencing, which can resolve only a fraction of transcripts within each cell (Hrvatin et al. 2018). Therefore, this requires that large numbers of cells be analyzed in differentiating cell types alone. Thus, single cell sequencing approaches are not yet amenable to studies with low sample number, as is typical of human or complex animal behavior studies. Indeed, state of the art techniques for hPTM analysis of single cells relies on global analysis by innovative mass spectrometry methods,

rather than next generation sequencing (Cheung et al. 2018). As for future studies, developing a method that leverages both cell sorting and S3EQ would be valuable to overcome these limitations in sample variance.

In conclusion, we described a within-sample ChIP- and RNA-seq method (S3EQ) that showed increased sensitivity for correlating chromatin modifications with gene expression, especially for lowly expressed transcripts. ChIP-S3EQ of neuronal nuclei reliably identified hPTM enrichment in the adult mouse NAc with high precision and revealed that the cytosolic transcriptome more closely resembles that of the whole cell than nuclear-enriched transcriptome alone. The S3EQ method can be applied to improve the quality of bulk tissue transcriptomic and epigenome studies of neuronal tissue.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

The authors thank Sonia Lombroso and Emilia Gocke for assistance in animal husbandry, as well as Katherine Webb for critical reading of the manuscript. We are greatly appreciative of Perelman School of Medicine NGSC members for their contributions. Financial support from Charles E Kaufman Foundation Young Investigator Award (EAH), Whitehall Foundation Grant (EAH), and NIH-NIDA Director's Pioneer Award - Avenir (EAH, DP1 DA044250) is gratefully acknowledged.

REFERENCES

- ApREA J et al., 2013 Transcriptome sequencing during mouse brain development identifies long non-coding RNAs functionally involved in neurogenic commitment. *EMBO Journal*, 32(24), pp.3145–3160. Available at: <http://dx.doi.org/10.1038/emboj.2013.245>. [PubMed: 24240175]
- Artegiani B et al., 2017 A Single-Cell RNA Sequencing Study Reveals Cellular and Molecular Dynamics of the Hippocampal Neurogenic Niche. *Cell Reports*, 21(11), pp.3271–3284. Available at: <https://doi.org/10.1016/j.celrep.2017.11.050>. [PubMed: 29241552]
- Bagot RC et al., 2016 Circuit-wide Transcriptional Profiling Reveals Brain Region-Specific Gene Networks Regulating Depression Susceptibility. *Neuron*, 90(5), pp. 969–983. [PubMed: 27181059]
- Bakken T et al., 2017 Cell type discovery and representation in the era of high-content single cell phenotyping. *BMC Bioinformatics*, 18(Suppl 17).
- Beyeler A et al., 2017 Amygdala During Memory Retrieval. , 90(2), pp.348–361.
- Cahill ME et al., 2016 Bidirectional Synaptic Structural Plasticity after Chronic Cocaine Administration Occurs through Rap1 Small GTPase Signaling. *Neuron*, 89(3).
- Cai X et al., 2014 Single-Cell, Genome-wide Sequencing Identifies Clonal Somatic Copy-Number Variation in the Human Brain. , 8(5), pp. 1280–1289.
- Calipari ES et al., 2016 In vivo imaging identifies temporal signature of D1 and D2 medium spiny neurons in cocaine reward. *Proceedings of the National Academy of Sciences*, 113(10), pp.2726–2731. Available at: <http://www.pnas.org/lookup/doi/10.1073/pnas.1521238113>.
- Cheung P et al., 2018 Single-Cell Chromatin Modification Profiling Reveals Increased Epigenetic Variations with Aging. *Cell*, 0(0), p. 1385–1397.e14. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0092867418304513>.
- Cheve M et al., 2018 Variation in Activity State, Axonal Projection, and Position Define the Transcriptional Identity of Individual Neocortical Projection Neurons. *Cell Reports*, 22(2), pp. 441–455. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S2211124717318703>. [PubMed: 29320739]

- Damez-Werno DM et al., 2016 Histone arginine methylation in cocaine action in the nucleus accumbens. *Proceedings of the National Academy of Sciences*, 113(34), pp.9623–9628. Available at: <http://www.pnas.org/lookup/doi/10.1073/pnas.1605045113>.
- David LE et al., 2011 The effect of formaldehyde fixation on RNA: Optimization of formaldehyde adduct removal. *Journal of Molecular Diagnostics*, 13(3), pp.282–288. Available at: <http://dx.doi.org/10.1016/j.jmoldx.2011.01.010>. [PubMed: 21497290]
- Day J & Carelli RM, 2007 The Nucleus Accumbens and Pavlovian Learning. *Neuroscientist*, 13(2), pp. 148–159. [PubMed: 17404375]
- Day JJ et al., 2013 DNA methylation regulates associative reward learning. *Nature neuroscience*, 16(10), pp. 1445–1452. [PubMed: 23974711]
- Dobin A et al., 2013 STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1), pp. 15–21. [PubMed: 23104886]
- Dulken BW et al., 2017 Single-Cell Transcriptomic Analysis Defines Heterogeneity and Transcriptional Dynamics in the Adult Neural Stem Cell Lineage. *Cell Reports*, 18(3), pp.777–790. [PubMed: 28099854]
- Feng J et al., 2014 Chronic cocaine-regulated epigenomic changes in mouse nucleus accumbens. *Genome biology*, 15(4), p.R65. [PubMed: 24758366]
- Feng J et al., 2014 Chronic cocaine-regulated epigenomic changes in mouse nucleus accumbens. *Genome Biology*, 15(4).
- Ferguson D et al., 2015 SIRT1-FOXO3a Regulate Cocaine Actions in the Nucleus Accumbens. *Journal of Neuroscience*, 35(7), pp.3100–3111. Available at: <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.4012-14.2015>. [PubMed: 25698746]
- Finegersh A & Homanics GE, 2016 Chromatin immunoprecipitation and gene expression analysis of neuronal subtypes after fluorescence activated cell sorting. *Journal of neuroscience methods*, 263, pp.81–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/26868730> [Accessed August 24, 2016], [PubMed: 26868730]
- Finegersh A & Homanics GE, 2015 Chromatin immunoprecipitation and gene expression analysis of neuronal subtypes after fluorescence activated cell sorting. , 344(6188), pp. 1173–1178.
- Gennarino VA et al., 2018 Pumilio1 Haploinsufficiency Leads to SCA1-like Neurodegeneration by Increasing Wild-Type Ataxin1 Levels. *Cell*, 160(6), pp. 1087–1098. Available at: <http://dx.doi.org/10.1016/j.cell.2015.02.012>.
- González-Porta M et al., 2013 Transcriptome analysis of human tissues and cell lines reveals one dominant transcript per gene. *Genome Biology*, 14(7), p.R70 Available at: <http://genomebiology.com/2013/14/7/R70>. [PubMed: 23815980]
- Guo JU et al., 2014 Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nature Neuroscience*, 17(2), pp.215–222. Available at: <http://dx.doi.org/10.1038/nn.3607>. [PubMed: 24362762]
- Hawk JD et al., 2012 NR4A nuclear receptors support memory enhancement by histone deacetylase inhibitors. *Journal of Clinical Investigation*, 122(10), pp.3593–3602. [PubMed: 22996661]
- Heinz S et al., 2010 Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular Cell*, 38(4), pp.576–589. [PubMed: 20513432]
- Heller EA et al., 2014 Locus-Specific Epigenetic Remodeling Controls Addiction- and Depression-Related Behaviors. *Nature Neuroscience*, 17(12), pp.1720–1727. [PubMed: 25347353]
- Hrvatin S et al., 2018 Single-cell analysis of experience-dependent transcriptomic states in the mouse visual cortex. *Nature Neuroscience*, 21(1), pp.120–129. Available at: <http://dx.doi.org/10.1038/s41593-017-0029-5>. [PubMed: 29230054]
- Hu Q et al., 2017 Histone posttranslational modifications predict specific alternative exon subtypes in mammalian brain. *PLoS Computational Biology*, 13(6), pp.1–18.
- Jiang Y et al., 2015 Kmtle regulates a large neuron-specific topological chromatin domain. *Nature Genetics*, 344(6188), pp. 1173–1178.
- Srinivasan Karpagam et al., 2016 Untangling the brain's neuroinflammatory and neurodegenerative transcriptional responses. *Nature Communications*, 48(7), pp. 829–834.

- Krishnaswami SR et al., 2016 Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nature Protocols*, 11(3), pp.499–524. [PubMed: 26890679]
- Labonté B et al., 2017 Sex-specific transcriptional signatures in human depression. *Nature Medicine*, 23(9), pp. 1102–1111.
- Lacar B et al., 2016 Nuclear RNA-seq of single neurons reveals molecular signatures of activation. *Nature Communications*.
- Liu J et al., 2017 Dynamics of RNA Polymerase II Pausing and Bivalent Histone H3 Methylation during Neuronal Differentiation in Brain Development. *Cell Reports*, 20(6), pp. 1307–1318. Available at: <http://dx.doi.org/10.1016/j.celrep.2017.07.046>. [PubMed: 28793256]
- Lobo MK et al., 2006 FACS-array profiling of striatal projection neuron subtypes in juvenile and adult mouse brains. *Nature neuroscience*, 9(3), pp.443–452. [PubMed: 16491081]
- Maze I et al., 2011 Cocaine dynamically regulates heterochromatin and repetitive element unsilencing in nucleus accumbens. *Proceedings of the National Academy of Sciences*, 108(7), pp.3035–3040. Available at: <http://www.pnas.org/cgi/doi/10.1073/pnas.1015483108>.
- Mews P et al., 2017 Acetyl-CoA synthetase regulates histone acetylation and hippocampal memory. *Nature*, 546(7658), pp.381–386. [PubMed: 28562591]
- Milne TA, Zhao K & Hess JL, 2009 Chromatin Immunoprecipitation (ChIP) for Analysis of Histone Modifications and Chromatin-Associated Proteins. , 538, pp.409–423. Available at: <http://link.springer.com/10.1007/978-1-59745-418-6>.
- Nott A et al., 2016 Histone deacetylase 3 associates with MeCP2 to regulate FOXO and social behavior. *Nature Neuroscience*, 19(July).
- Parikshak NN et al., 2016 Genome-wide changes in lncRNA, splicing, and regional gene expression patterns in autism. *Nature*
- Peña CJ et al., 2017 Early life stress confers lifelong stress susceptibility in mice via ventral tegmental area OTX2. *Science*, 1(June), pp. 1185–1188. Available at: <http://science.sciencemag.org/content/sci/356/6343/1185.full.pdf>.
- Rio DC et al., 2010 Preparation of cytoplasmic and nuclear RNA from tissue culture cells. *Cold Spring Harbor Protocols*, 5(6), pp.1–5.
- Robinson JT et al., 2011 Integrative Genome Viewer. *Nature Biotechnology*, 29(1), pp.24–6.
- Rompala GR, Finegersh A & Homanics GE, 2016 Paternal preconception ethanol exposure blunts hypothalamic-pituitary-adrenal axis responsivity and stress-induced excessive fluid intake in male mice.
- Ross-Innes CS et al., 2012 Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature*, 481(7381), pp.389–393. Available at: <http://dx.doi.org/10.1038/nature10730>. [PubMed: 22217937]
- Saddoris MP et al., 2015 Differential Dopamine Release Dynamics in the Nucleus Accumbens Core and Shell Reveal Complementary Signals for Error Prediction and Incentive Motivation. *Journal of Neuroscience*, 35(33), pp.11572–11582. Available at: <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.2344-15.2015>. [PubMed: 26290234]
- Sanders SJ et al., 2017 Whole genome sequencing in psychiatric disorders: The WGSPD consortium. *Nature Neuroscience*, 20(12), pp. 1661–1668. Available at: <http://dx.doi.org/10.1038/S41593-017-0017-9>. [PubMed: 29184211]
- Sati S et al., 2012 Genome-wide analysis reveals distinct patterns of epigenetic features in long non-coding RNA loci. *Nucleic Acids Research*, 40(20), pp.10018–10031. [PubMed: 22923516]
- von Schimmelmann M et al., 2016 Polycomb repressive complex 2 (PRC2) silences genes responsible for neurodegeneration. *Nature Neuroscience*, 2(August), pp.1–14.
- Schwartzentruber J et al., 2018 Molecular and functional variation in iPSC-derived sensory neurons. *Nature Genetics*, 50(1), pp.54–61. Available at: <http://dx.doi.org/10.1038/S41588-017-0005-8>. [PubMed: 29229984]
- Shen S et al., 2014 rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proceedings of the National Academy of Sciences*, 111(51), pp.E5593–E5601. Available at: <http://www.pnas.org/lookup/doi/10.1073/pnas.1419161111>.

- Shin J et al., 2015 Single-Cell RNA-Seq with Waterfall Reveals Molecular Cascades underlying Adult Neurogenesis. *Cell Stem Cell*, 17(3), pp.360–372. Available at: <http://dx.doi.org/10.1016/j.stem.2015.07.013>. [PubMed: 26299571]
- Simon W, 2010 Differential expression analysis for sequence count data. *Genome Biology*, 11(10). Available at: <http://genomebiology.com/2010/11/10/R106>.
- Stefanelli G et al., 2018 Learning and Age-Related Changes in Genome-wide H2A.Z Binding in the Mouse Hippocampus. *Cell reports*, 22(5), pp. 1124–1131. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/29386101>. [PubMed: 29386101]
- Taniguchi M, Carreira MB, Cooper YA, Bobadilla AC, Heinsbroek JA, Koike N, Larson EB, Balmuth EA, Hughes BW, Penrod RD, Kumar J, Smith LN, Guzman D, Takahashi JS, Kim TK, Kalivas PW, Self DW, Lin Y CC, 2017 HDAC5 and Its Target Gene, Npas4, Function in the Nucleus Accumbens to Regulate Cocaine-Conditioned Behaviors. *Neuron*, 96, pp. 130–144. [PubMed: 28957664]
- Taniguchi M et al., 2012 Histone Deacetylase 5 Limits Cocaine Reward through cAMP-Induced Nuclear Import. *Neuron*, 73(1), pp. 108–120. [PubMed: 22243750]
- Thorvaldsdottir H, Robinson JT & Mesirov JP, 2013 Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, 14(2), pp. 178–192. [PubMed: 22517427]
- Wang Z, Gerstein M & Snyder M, 2009 RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10, p.57 Available at: <http://dx.doi.org/10.1038/nrg2484>.
- Xu S et al., 2014 Spatial Clustering for Identification of ChIP-Enriched Regions (SICER) to Map Regions of Histone Methylation Patterns in Embryonic Stem Cells. *Methods in molecular biology* (Clifton, N.J.), 1150, pp.97–111. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4152844/>.
- Zaghlool A et al., 2013 Efficient cellular fractionation improves RNA sequencing analysis of mature and nascent transcripts from human tissues. *BMC Biotechnology*, 13. [PubMed: 23413993]
- Zhao W et al., 2014 Comparison of RNA-Seq by poly (A) capture, ribosomal RNA depletion, and DNA microarray for expression profiling. *BMC Genomics*, 15(1), pp. 1–11. [PubMed: 24382143]
- Zovkic IB et al., 2014 Histone H2A.Z subunit exchange controls consolidation of recent and remote memory. *Nature*, 515(7528), pp.582–586. Available at: <http://dx.doi.org/10.1038/nature13707>. [PubMed: 25219850]

Figure 1. Development of single sample sequencing (S3EQ) for within-sample epigenomic and transcriptomic profiling.

Schematic representation of the difference between S3EQ method and whole cell ChIP- and RNA- seq approaches. S3EQ method is used to generate cytosolic RNA-seq and nuclear ChIP-seq from the same NAc sample, while classic methods use two sets of samples to generate whole cell ChIP- and whole cell RNA-seq respectively.

Figure 2. ChIP-S3EQ performs in similar accuracy comparing to whole-cell ChIP-Seq
 (A) DNA fragment distributions from ChIP-S3EQ and ChIP approaches measured by Agilent Bioanalyzer. DNA fragment distributions showed no difference between the two methods.
 (B) Heatmap of H3K36me3 signal generated from nuclear S3EQ and whole-cell ChIP-seq was highly correlated using Spearman's rank-order method based on the number of mapped reads using a bin size of 50kb. (C) Heatmap of H3K36me3 enrichment on annotated gene body regions, 5kb up- and down-stream from the transcription start and end sites (TSS/TES). (D) Density plot of H3K36me3 reads per million using ChIP-seq (whole cell) and ChIP-S3EQ (nuclear). (E) IGV browser view of ChIP-seq and ChIP-S3EQ at gene ActinB region. (F) H3K36me3 enrichment in different genomic regions.

Figure 3. RNA-seq from different cellular compartments exhibits highly similar expression profile

(A) Heatmap of mapped reads identified by RNA-seq from cytosolic, nuclear fraction and whole cell using Spearman's rank-order method based on the number of mapped reads using a bin size of 50kb. (B) PCA plot of mapped reads identified by RNA-seq from cytosolic, nuclear fraction and whole cell. (C) Volcano plot of differential expression gene analysis between cytosolic and whole cell RNA. Each dot represents a single gene comparison. Statistically significant ($p < 0.05$) expression changes are marked in red, while fold changes > 5 are represented in orange. (D) Volcano plot of differential expression gene analysis between

nuclear and whole cell RNA. Each dot represents a single gene comparison. Statistically significant ($p < .05$) expression changes are marked in red, while fold changes > 5 are represented in orange. (E) Volcano plot of differential expression gene analysis between nuclear and cytosolic RNA. Each dot represents a single gene comparison. Statistically significant ($p < .05$) expression changes are marked in red, while genes significantly changed ($P < 0.05$) which also have a fold change > 5 are represented in green. (F) Go-term analysis on differential expression genes ($p < 0.05$) between cytosolic and nuclear RNA.

Figure 4. Cytosolic RNA resembles total RNA splicing profiles

(A) Schematic representation of different splicing events. Colored block indicates spliced exon. (B) Number of alternative splicing events identified in cytosolic RNA comparing to whole cell RNA and nuclear RNA compared to whole cell RNA (FDR < 0.05, $\Delta\text{PSI} > 0.1$). (C) Percentage of alternative splicing events (FDR < 0.05, $\Delta\text{PSI} > 0.1$) in total identified splicing events in comparisons between cytosolic RNA to total RNA and nuclear RNA to total RNA. (D) Function analysis of alternative splicing events of cytosolic and nuclear RNA compared to total RNA. Enrichment score calculated by DAVID function cluster analysis. (E) Genomic annotation of intron, exon and intergenic regions by Homer on all mapped reads from RNA-seq in whole cell, cytosolic and nuclear RNA (FDR < 0.05). (F)

Genomic annotation of micro and non-coding RNA regions by Homer on all mapped reads from RNA-seq in whole cell, cytosolic and nuclear RNA (FDR < 0.05).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 5. S3EQ H3K36me3 signal shows a stronger association in low-expressed RNA-S3EQ than whole cell ChIP-to RNA-seq

(A) Schematic representation of ChIP-seq reads in differential expression regions with 500bp flanking regions identified by RNA-seq. (B) H3K36me3 ChIP-seq enrichment levels of classic whole cell (red), S3EQ nuclear (blue) on highest differential expression gene (Sca-1) between cytosolic (dark green), nuclear (light blue) and total RNA (green) in RNA-seq. (C) Density plot of H3K36me3 ChIP-seq reads on whole cell RNA mapped reads of top 5% (red) and bottom 5% (dark red) in expression levels (left). Density plot of H3K36me3 ChIP-seq reads on nuclear RNA mapped reads of top 5% (green) and bottom 5% (dark green) in expression levels (middle). Density plot of H3K36me3 ChIP-S3EQ reads on RNA-S3EQ mapped reads of top 5% (blue) and bottom 5% (dark blue) in expression levels (right).

Figure 6. S3EQ accurately maps H3K36me3 to alternatively and differentially spliced transcripts
 (A) Schematic representation of ChIP-seq reads in alternatively spliced regions with 500bp flanking regions of skipped exons identified by rMATS(v3.4.2). (B) Density plot of H3K36me3 classic whole cell (red) and S3EQ nuclear (blue) ChIP-seq reads on all annotated skipped exons with 500bp up-and down-stream splicing junctions. (C) H3K36me3 ChIP-seq signals of whole cell (red), S3EQ nuclear (blue) on most frequently alternatively spliced gene (Tra_2a) between nuclear (light blue) and total RNA (green) in RNA-seq. Cytosolic (dark green) RNA-seq is also shown as control. (D) H3K36me3 ChIP-seq (blue) - S3EQ (red) signals of on alternatively spliced skipped exons (as in Fig 4B, left bar) between

nuclear and whole cell RNA. (E) H3K36me3 ChIP-seq (blue) -S3EQ (red) signals of alternatively spliced skipped exons (as in Fig 4B, right bar) between cytosolic and whole cell RNA.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript