HHS Public Access

Author manuscript

Mol Neurobiol. Author manuscript; available in PMC 2019 July 01.

Published in final edited form as:

Mol Neurobiol. 2018 July ; 55(7): 5611–5622. doi:10.1007/s12035-017-0781-2.

^{AZI2}3′UTR is a new *SLC6A3* downregulator associated with an epistatic protection against substance use disorders

Kefu Liu[1,2,#], **Jinlong Yu**[1,#], **Juan Zhao**[1,2,#], **Yanhong Zhou**[1,#], **Nian Xiong**[1,3], **Jie Xu**[4], **Tao Wang**[3], **Richard L. Bell**[5], **Hong Qing**[2], and **Zhicheng Lin**[1,*]

[1]Laboratory of Psychiatric Neurogenomics, Division of Basic Neuroscience, Mailman Neuroscience Research Center, McLean Hospital, Belmont, and Department of Psychiatry, Harvard Medical School, Massachusetts 02478, United States of America

[2]School of Life Science, Beijing Institute of Technology, 100081 Beijing, China

[3]Department of Neurology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, 430074 Hubei, China

[4]Department of Computer Information Systems, Bentley University, Waltham, Massachusetts 02452, United States of America

[5]Department of Psychiatry, Institute of Psychiatric Research, Indiana University School of Medicine, Indianapolis, Indiana 46202, United States of America

Abstract

Regulated activity of SLC6A3, which encodes the human dopamine transporter (DAT), contributes to diseases such as substance abuse disorders (SUDs); however, the exact transcription mechanism remains poorly understood. Here we used a common genetic variant of the gene, intron 1 DNP1B sequence, as bait to screen and cloned a new transcription active unit, ^{AZI2}3′UTR, for SLC6A3. ^{AZI2}3′UTR is a 3′ untranslated region (3′UTR) of the human 5-Azacytidine Induced 2 gene (AZI2) but appeared to be transcribed independently of AZI2. Found to be present in both human cell nuclei and dopamine neurons, this RNA was shown to down regulate promoter activity though a variant-dependent mechanism *in vitro* Both reduced RNA density ratio of ^{AZI2}3′UTR/AZI2 and increased DAT mRNA levels were found in ethanol-naïve alcohol-preferring rats. Secondary analysis of dbGaP GWAS datasets (Genome-Wide Association Studies based on the database of Genotypes and Phenotypes) revealed significant interactions between regions upstream of ^{AZI2}3′UTR and SLC6A3 in SUDs. Jointly, our data suggest that ^{AZI2}3′UTR confers variant-dependent transcriptional regulation of SLC6A3, a potential risk factor for SUDs.

Keywords

allele-dependent; dopamine transporter; kangaroo layout of genes; lncRNA; ribonucleic repressor; transcription factor

[*]Corresponding: Z. Carl Lin, zhicheng_lin@hms.harvard.edu, Tel:+1 617-855-2175.
[#]Equal contributors

Financial Disclosures
The authors declare no competing financial interests.

Introduction

Regulated dopamine transmission contributes to normal brain function but its dysregulation may result in pathophysiology associated with a range of neuropsychiatric disorders[1–5]. The dopamine transporter (DAT) is a primary regulator of dopamine neurotransmission. Its functional alteration has been consistently implicated in both normal brain function and dysfunction, especially in addiction especially in substances of abuse[6–10]. As such, the transcriptional regulation of the human gene *SLC6A3*, encoding DAT, plays an important role in brain activity by determining where, when, and how much functional DAT is expressed in the brain. Because of this, many studies have investigated *SLC6A3's* association with various brain disorders, including functional analysis of the gene's variants[11–15].

According to the 1000 Genomes Project, *SLC6A3* has more than 2,500 polymorphisms within its 70 kb gene regions among various populations[16,17]. To better understand how these variations modulate the gene's activity, it is essential to elucidate the transcription mechanisms. Previous studies have identified several transcription factors (TFs) for *SLC6A3*, including NURR1, HEY1, SP1, and SP3[18–21]. In addition, PITX3 was found to regulate the mouse DAT gene[22]. None of these have been linked to variant-dependent regulation associated with any diseases except limited evidence for NURR1's association with Parkinson's disease[23,24]. No other type of RNA-based regulation has been reported for *SLC6A3*.

In this study, we started the allele-dependent TF search and as a result, cloned a 3′ untranslated region (3′UTR).

Materials and Methods

Yeast One-Hybrid (Y1H) screen

By using human substantia nigra total RNA and Matchmaker One-Hybrid Library Screening Kit (Takara Bio USA Inc., Mountain View, CA, USA, Cat# 636560 and 630304), we screened for transcription factors in the following six steps per manufacturer's instruction: (a) constructing and testing the bait which was prepared by annealing two complementary 40 base DNA oligonucleotides harboring dinucleotide polymorphism in Intron 1 allele B(DNPiB) in the middle and carrying *EcoR* I and *Mlu* I at each end for inserting to reporter pHIS2.1 (see Supplementary Table 1 for all nucleotides used in this study); (b) generation of the cDNA for the library including first strand cDNA synthesis, amplifying cDNA using Long Distance PCR (LD-PCR), and purifying ds cDNA with CHROMA SPIN TE-400 Columns; (c) library screening; (d) confirmation of positive interactions and rescue of the prey plasmid by re-streaking and yeast colony PCR to eliminate duplicates, rescuing and isolation of library plasmids responsible for activation of reporters; (e) yeast transformation followed by streaking single candidate colony on new 100mM 3-AT, TDO (Triple Dropout Medium: SD/–His/–Leu/–Trp) plate for selection; and (f) screening by PCR of selected colonies and DNA re-sequencing.

Plasmid construction

Mammalian expression plasmid—F30 was freed from pGAD-T7- F30 which was identified by Y1H screen and cloned to pcDNA3.1+ to result in pcDNA3.1+F30 plasmid by *Eco*R I and *Xho* I restriction sites. The pcDNA3.1+F30 plasmid was used for all overexpression experiments. *Luciferase (Luc) reporters* SV40-A (pGL3-SV40-A-luc) and SV40-B (pGL3-SV40-B-luc) were constructed by directionally inserting 40 bp dsDNA oligo containing DNPi allele A or B in the middle of the oligo, after self-annealing the oligos carrying two restriction site ends, into *Hind*III/*Nco*I sites of Promega's pGL3 SV40 Promoter vector (Madison, WI, USA; see Supplementary Table 1 for a list of all nucleotides used in this study). Preparation of 7.9kb constructs has been described before[25]. 7.9kb-A (pGL3-hDAT7.9kb-A) and 7.9kb-B (pGL3-hDAT7.9kb-B) differed only by DNPi alleles and both covered a SLC6A3 promoter region (7.9kb) from 5,931bp upstream of transcription start site to 35 bp downstream of DNPi. For promoter screen, pGL3E-0.1kb, pGL3E-0.3kb and pGL3E-1.6kb were constructed by using high fidelity PfuUltra II Fusion HotStart DNA Polymerase (Agilent Genomics, Santa Clara, CA, USA) according to the manufacturer's instruction, and the primers for three pairs: 0.1kbF/F30R, 0.3kbF/F30R and 1.6kbF/F30R. PCR products were digested with *Kpn* I/*Nhe* I before cloning into the same sites of Promega's pGL3 Enhancer vector. All PCR products in the clones were confirmed for high fidelity by DNA-resequencing.

Cell Culture

BE(2)-M17, SK-N-AS, SN4741, IMR-32, HEK293T or HB248 of human origins were cultured in Dulbecco's modified eagle medium (DMEM, ATCC, VA, USA) containing 10% Fetal Bovine Serum (FBS, Atlanta biologicals, GA, USA) in a 75 cm^2 Falcon flask (Thermo Fisher Scientific, Waltham, MA, USA) in humidified air with 5% CO_2 at 37°C. For nucleus and cytoplasm separation, cells were split in Falcon 6-well plate for a next day confluence level of 80% and harvested by cell scrapers (Corning, NY, USA). SN4741 cells of mouse midbrain origin were cultured at 32°C for maintenance and at 37°C for gene expression after transfection.

Cell culture transfection for Luc activity assay or RNA extraction

Cells were split in a 24-well plate for a next day confluence level of 80%. For each well, plasmid DNA (from 0 ng–400 ng) was mixed with 2μl SuperFect Reagent (Qiagen, CA, USA) in DMEM and incubated at RT for 10 min. For Luc assay, cells in each well were transfected with 400 ng of the SV40-A, SV40-B or 7.9k-A or 7.9k-B plasmid DNA. Plasmid DNA of expression vector pcDNA3.1+ or pcDNA3.1+F30 was co-transfected with the transfection efficiency control plasmid pRL-TK reporter (Promega), respectively. For dose-dependence transfection and RNA extraction, 0 ng–400 ng plasmid DNA was used. Culture medium was removed from each well and the plasmid DNA/SuperFect mixture added. All plasmid DNA was isolated from *E. coli* culture at least three independent times for each plasmid by using PureYield™ Plasmid Miniprep System (Promega) and dissolved in TE buffer to a DNA concentration of 400ng/μl. Different preparations, each with an $OD_{260/280}$ ratio of about 1.8, were mixed to avoid quality variation. Three hours later, fresh growth medium replaced the transfection mixture. For Luc assay, 48 hrs later, cells were collected

and lysed and the firefly and renilla luciferase (RL) activities were measured with Dual-Luciferase Reporter Assay System (Promega) according to the manufacturer's instructions, using a Synergy HT luminometer (BioTek, Winooski, VT, USA). All firefly Luc activity was normalized to RL activity for transfection efficiency. For RNA extraction, the incubation time was 24 hrs. Each experiment in triplicate wells was repeated by the indicated number in figure legends.

RNAi on SV40-A or SV40-B

Cells were plated out with medium without antibiotics in a 24 well plate for a second day 80–90% confluence level. Just before transfection, the plasmid DNA mixture (160 ng SV40-A or SV40-B) and 5 pmol of F30-siRNA1, 2, or 3 were prepared, mixed gently in 50 μl medium (without serum), mixed with 1μl Lipofectamine™ 2000 diluted in 50 μl medium (without serum), mixed gently and incubated at room temperature (RT) for 20 min. The DNA-RNAi molecule-Lipofectamine 2000 complexes were added to each well containing cells and medium, followed by gentle mixing. After 48 hrs of incubation, cells were harvested for Luc activity assay, as described above.

Quantification of mRNA levels by quantitative real-time polymerase chain reaction (qRT-PCR)

For *in vitro* work, cells of each well in 24 well plate were extracted for total RNA in TRIZOL (Invitrogen, CA, USA; 250 μl for each well, 500 μl after combining duplicate wells), according to the TRIZOL User Guide, and RNA was quantified by NanoDrop Lite (Thermo Fisher Scientific, Waltham, MA, USA). Final mass from each extraction was about 4 μg RNA in 20 μl DEPC-treated water. For each sample, 200 ng RNA was subjected to cDNA synthesis in a volume of 10 μl, by Verso cDNA synthesis kit (Thermo Fisher Scientific).

For *in vivo* work, rodent brain tissue was sliced, and the left and right midbrain regions were micro-punched, transferred to RNase-free 1.5 ml microfuge tubes with 100 μl TRIZOL reagent, homogenized in ice by mechanical trituration followed by the addition of another 900 μl of TRIZOL reagent to the tubes to extract total RNA for cDNA synthesis.

Each qRT-PCR reaction was prepared by KAPA SYBR FAST qPCR Master Mix (KAPA biosystem, MA, USA) with 200 nM of primer mixture and 0.1μl of cDNA. The PCR program ran 45 cycles, with an annealing temperature of 56°C, in Bio-Rad CFX Connect Real-Time System (Bio-Rad, CA, USA) according to the manufacturer's protocol. The efficiency (an average coefficient of 2.0) was calculated using a series dilution method and Bio-Rad CFX Manager software and each coefficient was used in fold-change calculations for each primer pair. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as an internal control for input.

Northern blot analysis

Probe preparation—A 442 bp human probe (hF30p) was freed from 2 μg of pcDNA3.1+F30 plasmid DNA by double-digestion with *Nhe* I and *Hind* III restriction enzymes (New England Biolabs, Ipswich, MA, USA). A 339 bp mouse probe (mF30p) was

PCR amplified by using 0.1 μl mouse cDNA as the template and high fidelity PfuUltra II Fusion HotStart DNA Polymerase. Using the Verso cDNA synthesis kit, the mouse cDNA was synthesized in 10 μl from 200 ng mouse RNA extracted from SN4741 cells. The two DNA fragments were purified by Wizard® SV Gel and PCR Clean-Up System (Promega) and 100 ng of each DNA fragment was used for labeling with a North2South Biotin Random Prime Labeling kit (Thermo Fisher Scientific). mF30p was also used for rat RNA blotting because it had 84% identity with the rat AZI2 3′UTR, which was the only homology found throughout the entire rat genome.

RNA preparation—Fresh mouse or rat brain tissue was dissected into different subregions, including midbrain, hippocampus and cortex which were stored in −80°C. Cells from 6-well plate or 100 mg of brain tissue were homogenized in 1 ml TRIZOL Reagent by Argos Pestle Mixer (Argos Technologies, IL, USA) for RNA extraction as described above. The final RNA concentration was about 2μg/μl in 50μl of DEPC-treated water.

Cytoplasmic and nuclear RNA fractions of cells harvested in 200 μl ice-cold Lysis Buffer, J. Swirling buffer, from 6-well plate were separated using a Cytoplasmic and Nuclear RNA Purification Kit (Norgen Biotek, ON, Canada). The resultant RNA concentration was approximately 500 ng/μl for the cytoplasmic fraction and 90 ng/μl for the nuclear fraction, each in 50 μl of TE buffer.

Blotting—RNA was resolved to molecular weight by electrophoresis on a 1% denaturing agarose gel. Two μg of cytoplasmic RNA or 0.5 μg of nuclear RNA was loaded and after electrophoresis, RNA in the gel was capillary-transferred to BrightStar-Plus positive-charged Nylon membrane (Thermo Fisher Scientific). The membrane-RNA was crosslinked by UV and RNA was labeled by adding 40 ng labeled probe and detected by North2South Chemiluminescent Hybridization and Detection Kit (Thermo Fisher Scientific). The final results were visualized by a Chemi Doc XRS Molecular Imager, and band densities were quantified using ImageLab software (Bio-rad).

Western blot analysis

Cells were scrape-harvested from a 6-well plate, lysed in 200 μl ice-cold Lysis Buffer J. Swirling buffer for 5 min, followed by centrifugation at 12,000 rpm at 4°C for 3 min. The supernatant was collected as cytoplasmic proteins. The pellet was washed with ice-cold PBS, added to 200 μl RIPA buffer and disrupted by sonication for the nuclear proteins. Protein concentration was measured by Bradford Protein Assay (Bio-Rad, Hercules, CA, USA). 20 μl sample was loaded on a 10% Criterion TGX Precast Midi Protein Gel (Bio-Rad). Proteins were then transferred to a PVDF membrane (Bio-Rad) and stained with antibodies for GAPDH (1:1000 dilution, Abcam, Cambridge, MA, USA), or a monoclonal antibody to RNA Pol II (1:1000 dilution, Santa Cruz Biotechnology Inc., CA, USA), followed by incubation with appropriate secondary antibodies and the staining was visualized with Pierce ECL substrate (Thermo Fisher Scientific). Images were captured as for the Northern blots.

RNAscope and immunofluorescence (IF) staining

RNAscope analysis was carried out by using the RNAscope 2.5 HD Reagent Kit – RED (Advance Cell Diagnosis, Newark, CA, USA). The sample was fixed, dehydrated, and pretreated. BE(2)-M17 cells cultured on coverslips or frozen human *post mortem* nigra brain tissue were fixed with 4% paraformaldehyde (PFA) for 30 min and rinsed 3 times with Tris-buffered saline (TBS). Fixed cells or sections of the fixed tissue were sequentially immersed in 50% ethanol (EtOH), 70% EtOH and then 100% EtOH for dehydration. Then 2–4 drops of RNAscope Protease IV were applied to pretreat the sample. At the probe incubation step, the sections were incubated with the desired probe (~2–3 drops/section or coverslip) at 40°C for 2 hrs and then washed four times in 1x wash buffer for 1 min each. The following probes were used: 5′ probe was a 10ZZ probe named Hs-AZI2-O2 targeting 546-1160 of NM_022461.4; 3′ probe was a 13ZZ probe named Hs-AZI2-O1 targeting 2384-3286 of NM_022461.4; scramble probe was RNAscope 3-plex Negative Control Probe. Amplification steps were performed by sequential incubation of Amp1 (pre-amplifier), Amp2 (background reducer), Amp3 (amplifier), Amp4 (label probe) at 40°C for 30 min each and then by incubation with Amp5 and Amp6 at RT for 30 min. The signal was visualized by mixing 1 part of FAST-A solution to 60 parts of FAST-B solution and then incubated at RT for 10 min.

IF staining was performed immediately after RNAscope. Samples were incubated in blocking buffer (Life Technology) at RT for 30 min and then incubated with rabbit polyclonal anti-tyrosine hydroxylase (anti-TH, Abcam, Cambridge, MA, USA, 2 mg/ml) primary antibody (1:500 dilution) at 4°C overnight. The following day, samples were washed four times in TBS for 4 min each, incubated with fluorescent secondary antibody (a 488 labeled goat anti-mouse secondary antibody) at 1:500 dilution with 0.05% Triton 100 at RT for 2 hrs. Sections/coverslips were washed and mounting buffer with DAPI (VECTOR LABORATORIES, Burlingame, CA, USA) was applied. Images were collected by a confocal microscope, Leica TCS SP8 (Leica Microsystems Inc. IL, USA).

RNA-guided Chromosome conformation capture (R3C) protocol[26]

Cells from a 75 cm^2 flask were treated with 2% formaldehyde at RT for 10 min and added to cold 1 M glycine for a final concentration of 0.125M on ice. After scraping and centrifuging, cell pellets were lysed in 1ml lysis buffer (10mM Tris, pH 8.0, 10mM NaCl, 0.2% NP-40, protease inhibitors and RNA inhibitors) on ice for 10 min. Pelleted by centrifugation at 12,000 rpm and 4°C for 10 min, SDS buffer was added to dissolve the pellet in order to detach non-crosslinked proteins from DNA, followed by biotin-modified reverse transcription for dsDNA. The first-strand cDNAs were synthesized by using NEBNext RNA First Strand Synthesis Module (New England Biolabs, Ipswich, MA) and second-strand cDNAs synthesized by using NEBNext mRNA Second Strand Synthesis Module (NEB). dsDNA was digested with 600 U *BspH* I (NEB) at 37°C overnight with shaking (900 rpm) and then ligated with T4 DNA ligase(NEB) at 16°C for 4 hrs. Next, samples were reverse crosslinked by proteinase K at 65°C overnight and digestion of remaining RNA by RNase A at 37°C for 30 min, followed by purification with phenol–chloroform method and streptavidin Dynabeads (Invitrogen). Purified DNA was used for PCR amplification using four pairs of primers: r3cF30F/r3cF30R, r3cF30F/r3cSLC6A3R, r3cSLC6A3F/r3cF30R,

and r3cSLC6A3F/r3cSLC6A3R. The PCR program ran 45 cycles with the annealing temperature of 56°C in Veriti 96 well thermo cycler (Applied Biosystems, CA, USA).

Secondary analysis of dbGaP GWAS genotype

Genetic analysis—Secondary analysis used three datasets, Collaborative Study on the Genetics of Alcoholism (COGA, phs000125.v1.p1), Study of Addiction: Genetics and Environment (SAGE, phs000092.v1.p1) and the Australian twin-family study of alcohol use disorder (OZALC, phs000181.v1.p1). COGA covered two different US populations, that is, European Americans and African Americans, so that after data quality control (QC) using a published protocol[27], the dataset was split into two datasets, European Americans and African Americans. From the OZALC sample, only the unrelated individuals were extracted for the purpose of epistasis analysis. After QC, a total of 6,596 subjects were kept as unrelated and 41% of them were cases. Imputation was carried for each of these four cleaned datasets as described before[28], in order to extend the association power. Data manipulation, allelic association, and meta-analysis were all performed by using PLINK[29]; logistic regression analysis of single nucleotide polymorphism (SNP)-SNP interactions used the function implemented in the CASSI 2.50 software[30].

Genetic result display used a visualization program to draw the graph. The visualization program is implemented using the Java Swing package. The program read two input files in text format. The first file contains six columns for gene information: Node Name, Start Marker, End Marker, Lower Bound, Upper Bound, and Direction. The start and end markers are the beginning and end coordinates of a gene. The second file contains relation information recorded in three columns: Coordinate 1, Coordinate 2, and Strength. The strength could be any statistics measuring the weight or confidence of relations.

Data analysis

For rat AZI2 mRNA (5′ primers) or 3′UTR (F30) mRNA qRT-PCR data analysis, GAPDH mRNA was used to normalize for input, followed by calculating a relative expression ratio between alcohol-nonpreferring, NP, and alcohol-preferring, P, rats. The 3′UTR ratio was further normalized by the AZI2 ratio to obtain an ^{AZI2}3′UTR enrichment ratio of 3′UTR/AZI2, in order to compare the ^{AZI2}3′UTR/AZI2 ratio between Northern blotting and qRT-PCR findings. One-way or two-way ANOVAs with post hoc Tukey tests or *Student's* t-tests were used in this study to assess statistical significance with the criteria of $P < 0.05$. All data are represented by mean ± S.E.M.

Results

F30 cloning

In a previous study, we found that a 40 bp Intron 1 DNA sequence (fragment III) displayed inhibitory activity in human neuroblastoma SK-N-AS cells[25]. That fragment was of haplotype A, carrying allele A of a dinucleotide polymorphism in Intron 1 (DNPi or rs67175440). This DNPi had two alleles, A (AG) and B (GA). Fragment III carried allele A and was not bound by any nuclear protein. Contrarily, the allele B fragment is bound by

nuclear protein (unpublished observations). Therefore we used this allele B (DNPiB) containing 40 bp fragment as bait to search for allele-dependent TFs targeting *SLC6A3*.

In the yeast one-hybrid (Y1H) screen of a cDNA library which was prepared with total RNA isolated from *post mortem* human nigral tissue, we picked 92 single colonies for PCR amplification and the PCR spotted three candidate positive colonies, F30, F34 and S93, with the length of PCR product larger than 600bp for DNA re-sequencing. S93 had no homology to any known human RNA except an 89% identity with a Papio anubis cDNA clone (GenBank: GE901207.1) and was not pursued. Purified F30 plasmid enabled colony growth as shown in selection plate of Supplementary Fig. 1A. Interestingly, both F30 and F34 represented the same human 5-Azacytidine Induced 2 (AZI2) cDNA. Specifically, the sequences matched the second half of the AZI2 3′UTR (Supplementary Fig. 1B).

To find out whether F30 was a truncated cDNA fragment of AZI2, Northern blotting was carried out to examine the presence of F30 in six cultured human cell lines. The results suggested that F30 was not a truncated cDNA and confirmed F30 as an independent transcript in all of the examined human cell lines (Supplementary Fig. 1C). This observation was unexpected given that a protein would make better sense as the cloned product of Y1H rather than an RNA fragment. We then studied whether F30, a part of 3′UTR, actually conferred any regulatory activity. F30 and 7.9kb-B, which was a 7.9 kb *SLC6A3* promoter fragment carrying DNPiB and cloned into a luciferase reporter vector, were co-overexpressed in the human neuroblastoma cell line SH-SY5Y. The results showed that F30 was able to downregulate *SLC6A3* promoter activity (Supplementary Fig. 1D). These data suggested that F30 had independent regulatory activity in human systems and this 1,016 nucleotide RNA has been termed ^{AZI2}3′UTR here due to its identity to a portion of the human AZI2's 3′UTR. We will still use "F30" to describe the cDNA clone for conciseness in the rest of the description of these experiments.

Allele-dependent regulation of exogenous promoters by ^{AZI2}3′UTR

Next we examined whether ^{AZI2}3′UTR/F30 could regulate promoter activity in an allele-dependent manner. We tested two promoters, SV40 and *SLC6A3* in the human neuroblastoma cell line SK-N-AS. The results showed that, indeed, F30 downregulated both promoter activities in a DNPiB-dependent manner (Fig. 1a–d). This DNPiB-dependent downregulation of *SLC6A3* promoter by F30 was also observed in HEK293T cells (Supplementary Fig. 2). To further verify allele-dependence, RNAi was used to target 3′UTR for effects on DNPi-regulated SV40 promoter activity. Expectedly, 3′UTR RNAi upregulated SV40 promoter activity only through DNPiB and not DNPiA. This DNPiB-dependence of the RNAi effects was also observed in SH-SY5Y cells (Supplementary Fig. 3). Although these RNAi results could be also attributable to AZI2 downregulation, together, these observations suggest that ^{AZI2}3′UTR downregulates promoter activity via DNPiB in human transient expression systems.

^{AZI2}3′UTR regulation of endogenous *SLC6A3* activity

In the human neuroblastoma cell line BE(2)-M17, overexpression of F30 downregulated endogenous *SLC6A3* activity dose-dependently (Fig. 1e–f). We chose BE(2)-M17 because it

allowed examination of haplotype-dependent regulation[31]. Whereas we initially looked at possible allele-dependence in $^{AZI2}3'$UTR regulation of endogenous gene activity, this assay lacked the needed sensitivity to detect this effect.

Localization of $^{AZI2}3'$UTR to nuclei of BE(2)-M17 cells

To exert transcriptional activity, $^{AZI2}3'$UTR must be present in cell nuclei. We therefore used two approaches to identify the subcellular location of $^{AZI2}3'$UTR in BE(2)-M17. In the first approach, cytoplasmic and nuclear fractions were separated from each other and subjected to Northern blotting. As shown in Fig. 2A–B, the main form of the 3′UTR was $^{AZI2}3'$UTR, not AZI2. AZI2 and $^{AZI2}3'$UTR were both present in the cytoplasm but $^{AZI2}3'$UTR was enriched by 3-fold in the nuclei in comparison (Fig. 2C). The second approach used RNAscope analysis combined with immunofluorescence (IF) staining. This dual approach revealed that the 5′ probe, which selectively targeted AZI2, was more highly sensitive. A comparison of the 3′ to 5′ probe labeled AZI2 labeled mRNA revealed that $^{AZI2}3'$UTR was also enriched in the nuclei. (Fig. 2D vs 2E). This result is supported semi-quantitatively by IF intensity analysis (Supplementary Fig. 4).

Localization of $^{AZI2}3'$UTR to nuclei of human *post mortem* dopamine neurons

We used the human probes and validated the nuclear localization in dopamine neurons of human *post mortem* tissue from two subjects. In these *post mortem* studies, the 5′ probe again displayed greater sensitivity and labeled the AZI2 mRNA both in the cytoplasm and nucleus. In contrast, the 3′ probe, which targets the 3′UTR, revealed labeling mainly in the nuclei of TH-positive neurons from both subjects (Fig. 3). This nuclear enrichment of $^{AZI2}3'$UTR, as observed by the semi-quantification of IF mean intensity or labeling clusters (Supplementary Fig. 5), was consistent with the Northern blotting finding that $^{AZI2}3'$UTR was enriched in human midbrain tissue (Fig. 3 ***Insert***).

Promoter activity upstream of $^{AZI2}3'$UTR

$^{AZI2}3'$UTR expression was enriched, compared with AZI2, particularly in the midbrain, compared with the hippocampus and cortex of mice, rats and humans using Northern blot analyses (Supplementary Fig. 6). These findings suggested that $^{AZI2}3'$UTR was possibly transcribed independently of AZI2. To investigate this possibility, three human chromosomal DNA fragments (0.1kb, 0.3 kb, and 1.6 kb) immediately upstream of $^{AZI2}3'$UTR were used. Transcriptional regulatory activity was explored using the promoter report vector pGL3 Enhancer in two dopamine neuron models, SK-N-AS and SN4741. *Student's* t-test indicated that both the 0.1 kb and the 1.6 kb fragments inhibited the basal promoter activity carried by the vector down to 0.64- ($P < 0.0001$) and 0.52-fold ($P < 0.0001$) in SK-N-AS, and 0.90- ($P = 0.04$) and 0.66-fold ($P < 0.0001$) of the vector activity in SN4741. In contrast, the 0.3 kb fragment displayed clear promoter activity in both cell lines (Fig. 4). These data suggested that $^{AZI2}3'$UTR had its own promoter which was under different inhibitory regulations.

Implication of $^{AZI2}3'$UTR/AZI2 in SUDs

To explore the possible association of this AZI2/$^{AZI2}3'$UTR locus with substance abuse, two approaches were used. The first used a well-established rat animal model of alcoholism and

the second involved a secondary analysis of GWAS datasets in dbGaP[32]. The rat model used was the bidirectionally selected alcohol-preferring P *vs* the alcohol-nonpreferring NP rats[33,34]. Northern blotting analysis showed that P rats had significantly decreased $^{AZI2}3'$UTR/AZI2 expression ratio compared with NP rats in midbrain.(Fig. 5A–B), consistent with findings from qRT-PCR analysis of mRNA levels (Fig. 5C). At the same time, the P rats had significantly increased DAT mRNA levels in the midbrain, compared with NP rats (Fig. 5D).

The GWAS secondary analysis used four cohorts from three datasets, Collaborative Study on the Genetics of Alcoholism[35] (COGA, dbGaP accession#: phs000125.v1.p1), Study of Addiction: Genetics and Environment (SAGE, polysubstance abuse but mainly with cigarette smoking, dbGaP accession#: phs000092.v1.p1) and the Australian twin-family study of alcohol use disorder (OZALC, dbGaP accession#: phs000181.v1.p1). COGA had two ethnicities, US Caucasians of European descent and African Americans, so after data quality control (QC), this dataset was split into the two separate ethnic cohorts. Overlapping individuals, those included in both the COGA and SAGE studies, were removed from the SAGE datasets. The OZALC study used family genotype but only the unrelated individuals were extracted for the purpose of epistasis analysis. After QC, a total of 6,596 informative subjects remained, including 41.0% cases: 1,368 from COGA_EA, 753 from COGA_AA, 2,063 from the OZALC study, and 2,412 from the SAGE study. Logistic regression was used for the epistasis analysis of gene-gene interaction. These cleaned datasets showed no significant main effects at the AZI2 locus (data not shown). By contrast, epistasis analysis revealed several interactions between the DNPi/core promoter region of *SLC6A3* and the AZI2 locus (Fig. 5D). One of them, the most significant one in double-arrow curve, was between rs2617605 (1 kb downstream of DNPi or at 1,442,521 in chr5) and rs12054402 (25 kb upstream of $^{AZI2}3'$UTR or at 28,384,051 in chr3), with a *P*-value of 1.25×10^{-11} and an odds ratio (OR) of 0.42 using a random effects model, followed by Bonferroni corrections. It is noteworthy that the most significant interactions were from the upstream region of $^{AZI2}3'$UTR, not upstream of AZI2 (the DNPi alleles were not typed in the original GWAS studies nor were they captured by imputation).

DISCUSSION

$^{AZI2}3'$UTR is the first inhibitory regulator, the first allele-dependent regulator and also the first ribonucleic regulator of *SLC6A3* based on the *in vitro* data. The data presented here suggests AZI23′UTR protects against SUD. It appears to be the first regulatory ribonucleic acid shared by a 3′UTR of another gene. Therefore, this discovery describes the first kangaroo layout of two genes, a protein gene and a RNA gene (a small RNA gene in the end of a large protein gene like a pup in kangaroo's pocket). It is unclear how the Y1H system detected it because Y1H has been designated to clone proteinaceous activity. One possibility for its detection could be that the DNPiB sequence may have been recognized by a yeast endogenous protein so that $^{AZI2}3'$UTR regulated the recognition for DNPiB specifically. Thus, the putative proteinaceous activity of this assay might help explain allele-dependent transcriptional regulation in exogenous expression systems. An attempt to investigate possible RNA-DNPiB binding activity by R3C failed, possibly because of the indirect regulation. Nested deletions from either the 5′ or the 3′ end aimed to screen for possible

functional domains or small open reading frames in two human neuroblastoma cell lines (SK-N-AS and SH-SY5Y) revealed no conclusive results. The question is still under investigation.

By molecular size and location, $^{AZI2}3'$UTR belongs to the category of long non-coding RNAs (lncRNAs). The genome encodes thousands of lncRNAs and some of them are characterized suggesting spatial or structural regulatory activity[36]. However, the exact mechanisms by which these lncRNAs regulate gene expression are still under investigation in various laboratories[37–41]. It is likely that $^{AZI2}3'$UTR is expressed under the control of its own promoter which seems highly regulated. The disproportion between $^{AZI2}3'$UTR and AZI2 RNA abundance in the midbrain of different species, supports this view of an independent promoter. $^{AZI2}3'$UTR appeared ubiquitously expressed but the biological activity of this widespread RNA requires continued research. For instance, it will be important to clarify if it plays a significant role in the predisposition to the development of SUDs.

Full-gene epistasis analysis found extensive and statistically significant single nucleotide polymorphism (SNP)-SNP interactions between the two genes in the genetic etiology of SUDs. Many SNPs throughout the AZI2 gene interacted with those located in the downstream, or near the 3′UTR, region of *SLC6A3*. The most significant interactions were from the upstream region of $^{AZI2}3'$UTR, rather than upstream of AZI2 (Supplementary Fig. 7*,* see Supplementary Table 2 for details). These data suggest two possibilities. One possibility is that transcriptional regulation of *SLC6A3* plays a greater role in psychopathology than previously thought, or AZI2 mRNA contributes to transcriptional regulation of *SLC6A3* as well. This would be consistent with the observation that the AZI2 RNA was detected in the nuclei of both BE(2)-M17 cells and human *post mortem* dopamine neurons. This data prepares a cellular platform for future study of this lncRNA. Moreover, a potential role for AZI2 in SUDs is supported by a report that the mouse AZI2 is related to sensitivity to amphetamine[42].

The present study highlights the need for further research in this area. We need to understand whether $^{AZI2}3'$UTR overexpression regulates animal DAT gene *in vivo*, reduces ethanol intake in a genetic animal model of alcoholism, the P rat, and even is able to attenuate relapse in models for SUDs. Additionally, delineation of other functional elements underlying these human gene-gene interactions will be another new project. Thus, continued functional studies of $^{AZI2}3'$UTR is warranted. In conclusion, we have identified $^{AZI2}3'$UTR as a new and independent transcriptional regulator of *SLC6A3*, and found that this transcriptional regulation is associated with an epistatic protection against SUDs in humans.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank NIH for granting ZL access to dbGaP, under project #1542: “Functional genomics of dopamine-related diseases” as well as NIDA funding of DA021409. We are grateful to Kerry J. Ressler and his Neurobiology of Fear

Laboratory at McLean Hospital for assistance in RNAscope analysis, and to the Harvard Brain Tissue Resource Center for providing the human brain tissue samples and associated information for these investigations. We thank International Graduate Exchange Program of Beijing Institute of Technology for supporting KL's training at McLean Hospital in U.S.A.".

References

1. Girault JA, Greengard P. The neurobiology of dopamine signaling. Archives of neurology. 2004; 61(5):641–644. DOI: 10.1001/archneur.61.5.641 [PubMed: 15148138]
2. Rosen ZB, Cheung S, Siegelbaum SA. Midbrain dopamine neurons bidirectionally regulate CA3-CA1 synaptic drive. Nat Neurosci. 2015; 18(12):1763–1771. DOI: 10.1038/nn.4152 [PubMed: 26523642]
3. Howe MW, Tierney PL, Sandberg SG, Phillips PE, Graybiel AM. Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. Nature. 2013; 500(7464):575–579. DOI: 10.1038/nature12475 [PubMed: 23913271]
4. Perry CJ, Baciadonna L, Chittka L. Unexpected rewards induce dopamine-dependent positive emotion-like state changes in bumblebees. Science (New York, NY). 2016; 353(6307):1529–1531. DOI: 10.1126/science.aaf4454
5. Gadagkar V, Puzerey PA, Chen R, Baird-Daniel E, Farhang AR, Goldberg JH. Dopamine neurons encode performance error in singing birds. Science (New York, NY). 2016; 354(6317):1278–1282. DOI: 10.1126/science.aah6837
6. Vaughan RA, Foster JD. Mechanisms of dopamine transporter regulation in normal and disease states. Trends Pharmacol Sci. 2013; 34(9):489–496. DOI: 10.1016/j.tips.2013.07.005 [PubMed: 23968642]
7. Calipari ES, Juarez B, Morel C, Walker DM, Cahill ME, Ribeiro E, Roman-Ortiz C, Ramakrishnan C, Deisseroth K, Han MH, Nestler EJ. Dopaminergic dynamics underlying sex-specific cocaine reward. Nat Commun. 2017; 8:13877.doi: 10.1038/ncomms13877 [PubMed: 28072417]
8. Wang KH, Penmatsa A, Gouaux E. Neurotransmitter and psychostimulant recognition by the dopamine transporter. Nature. 2015; 521(7552):322–327. DOI: 10.1038/nature14431 [PubMed: 25970245]
9. Volkow ND, Wang GJ, Fischman MW, Foltin RW, Fowler JS, Abumrad NN, Vitkun S, Logan J, Gatley SJ, Pappas N, Hitzemann R, Shea CE. Relationship between subjective effects of cocaine and dopamine transporter occupancy. Nature. 1997; 386(6627):827–830. DOI: 10.1038/386827a0 [PubMed: 9126740]
10. Tiihonen J, Kuikka J, Bergstrom K, Hakola P, Karhu J, Ryynanen OP, Fohr J. Altered striatal dopamine re-uptake site densities in habitually violent and non-violent alcoholics. Nat Med. 1995; 1(7):654–657. [PubMed: 7585146]
11. Faraone SV, Spencer TJ, Madras BK, Zhang-James Y, Biederman J. Functional effects of dopamine transporter gene genotypes on in vivo dopamine transporter functioning: a meta-analysis. Molecular psychiatry. 2014; 19(8):880–889. DOI: 10.1038/mp.2013.126 [PubMed: 24061496]
12. van der Voet M, Harich B, Franke B, Schenck A. ADHD-associated dopamine transporter, latrophilin and neurofibromin share a dopamine-related locomotor signature in Drosophila. Molecular psychiatry. 2016; 21(4):565–573. DOI: 10.1038/mp.2015.55 [PubMed: 25962619]
13. Hamilton PJ, Campbell NG, Sharma S, Erreger K, Herborg Hansen F, Saunders C, Belovich AN, Sahai MA, Cook EH, Gether U, McHaourab HS, Matthies HJ, Sutcliffe JS, Galli A. De novo mutation in the dopamine transporter gene associates dopamine dysfunction with autism spectrum disorder. Molecular psychiatry. 2013; 18(12):1315–1323. DOI: 10.1038/mp.2013.102 [PubMed: 23979605]
14. Kurian MA, Zhen J, Cheng SY, Li Y, Mordekar SR, Jardine P, Morgan NV, Meyer E, Tee L, Pasha S, Wassmer E, Heales SJ, Gissen P, Reith ME, Maher ER. Homozygous loss-of-function mutations in the gene encoding the dopamine transporter are associated with infantile parkinsonism-dystonia. The Journal of clinical investigation. 2009; 119(6):1595–1603. DOI: 10.1172/jci39060 [PubMed: 19478460]

15. Kurian MA, Li Y, Zhen J, Meyer E, Hai N, Christen HJ, Hoffmann GF, Jardine P, von Moers A, Mordekar SR, O'Callaghan F, Wassmer E, Wraige E, Dietrich C, Lewis T, Hyland K, Heales S Jr, Sanger T, Gissen P, Assmann BE, Reith ME, Maher ER. Clinical and molecular characterisation of hereditary dopamine transporter deficiency syndrome: an observational cohort and experimental study. The Lancet Neurology. 2011; 10(1):54–62. DOI: 10.1016/s1474-4422(10)70269-6 [PubMed: 21112253]

16. Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR. A global reference for human genetic variation. Nature. 2015; 526(7571):68–74. DOI: 10.1038/nature15393 [PubMed: 26432245]

17. Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, Zhang Y, Ye K, Jun G, Hsi-Yang Fritz M, Konkel MK, Malhotra A, Stutz AM, Shi X, Paolo Casale F, Chen J, Hormozdiari F, Dayama G, Chen K, Malig M, Chaisson MJ, Walter K, Meiers S, Kashin S, Garrison E, Auton A, Lam HY, Jasmine Mu X, Alkan C, Antaki D, Bae T, Cerveira E, Chines P, Chong Z, Clarke L, Dal E, Ding L, Emery S, Fan X, Gujral M, Kahveci F, Kidd JM, Kong Y, Lameijer EW, McCarthy S, Flicek P, Gibbs RA, Marth G, Mason CE, Menelaou A, Muzny DM, Nelson BJ, Noor A, Parrish NF, Pendleton M, Quitadamo A, Raeder B, Schadt EE, Romanovitch M, Schlattl A, Sebra R, Shabalin AA, Untergasser A, Walker JA, Wang M, Yu F, Zhang C, Zhang J, Zheng-Bradley X, Zhou W, Zichner T, Sebat J, Batzer MA, McCarroll SA, Mills RE, Gerstein MB, Bashir A, Stegle O, Devine SE, Lee C, Eichler EE, Korbel JO. An integrated map of structural variation in 2,504 human genomes. Nature. 2015; 526(7571):75–81. DOI: 10.1038/nature15394 [PubMed: 26432246]

18. Bannon MJ, Pruetz B, Manning-Bog AB, Whitty CJ, Michelhaugh SK, Sacchetti P, Granneman JG, Mash DC, Schmidt CJ. Decreased expression of the transcription factor NURR1 in dopamine neurons of cocaine abusers. Proceedings of the National Academy of Sciences of the United States of America. 2002; 99(9):6382–6385. DOI: 10.1073/pnas.092654299 [PubMed: 11959923]

19. Wang J, Bannon MJ. Sp1 and Sp3 activate transcription of the human dopamine transporter gene. J Neurochem. 2005; 93(2):474–482. DOI: 10.1111/j.1471-4159.2005.03051.x [PubMed: 15816870]

20. Jacobs FM, van der Linden AJ, Wang Y, von Oerthel L, Sul HS, Burbach JP, Smidt MP. Identification of Dlk1, Ptpru and Klhl1 as novel Nurr1 target genes in meso-diencephalic dopamine neurons. Development (Cambridge, England). 2009; 136(14):2363–2373. DOI: 10.1242/dev.037556

21. Kanno K, Ishiura S. Differential effects of the HESR/HEY transcription factor family on dopamine transporter reporter gene expression via variable number of tandem repeats. J Neurosci Res. 2011; 89(4):562–575. DOI: 10.1002/jnr.22593 [PubMed: 21290414]

22. Hwang DY, Hong S, Jeong JW, Choi S, Kim H, Kim J, Kim KS. Vesicular monoamine transporter 2 and dopamine transporter are molecular targets of Pitx3 in the ventral midbrain dopamine neurons. J Neurochem. 2009; 111(5):1202–1212. DOI: 10.1111/j.1471-4159.2009.06404.x [PubMed: 19780901]

23. Jankovic J, Chen S, Le WD. The role of Nurr1 in the development of dopaminergic neurons and Parkinson's disease. Progress in neurobiology. 2005; 77(1–2):128–138. DOI: 10.1016/j.pneurobio.2005.09.001 [PubMed: 16243425]

24. Lou X, Liao W. Association of Nurr1 gene mutations with Parkinson's disease in the Han population living in the Hubei province of China. Neural regeneration research. 2012; 7(23):1791–1796. DOI: 10.3969/j.issn.1673-5374.2012.23.005 [PubMed: 25624803]

25. Zhao Y, Zhou Y, Xiong N, Lin Z. Identification of an intronic cis-acting element in the human dopamine transporter gene. Mol Biol Rep. 2012; 39(5):5393–5399. DOI: 10.1007/s11033-011-1339-4 [PubMed: 22160470]

26. Zhang H, Zeitz MJ, Wang H, Niu B, Ge S, Li W, Cui J, Wang G, Qian G, Higgins MJ, Fan X, Hoffman AR, Hu JF. Long noncoding RNA-mediated intrachromosomal interactions promote imprinting at the Kcnq1 locus. J Cell Biol. 2014; 204(1):61–75. DOI: 10.1083/jcb.201304152 [PubMed: 24395636]

27. Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT. Data quality control in genetic case-control association studies. Nat Protoc. 2010; 5(9):1564–1573. DOI: 10.1038/nprot.2010.116 [PubMed: 21085122]

28. Kennedy JL, Xiong N, Yu J, Zai CC, Pouget JG, Li J, Liu K, Qing H, Wang T, Martin E, Levy DL, Lin Z. Increased Nigral SLC6A3 Activity in Schizophrenia Patients: Findings From the Toronto-McLean Cohorts. Schizophr Bull. 2016; 42(3):772–781. DOI: 10.1093/schbul/sbv191 [PubMed: 26707863]

29. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81(3):559–575. DOI: 10.1086/519795 [PubMed: 17701901]

30. Ueki M, Cordell HJ. Improved statistics for genome-wide interaction analysis. PLoS Genet. 2012; 8(4):e1002625.doi: 10.1371/journal.pgen.1002625 [PubMed: 22496670]

31. Zhou Y, Michelhaugh SK, Schmidt CJ, Liu JS, Bannon MJ, Lin Z. Ventral midbrain correlation between genetic variation and expression of the dopamine transporter gene in cocaine-abusing versus non-abusing subjects. Addict Biol. 2014; 19(1):122–131. DOI: 10.1111/j.1369-1600.2011.00391.x [PubMed: 22026501]

32. Mailman MD, Feolo M, Jin Y, Kimura M, Tryka K, Bagoutdinov R, Hao L, Kiang A, Paschall J, Phan L, Popova N, Pretel S, Ziyabari L, Lee M, Shao Y, Wang ZY, Sirotkin K, Ward M, Kholodov M, Zbicz K, Beck J, Kimelman M, Shevelev S, Preuss D, Yaschenko E, Graeff A, Ostell J, Sherry ST. The NCBI dbGaP database of genotypes and phenotypes. Nat Genet. 2007; 39(10):1181–1186. DOI: 10.1038/ng1007-1181 [PubMed: 17898773]

33. Bell RL, Rodd ZA, Lumeng L, Murphy JM, McBride WJ. The alcohol-preferring P rat and animal models of excessive alcohol drinking. Addict Biol. 2006; 11(3–4):270–288. DOI: 10.1111/j.1369-1600.2005.00029.x [PubMed: 16961759]

34. Mello NK, Mendelson JH. OPERANT PERFORMANCE BY RATS FOR ALCOHOL REINFORCEMENT. A COMPARISON OF ALCOHOL-PREFERRING AND NONPREFERRING ANIMALS. Quarterly journal of studies on alcohol. 1964; 25:226–234. [PubMed: 14161785]

35. Grucza RA, Bierut LJ. Co-occurring risk factors for alcohol dependence and habitual smoking: update on findings from the Collaborative Study on the Genetics of Alcoholism. Alcohol research & health: the journal of the National Institute on Alcohol Abuse and Alcoholism. 2006; 29(3):172–178. [PubMed: 17373405]

36. Engreitz JM, Ollikainen N, Guttman M. Long non-coding RNAs: spatial amplifiers that control nuclear structure and gene expression. Nature reviews Molecular cell biology. 2016; 17(12):756–770. DOI: 10.1038/nrm.2016.126 [PubMed: 27780979]

37. Kotzin JJ, Spencer SP, McCright SJ, Kumar DB, Collet MA, Mowel WK, Elliott EN, Uyar A, Makiya MA, Dunagin MC, Harman CC, Virtue AT, Zhu S, Bailis W, Stein J, Hughes C, Raj A, Wherry EJ, Goff LA, Klion AD, Rinn JL, Williams A, Flavell RA, Henao-Mejia J. The long non-coding RNA Morrbid regulates Bim and short-lived myeloid cell lifespan. Nature. 2016; 537(7619):239–243. DOI: 10.1038/nature19346 [PubMed: 27525555]

38. Anderson KM, Anderson DM, McAnally JR, Shelton JM, Bassel-Duby R, Olson EN. Transcription of the non-coding RNA upperhand controls Hand2 expression and heart development. Nature. 2016; 539(7629):433–436. DOI: 10.1038/nature20128 [PubMed: 27783597]

39. Sallam T, Jones MC, Gilliland T, Zhang L, Wu X, Eskin A, Sandhu J, Casero D, Vallim TQ, Hong C, Katz M, Lee R, Whitelegge J, Tontonoz P. Feedback modulation of cholesterol metabolism by the lipid-responsive non-coding RNA LeXis. Nature. 2016; 534(7605):124–128. DOI: 10.1038/nature17674 [PubMed: 27251289]

40. Leucci E, Vendramin R, Spinazzi M, Laurette P, Fiers M, Wouters J, Radaelli E, Eyckerman S, Leonelli C, Vanderheyden K, Rogiers A, Hermans E, Baatsen P, Aerts S, Amant F, Van Aelst S, van den Oord J, de Strooper B, Davidson I, Lafontaine DL, Gevaert K, Vandesompele J, Mestdagh P, Marine JC. Melanoma addiction to the long non-coding RNA SAMMSON. Nature. 2016; 531(7595):518–522. DOI: 10.1038/nature17161 [PubMed: 27008969]

41. Gong C, Maquat LE. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3′ UTRs via Alu elements. Nature. 2011; 470(7333):284–288. DOI: 10.1038/nature09701 [PubMed: 21307942]

42. Parker CC, Gopalakrishnan S, Carbonetto P, Gonzales NM, Leung E, Park YJ, Aryee E, Davis J, Blizard DA, Ackert-Bicknell CL, Lionikas A, Pritchard JK, Palmer AA. Genome-wide association

study of behavioral, physiological and gene expression traits in outbred CFW mice. Nat Genet. 2016; 48(8):919–926. DOI: 10.1038/ng.3609 [PubMed: 27376237]

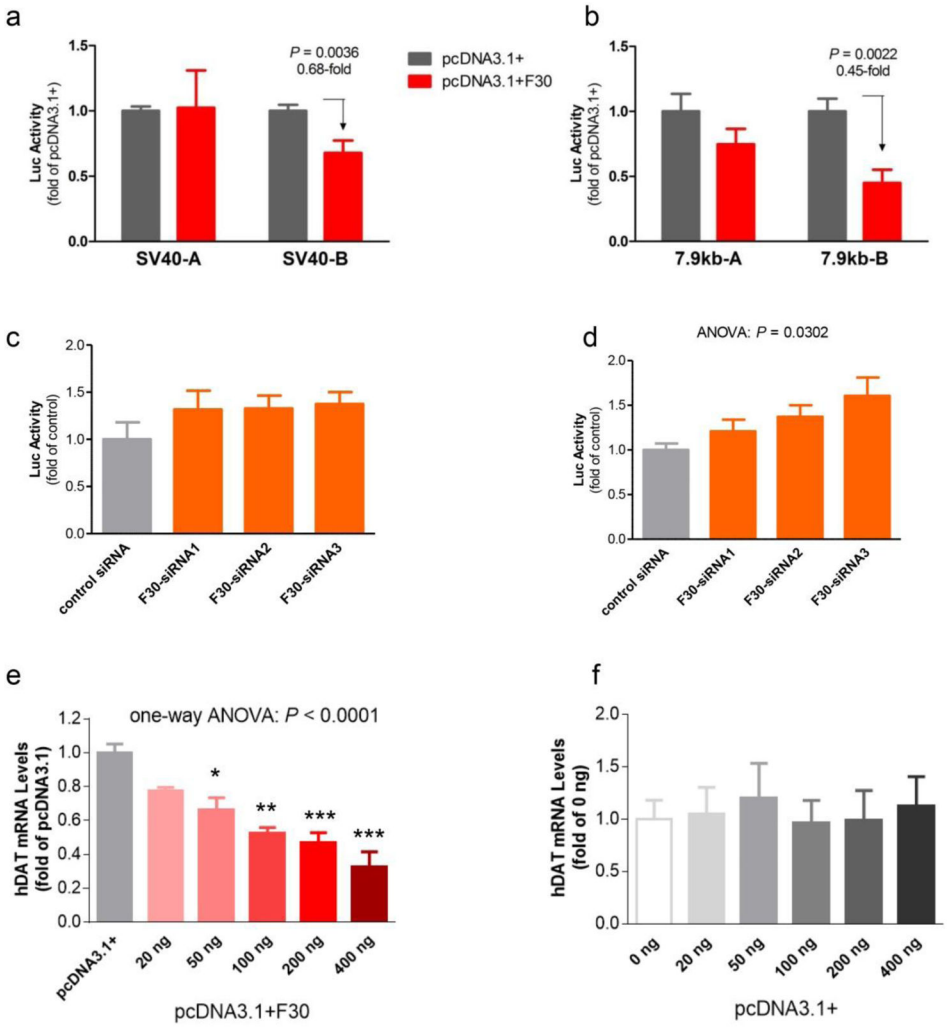

Figure 1. F30 (^{AZI2}3′UTR) allele-dependent downregulation promoters activity by luciferase reporting in the human neuroblastoma SK-N-AS cells and downregulation of endogenous *SLC6A3* mRNA level in the human neuroblastoma BE(2)-M17 cells

(a) SV40 promoter or (b) on 7.9kb *SLC6A3* promoter, based on effects of F30 vs. vector pcDNA3.1 overexpression. A and B represent promoter carrying DNPiA or DNPiB alleles, respectively. Indicated *P*-value was obtained from *Student*'s t-tests; ANOVA $P = 0.0028$ for F30 effects on both SV40 and 7.9kb SLC6A3 promoters, compared to vector activity ($n = 7$–10); (c) F30 DsiRNA exerts enhancing effects on the SV40-B promoter but not the SV40-A promoter (d), based on one-way ANOVA ($n = 8$–12). Three different F30 DsiRNA molecules were used (sequences listed in Supplementary Table 1). The SV40 data in panel (a) and DsiRNA effects in panels (c) and (d) were replicated in HEK293T and SH-SY5Y cells (see Supplementary Figs. 1 and 2).(e) different doses pcDNA3.1-F30 vector and (f) pcDNA3.1 vector. One-way ANOVA: *, $P < 0.05$; **, $P < 0.001$; ***, $P < 0.0001$ ($n = 3$–6).

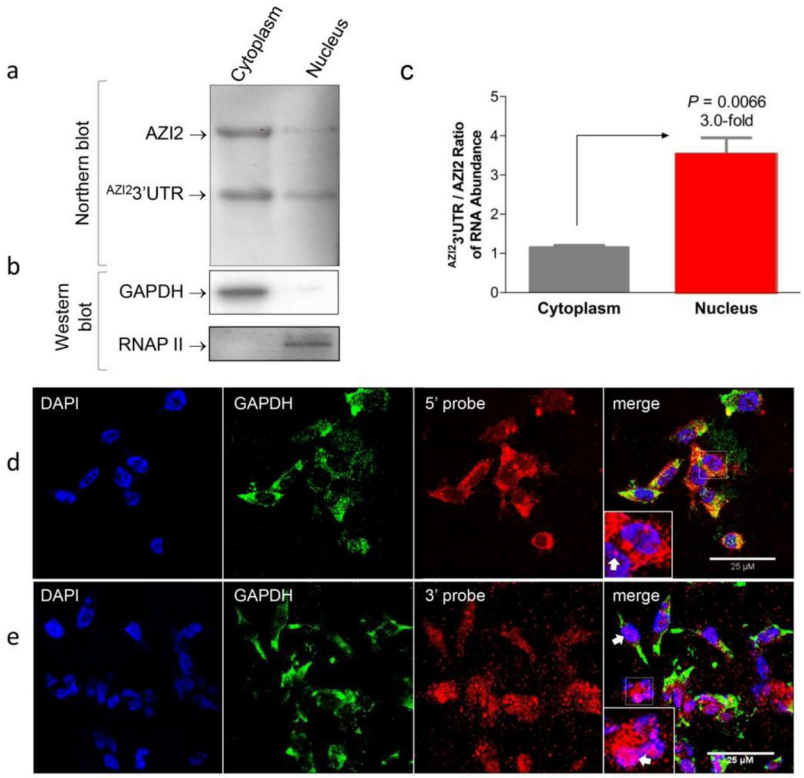

Figure 2. Nuclear expression of $^{AZI2}3'$UTR in human neuroblastoma BE(2)-M17 cells

(a) Nuclear enrichment of $^{AZI2}3'$UTR expression, compared to cytoplasmic expression, by F30-based Northern blotting. Noticed that loaded RNA amount for nuclear fraction was only 25% of what was for cytoplasmic fraction. (b) Clean nuclear fraction separation from the cytoplasmic fraction by GAPDH- and RNAP II-based Western blots. (c) Quantification of Northern blot-based nuclear enrichment. *P*-value was obtained by *Student*'s t-tests ($n = 3$). (d) Low density of nuclear AZI2 RNA and (e) enriched nuclear expression of $^{AZI2}3'$UTR RNA both revealed by confocal images from RNAscope analyses. DAPI, nucleus; GAPDH, cytoplasmic control in immunostaining; 5′ probe, for targeting AZI2 mRNA only; 3′ probe, for targeting 3′UTR (both AZI2 and $^{AZI2}3'$UTR). *Insert*, close-up showing nuclear expression of 3′UTR with GAPDH staining filtered out. White arrow, nuclear expression of 3′UTR. Scale bars, 25 μm.

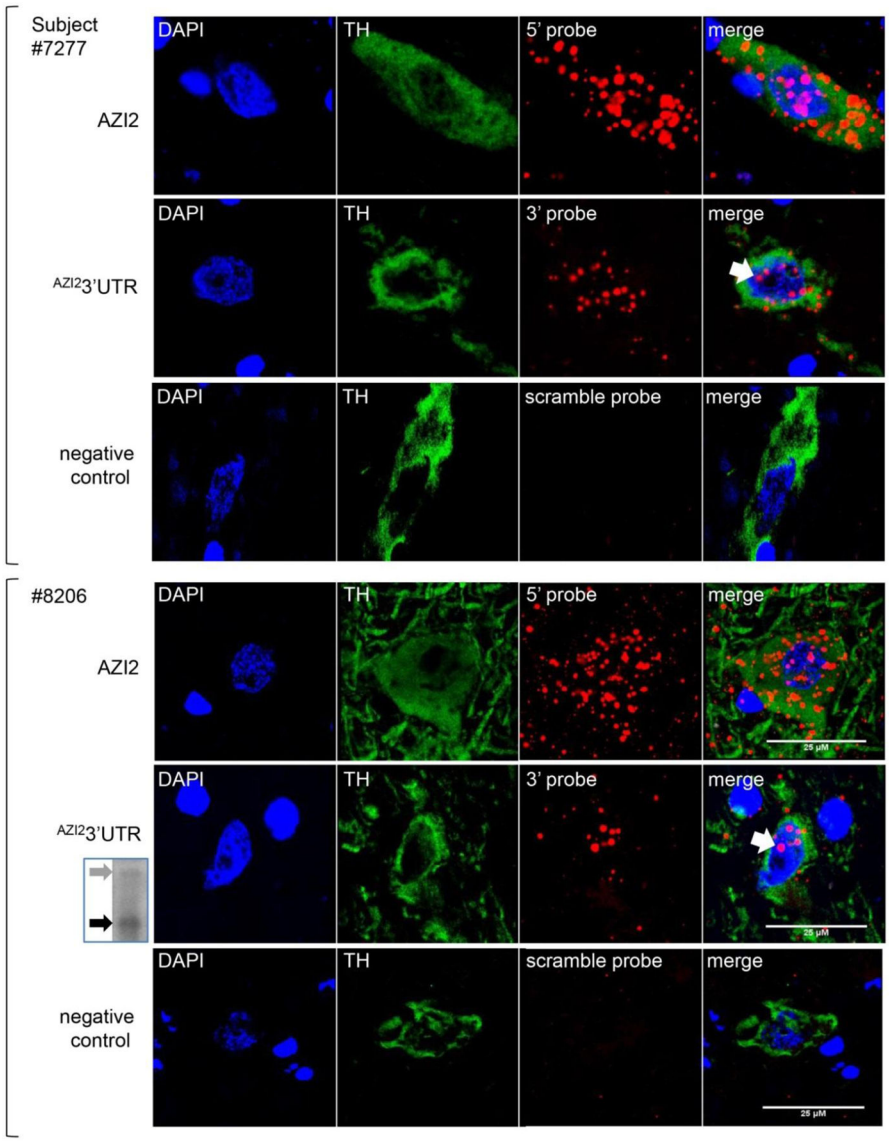

Figure 3. Nuclear localization of 3′UTR expression in human *post mortem* nigral dopamine neurons from two individuals (upper three rows from #7277, a 58 year old male and lower three rows from #8206, a 93 year old female) based on confocal images from RNAscope analysis

The consistent findings, despite different demographics, suggests the generalizability of these findings. Three up-to-down rows represent three RNA probes, 5′ probe for AZI2, 3′ probe for 3′UTR and scrambled probe for negative control; four left-to-right columns of images are DAPI for nucleus, TH (tyrosine hydroxylase protein immunostaining) for labeling DA neurons, RNA probing, and merging all three. White arrow, nuclear expression of 3′UTR by 3′ probe; Scale bars, 25 μm. *Insert on the lower left*: Northern blot for enriched expression of ^{AZI2}3′UTR (black arrow), compared to AZI2 (gray arrow), in the same female (see supplementary data for an enlarged view).

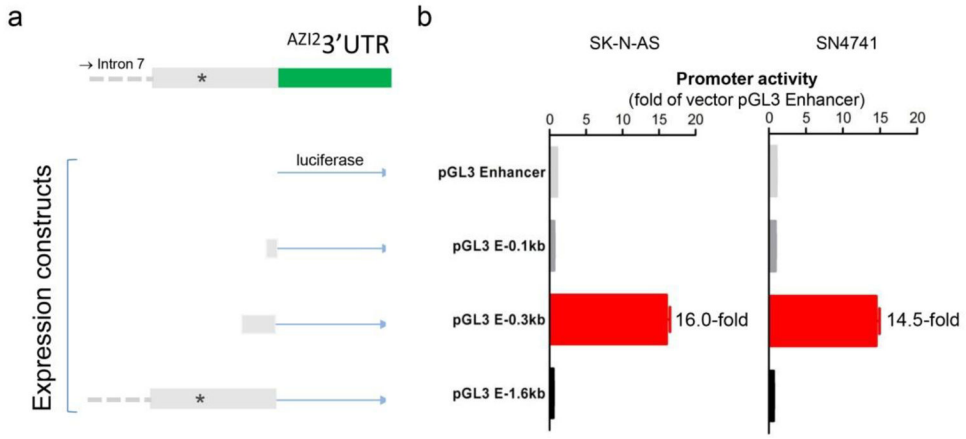

Figure 4. Identification of a regulated promoter upstream of $^{AZI2}3'$UTR

(a) Construction of three promoter reporters, with immediately upstream 0.1 kb, 0.3 kb and 1.6 kb (1,585 bp) fragments, in Promega's pGL3 Enhancer vector carrying the firefly luciferase gene. (b) Luciferase activity detected in SK-N-AS (left) and SN4741 (right) cell lines. Each dataset came from four independent experiments.

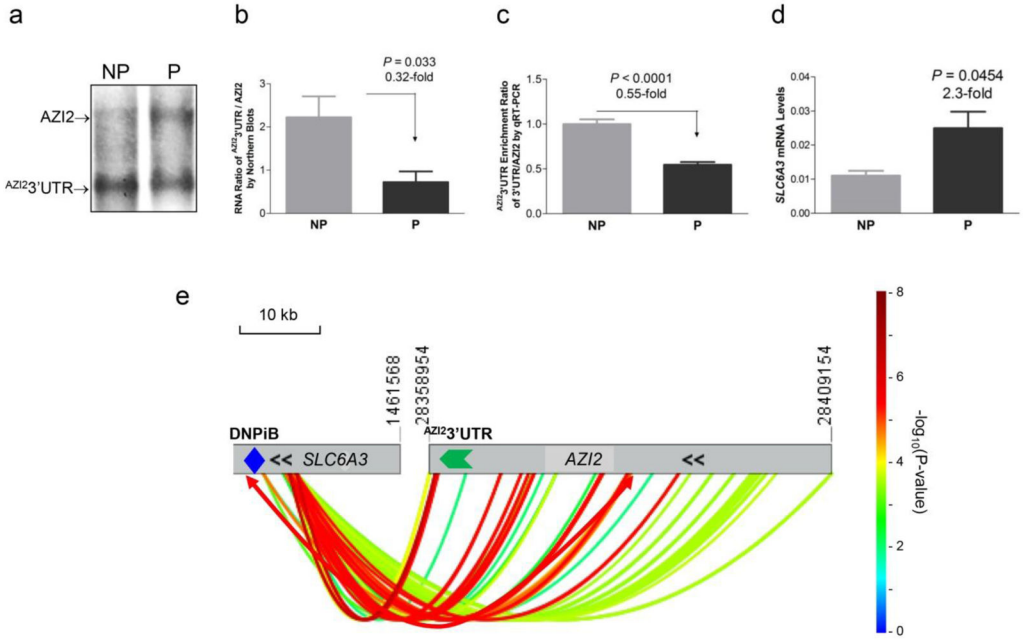

Figure 5. Implication of the $^{AZI2}3'$UTR locus in genetic etiology of SUDs

(a–c) Reciprocal correlation between $^{AZI2}3'$UTR and DAT expression in midbrain from a genetic animal model of alcoholism male P vs NP rats. (a) A representative Northern blot showing increased $^{AZI2}3'$UTR expression in the NP, compared to P rats. (b) Quantification of Northern blot data as in panel (a). (c) $^{AZI2}3'$UTR enrichment ratio in NP vs P based on qRT-PCR analysis of mRNA levels in midbrain VTA. (d) Increased DAT mRNA levels in P, compared to NP rats. *P*-values obtained from *Student*'s tests ($n = 3–4$). (**e**) Human genetics of SUDs via SNP interactions (colored curves) between a *SLC6A3* intron 1/core promoter region around DNPi (blue diamond) and the $^{AZI2}3'$UTR (green arrow)/AZI2 locus. Data are from meta-analysis of logistic regression results by using three dbGaP datasets of SUDs (alcohol and cigarette smoking). For *SLC6A3*, only the $^{AZI2}3'$UTR target or DNPi surrounding region is shown (see Supplementary Fig. 7 for full scale interactions between the two loci/three genes). Gray horizontal bars, on the left is *SLC6A3* in chr5 and on the right is the $^{AZI2}3'$UTR/AZI2 loci in chr3 with coordinates indicated above the bars; double black arrow symbols, location of transcription start site of SLC6A3 or AZI2 (both run on the minus strands of the chromosomes); thermometer bar on the right, interaction strength (*P*-value) in the form of $-\text{Log}_{10}$(*P*-value); double-arrowed curve, the most significant interaction (rs2617605 at 1442521 of chr5 vs rs12054402 at 28384051 of chr3) which is outside the range of the indicated thermometer bar scale; chromosomal scale, 10 kb.