Transcriptome analysis of genetically matched human induced pluripotent stem cells disomic or trisomic for chromosome 21

Patrick K. Gonzales[1,2], Christine M. Roberts[1,2], Virginia Fonte[1,2], Connor Jacobsen[2], Gretchen H. Stein[1,3], Christopher D. Link[1,2]*

1 Linda Crnic Institute for Down Syndrome, University of Colorado, Aurora, Colorado, United States of America, 2 Department of Integrative Physiology, University of Colorado, Boulder, Colorado, United States of America, 3 Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, Colorado, United States of America

* linkc@colorado.edu

Abstract

Trisomy of chromosome 21, the genetic cause of Down syndrome, has the potential to alter expression of genes on chromosome 21, as well as other locations throughout the genome. These transcriptome changes are likely to underlie the Down syndrome clinical phenotypes. We have employed RNA-seq to undertake an in-depth analysis of transcriptome changes resulting from trisomy of chromosome 21, using induced pluripotent stem cells (iPSCs) derived from a single individual with Down syndrome. These cells were originally derived by Li et al, who genetically targeted chromosome 21 in trisomic iPSCs, allowing selection of disomic sibling iPSC clones. Analyses were conducted on trisomic/disomic cell pairs maintained as iPSCs or differentiated into cortical neuronal cultures. In addition to characterization of gene expression levels, we have also investigated patterns of RNA adenosine-to-inosine editing, alternative splicing, and repetitive element expression, aspects of the transcriptome that have not been significantly characterized in the context of Down syndrome. We identified significant changes in transcript accumulation associated with chromosome 21 trisomy, as well as changes in alternative splicing and repetitive element transcripts. Unexpectedly, the trisomic iPSCs we characterized expressed higher levels of neuronal transcripts than control disomic iPSCs, and readily differentiated into cortical neurons, in contrast to another reported study. Comparison of our transcriptome data with similar studies of trisomic iPSCs suggests that trisomy of chromosome 21 may not intrinsically limit neuronal differentiation, but instead may interfere with the maintenance of pluripotency.

Introduction

Down Syndrome (DS) results from an extra copy of chromosome 21, and this change in gene dosage has been proposed to alter chromosome 21 gene expression. Chromosome 21 trisomy also has the potential to alter the global transcriptome, either by secondary effects of chromosome 21 gene over-expression, or as a byproduct of additional genetic material itself [1]. In

addition to the possible perturbation of specific cellular pathways by altered expression of chromosome 21 genes, chromosome 21 also contains genes that impact the global transcriptome directly. These include *ADARB1*, encoding one of two genes responsible for adenosine-to-inosine RNA editing [2]; U2AF1, a constitutive splicing factor [3]; and DYRK1A, a kinase known to target splicing factors [4]. Identifying transcriptome changes reproducibly caused by chromosome 21 trisomy may be crucial to understanding the origins of the clinical features of Down Syndrome.

Initial gene expression studies that sought to establish chromosome 21-dependent transcriptome changes using human material were confounded by the effects of different genetic backgrounds in the DS and control individuals who were studied [5]. An alternative approach is to characterize gene expression in inbred mouse DS models, which reproduce some of the developmental and behavioral phenotypes of DS individuals [6]. Transcriptome studies have been done in multiple mouse DS models [7–11], but no consistent pattern of transcriptome changes has been identified, perhaps due to the different tissues and developmental stages analyzed in these studies. In addition, these mouse DS models have triplication of chromosomal regions syntenic to only part of human chromosome 21 as well as triplication of non-chromosome 21 homologous genes, and thus may not capture the full effects of human chromosome 21 trisomy.

Recently, the development of paired disomic/trisomic induced pluripotent stem cells (iPSCs) derived from the same trisomic individual (or from discordant monozygotic twins) has allowed the measurement of human trisomy 21-dependent gene expression independent of genetic background differences [12–15]. These studies have consistently observed a general upregulation of genes on chromosome 21, as well as highly variable significant gene expression differences on other chromosomes. Li et al [12] used an elegant genetic selection system to derive disomic iPSC subclones from chr21 trisomic iPSCs that were derived from an individual with Down syndrome. Based on an initial microarray analysis, these researchers also reported a general upregulation of chromosome 21 genes, but could identify only 3 non-chromosome 21 genes with a > 2 fold dysregulation between disomic and trisomic iPSC pairs. We have used RNA-seq to significantly extend this analysis using both disomic/trisomic iPSC pairs and cortical neuron cultures derived from them. In addition to assaying transcript accumulation, we have used the RNA-seq data to look for other possible trisomy-induced transcriptome changes, including possible alterations in adenosine-to-inosine RNA editing, alternative splicing, and repetitive element expression.

In contrast to the Jiang et al study [14], the trisomic iPSCs we characterized were not restricted in their ability to differentiate into cortical neuron cultures, and in fact showed higher expression of neuronal transcripts when compared to disomic iPSCs. To investigate differences between our analyses and previous studies, we have compared global gene expression data among the related studies, and find evidence for at least two different states for trisomic iPSC. We suggest that these different states may reflect an inherent inability for trisomic iPSCs to maintain full pluripotency.

Materials and methods

iPSC cell culture

Trisomic (C2 and C3) and disomic iPSC cell lines (C2-4-3, C2-4-4, C3-5-11, and C3-5-13), which had been developed from an individual with Down syndrome by Li et al. [12], were obtained through the Linda Crnic Institute for Down Syndrome (University of Colorado, Aurora, CO). These iPSC lines were grown on Geltrex (Life Technologies) under feeder-free conditions in mTeSR1 medium according to the protocol of Stem Cell Technologies. iPSC

colonies were passaged with dispase, and differentiating colonies were removed manually by scraping. Passage numbers quoted in the Results section refer to passages in our laboratory; we are unable to determine passages performed before we obtained the cells. Periodically, live iPSC colonies were stained with anti-Tra-1-60 or Tra-1-81 conjugated to Dylight fluor 488 to confirm the stem cell character of the colonies. Likewise, fixed cells were shown to express Oct4 by immunostaining. We used ABI/Thermo Fisher TaqMan Copy Number Assays (Hs01180853_cn; Hs02928366_cn; Hs01533676 corresponding to APP, DYRK1A and RCAN1 respectively) to evaluate the number of copies of Chr. 21 in the iPSC lines. We found that C2 was stably trisomic, but that C3 appeared initially to be a mixture of trisomic and disomic cells. At passage 3 of clone C3, our quantitative PCR analysis for Chr 21 loci indicated only a 1.25X relative level of Chr 21 genes rather than the expected 1.5X representation (which we invariably observed for trisomic clone C2). We continued to maintain the C3 clone, and rechecked it at passage 15. At this point all the qPCR analyses (targeting APP, DYRK1A, and RCAN exons) indicated this clone was disomic for Chr 21 loci, which repeated in all subsequent assays. This iPSC clone still expressed the Tra-1-60 or Tra-1-81 markers we routinely used to assay "stemness". We designated this disomic derivative cell line C3-D21.

Neuronal differentiation

iPSC cells were induced to differentiate to cortical neurons by a modification of the method of Espuny-Camacho et al. [16]. Briefly, iPSCs were dissociated with Accutase, and plated on Geltrex-coated dishes at $10–20 \times 10^3$ cells/cm^2 in mTeSR1 medium + 10 µM Y27632 ROCK inhibitor (Tocris). After 2–4 days, the cells formed a meshwork of dense ridges with mostly open spaces in between. The differentiation process was begun (day 0) by feeding the cells every other day for 12 days with NIM neural induction medium (Stem Cell Technologies), followed by DDM medium supplemented with 2% B27 [17] every other day for 12 days. On day 24, the dense ridges or islands containing neural progenitors were manually dissociated, resuspended in DDM + 2% B27 + 10 µM Y27632, and plated on dishes and chamber slides coated with poly-ornithine and laminin [18]. Every 4–5 days, half the medium was changed to fresh Neurobasal + 2% B27 + 2mM GlutaMax (Invitrogen). Colonies that were clearly nonneuronal were removed by scraping, rinsing and refeeding the cultures. After 40 days of differentiation, the cells were rinsed with PBS and harvested with ice-cold Trizol or fixed with 4% para-formaldehyde + 4% sucrose in PBS.

Successful differentiation of iPSCs into cortical neuronal cultures was confirmed by immunofluorescence staining for markers of cortical layers V (transcription factor Ctip2, encoded by BCL11B) and VI (transcription factor TBR1), as well as for glutamatergic markers vGLUT1 (encoded by SLC17A7) and GRIK2 (Fig 1).

Immunofluorescence staining

For live cell staining, StemGent antibodies conjugated to Dylight 488: Tra1-60 or Tra1-81, were applied to iPSC cultures for 30–60 min according the manufacturer's instructions, and the stained colonies were examined with an Evos FL microscope.

For staining of fixed cells, cells were grown on chamber slides (LabTek II CC2), rinsed with Dulbecco's PBS, fixed 15 min in 4% p-formaldehyde/4% sucrose in PBS, permeabilized 7 min in 0.25% Triton X-100 in PBS, and blocked 1–2 hrs with 5% sheep or goat serum in PBS. Primary antibodies, applied overnight at 4˚C, were from Abcam: Oct4 (ab19857), Tbr1 (ab31940), Ctip2 (ab18465), GRIK2 (ab53092), vGlut1 (ab72311) and βIII-tubulin (ab7751); from Millipore: MAP2 (polyclonal AB5622), MAP2 (monoclonal MAB3418). The slides were rinsed (5X PBS), secondary antibodies (Life Technologies) were applied for 2 hrs at room
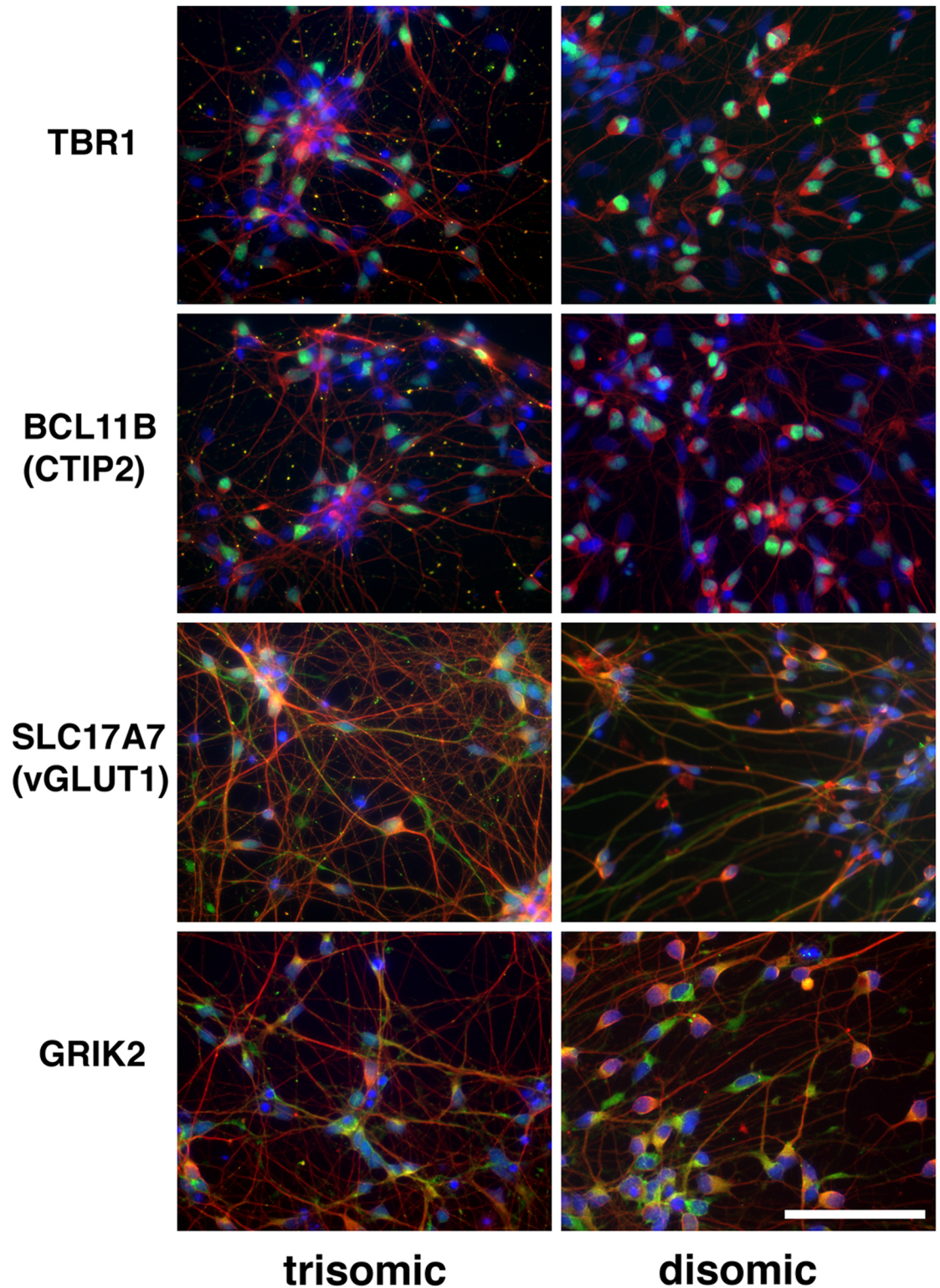
Fig 1. Confirmation of differentiation of iPSCs into cortical neuronal cultures. Images taken from clones C2 (trisomic) and C3 -D21 (disomic) 40 days after initiation of the differentiation protocol. Fixed cells on chamber slides were probed with antibodies against the marker proteins listed in the left column. Neuronal marker Beta III tubulin is red; DAPI is blue; and Ctip2, TBR1, SLC17A7 (vGLUT1) and GRIK2 are green. Size bar = 20 μ.

https://doi.org/10.1371/journal.pone.0194581.g001

temperature, rinsed (4X PBS), stained with DAPI (Sigma D9542), and mounted with ProLong Gold AntiFade (Life Technologies). Images were captured using an epifluorescence microscope (Zeiss Axioskop) equipped with a CCD camera and Slidebook image analysis software (3i). All images comparing trisomic and disomic cells were captured using identical exposures and digital adjustments.

Immunoblotting

Lysates of disomic and trisomic iPSCs were prepared with RIPA buffer containing protease and phosphatase inhibitor cocktails (Sigma), and the protein content of each lysate was determined (Bradford assay). Using BioRad reagents and equipment, an equal amount of protein from each lysate was denatured by boiling 5 minutes in 4x Sample Buffer containing DTT; run on a Criterion XT pre-cast gradient gel (4%-12%) in XT-MOPS running buffer at 150 volts for 1.5–2 hours; and blotted via a Trans-Blot Turbo system for 7 minutes using the Mixed Molecular Weight protocol (2.5A, up to 25V). The blot was washed in TBS-T, (tris-buffered saline containing Tween 20), stained with Ponceau S to affirm protein transfer, washed again and blocked in 5% milk in TBS-T before being stained sequentially with MAP2 rabbit polyclonal antibody (1:1000, Abcam ab32454) and goat anti-rabbit HRP (1:1000). The washed blot was then treated with SuperSignal West Femto Maximum Sensitivity substrate (Life Technologies), diluted 1:10 in distilled water, for 5 minutes while covered, and images were taken using a ChemiDoc-it imaging system and VisionWorks LS software. The blot was stripped and stained for GAPDH (1:2000) as a loading control. Images were analyzed using Fiji (ImageJ) and Excel; integrated density of each band was measured and normalized to the GAPDH signal before comparing the amount of MAP2 in each protein sample.

RNA isolation, cDNA library preparation, high-throughput sequencing

RNA for poly(A) RNA sequencing libraries was extracted from IPSC and neuronal cultures via TRIzol extraction. Genomic DNA was removed using TURBO DNase (Invitrogen). RNA integrity (RIN) was verified on an Agilent 2100 bioanalyzer and only samples with RIN ≥ 7.0 were used for sequencing. Single-end sequencing libraries were prepared by the UCCC Genomics Core (Aurora, CO) using the TruSeq Stranded mRNA Library Prep Kit (Illumina). Cluster generation and sequencing were performed on the Illumina HiSeq 2000 platform. The reads were de-multiplexed and converted to FASTQ format using CASAVA software from Illumina. Read numbers per sample ranged from 78,091,984–158,679,037, and uniquely aligned reads ranged from 56,727,581 to 100,169,823.

Analysis of RNA-seq data

Analysis of library complexity and per-base sequence quality (i.e. q > 30) was assessed using FastQC (v0.11.2) software [19]. Low quality bases (q < 10) were trimmed from the 3' end of reads. Adaptor sequences and reads shorter than 40 nucleotides were removed using Trimmomatic [20]. Reads were aligned to GRCh37/hg19 using TopHat2 (v2.0.14, —b2-very-sensitive —keep-—no-coverage-search—library-type fr-firststrand) [21]. High quality mapped reads (MAPQ > 10) were filtered with SAMtools (v0.1.19) [22]. Gene level counts were obtained using the GRCh37 Ensembl annotation with Rsubread (v 1.18.0, strandSpecific = 2, GTF.featureType = "exon", countMultiMappingReads = FALSE) [23]. Differential expression of genes was determined using DESeq (v1.24.0) software with the options "per-condition", "maximum", and "local" for dispersion calculations [24]. Significance was assigned to genes with a Benjamini-Hochberg FDR less than 10%.

Global RNA editing

We identified A-G RNA editing sites in the 6 (3 disomic and 3 trisomic) RNA-seq cortical neuronal samples using the SPRINT RNA editing software [25]. First, a mapping index was made using the SPRINT prepare command with the Ensembl GRCh38 human genome. We then used trimmed and filtered reads as input for SPRINTS main command with the strand specific parameter set to 0. In brief, SPRINT aligned the reads to the genome two times. The first alignment takes reads that aligned with high mapping quality and produces a list of RNA editing sites. To capture hyper edited regions SPRINT remaps the previously unmapped reads to a reference genome with all A (adenosines) masked by G (guanines) and produces a list of hyper edited sites. We combined the list from both mapping steps and required more than two A-G RNA edits per site. An edit ratio was then determined for each location by summing the total number of edits (A to G changes) divided by the total possible edits (read depth at edit location). A minimum read depth of 10 was required for the edit ratio to be determined. The global editing percentage for each sample was then ascertained by averaging all edit ratios considered. A Student's two-tailed paired T-test was used to determine significance (P value < 0.05).

Alternative splicing analysis

Splicing analysis was performed using Hartley's QoRTs / JunctionSeq pipeline [26]. Tophat aligned reads and an Ensembl annotation file (GRCh37) were used as input to generate gene counts with QoRTs software using the options "—stranded", "—singleEnded", and "—min-MAPQ 50". A flat annotation file containing known splice junctions was produced in QoRTs using the same Ensembl annotation above with the "—stranded" option set. The QoRTs generated files were both used as input into JunctionSeq for differential exon and splice junction analysis. Exons or splice junctions were considered differentially expressed in the disomic vs. trisomic comparison if the adjusted P value was less than 10%.

Semi-quantitative PCR

To capture the exclusion of APOO exon 4 in trisomic samples we designed primers bordering exon 3 and 5 (FWD 5' `TCTCACAGCTCCGACACTAT` 3', REV 5' `GAGTCCAATAAGGCCA GCAA` 3'). For the APOO loading control, primers bordering exon 6 and 7 were used (FWD 5' `GCCTCCCTCTATTATCCACAAC` 3', REV 5' `TCTCCTTCCCAAATCTTCTATGAC` 3'). 30ng/μl of cDNA made from a mix of random hexamers, oligo dt, and Superscript IV reverse transcriptase (Invitrogen,Carlsbad, CA, USA) were used as input for all samples. PCR cycling conditions were set as follows: denaturation for 5 min at 95 ˚C, followed by 40 cycles of denaturation for 30 s at 94 ˚C, annealing for 30 s at 54 ˚C, and polymerization for 30s at 72 ˚C, and a final extension for 10 min at 72 ˚C. APOO exon 4 exclusion was resolved on 2% agarose gel.

Quantitative PCR

To further confirm splicing changes in APOO we performed quantitative PCR on the remaining cDNA from IPSC disomic samples C244, C243, and IPSC trisomic sample C2B. Following standard protocols, quantitative real-time PCRs (qRT-PCR) were conducted using PerfeCTa SYBR Green SuperMix, Low ROX (Quanta Biosciences, Gaithersburg, MD, USA) for all samples in technical triplicates. Primers used to capture the trisomic exon inclusion were: FWD 5' `CCTGAGGGTCAATCGAAGTATG` 3' and REV 5' `TGACACCAGGTTGTGTATGG` 3'. For the control, primers bordering APOO exon 6 and 7 were used (FWD 5' `GCCTCCCTCTA TTATCCACAAC` 3', REV 5' `TCTCCTTCCCAAATCTTCTATGAC` 3'). qRT-PCRs were run

in an ABI 7500 Real-Time PCR System (Applied Biosystems, Foster City, CA, USA). Relative quantification was determined using the ΔΔCt method.

Expression of repetitive elements

RepEnrich software was used to generate a count table containing uniquely aligned reads assigned to repetitive elements [27]. An hg19 repeatmasker file was used to build the repetitive element annotation. Size factors generated from the DESeq gene expression analysis were used to normalize the samples. DESeq (v1.24.0) was used to determine differential expression of repetitive element transcripts. Significance was assigned to genes with an FDR less than 10%.

Cluster analysis of datasets

The full datasets of normalized transcript accumulation values for the iPSC studies by Weick et al [13], Jiang et al [14], and LeTourneau et al [15] were downloaded from the Gene Expression Omnibus (GSE48611, GSE47014, and GSE55504, respectively). Trisomic/disomic ratios were calculated using the averaged transcript levels for all genes assayed in each study. Ratios were imported into Cluster 3.0 [28] for genes identified as significantly differentially expressed in our study, and present in 6/7 of the studies being compared. The ratios were log-transformed, and hierarchical clustering was performed using centroid linkage and centered correlation for the similarity matrix. The clustering output was visualized using Java TreeView [29].

All raw sequencing data has been deposited in GEO (GSE101942).

Results

Generation of RNA-seq data

Paired trisomic/disomic iPSCs were obtained from the Russell lab and maintained in feeder-free media. Cortical neuronal cultures were derived from the iPSCs following the procedure of Espuny-Camacho et al [16] with minor modifications (see Materials and methods) and harvested after 40 days in culture. All sequence data was generated from polyA-selected, strand-specific sequencing libraries by Illumina sequencing (100–115 bp single end reads). The iPSC data was generated from 3 independent preps of trisomic clone C2 (harvested at passages 3, 5, and 12 from our receipt of the cells), and 3 independent preps of disomic sub-clones derived from the C2 lineage (two C2-4-4 and one C2-4-3, harvested at passages 3, 5, and 10, respectively, from our receipt of the cells). Our analysis of the transcriptomes of cortical neurons compared 3 biologically independent sets of differentiated cortical neurons derived from the trisomic C2 line and from disomic C3 iPSC cells (C3-D21, which arose spontaneously from trisomic clone C3). Differentially expressed transcripts were identified using the DESeq algorithm. Principal component analysis demonstrated that the 12 transcriptome datasets partitioned as expected (Fig 2).

Gene expression in trisomic cells

Using a false discovery rate (FDR) of 10%, we identified 1644 transcripts (810 up, 834 down) with differential expression between trisomic and disomic iPSCs (see S1 Table). Gene ontology analysis revealed that trisomic iPSCs had increased transcript levels for genes involved in neurogenesis and neuronal function (Table 1). In contrast, genes with decreased transcripts in the trisomic iPSCs were heavily over-represented in cell adhesion function and germ layer/mesoderm development. The germ layer category includes genes well-established to function in maintenance of stem cell fates (e.g., KLF4, NODAL, LEFTY1 and 2, etc).
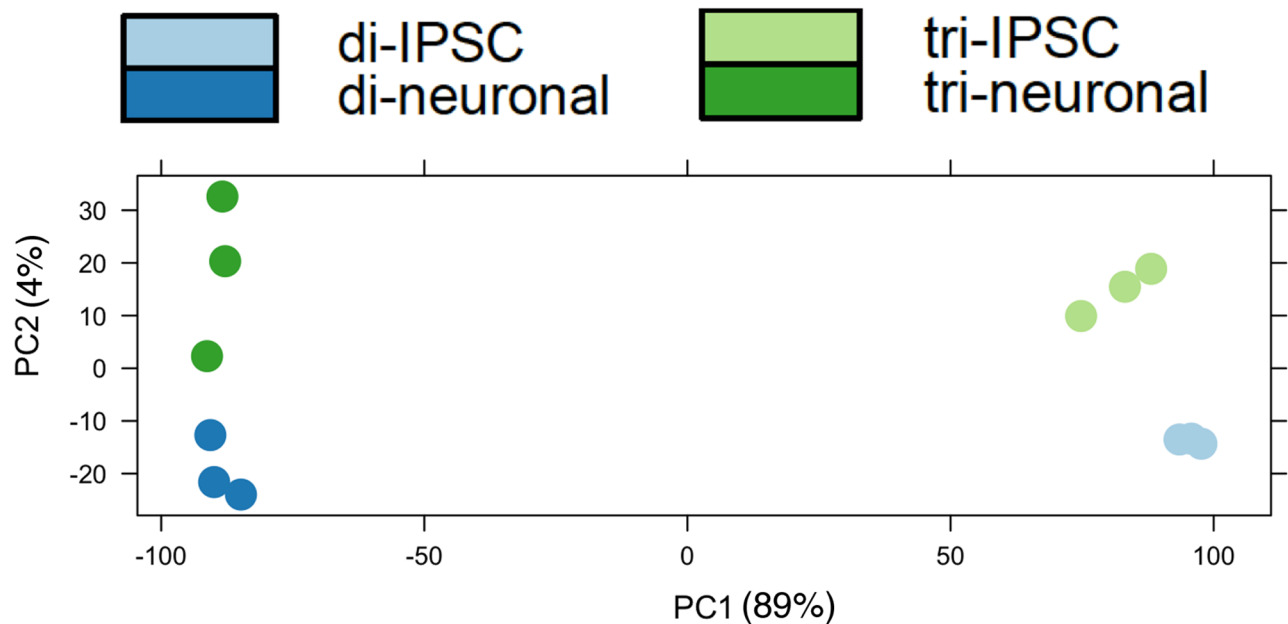
Fig 2. Principal component analysis of transcriptome datasets. Note principal component 1 clearly segregates iPSC from neuronal cultures, while component 2 partitions trisomic from disomic cells.

To confirm the surprising observation that the trisomic iPSCs appeared to express neuronal transcripts at higher levels than the paired disomic iPSCs, we performed quantitative RT-PCR on a set of neuronal-specific transcripts that showed significant upregulation. Specifically, we chose neurotransmitter receptors (GRIN3A, glutamate receptor; GABRA2, GABA receptor), a neuronal cytoskeletal regulator (STMN2, stathmin 2), and two neuronal transmembrane proteins (UNC5C, netrin receptor; LRFN2, PSD95-binding protein). As shown in Fig 3A, 4/5 of these genes had transcript levels > 5 fold higher in trisomic iPSCs than in paired disomic iPSCs.

To further confirm the more neuronal character of the trisomic iPSCs, we focused on MAP2, a classic neuronal-specific microtubule-binding protein whose transcript was increased an average of 6 fold in the trisomic iPSCs compared to the disomic iPSCs. As shown in Fig 2, MAP2 protein was also significantly increased in the trisomic iPSCs as assayed by immunofluorescence (Fig 3B) or by immunoblot (Fig 3C and 3D). Note that the trisomic iPSC clone shown still strongly expressed pluripotency marker OCT4, despite the increased expression of MAP2.

Letourneau et al [15] claimed to identify gene expression dysregulation domains (GEDDs) in trisomic iPSCs and fibroblasts. We applied the same algorithmic approach to our iPSC data, and either did not see the same patterns of expression domains (see S1 Fig) or did not detect statistically significant expression domains.

Using the same parameters applied in the iPSC comparison to the RNA-seq data from the neuronal cultures, we identified 1,377 transcripts (212 up in trisomic cells, 1165 down) differentially expressed in the trisomic and disomic neuronal cultures (See S2 Table). Gene ontology analysis of differentially expressed transcripts (Table 1) identified similar functional categories as observed in the precursor iPSCs, even though there was not a perfect correlation of trisomic/disomic gene expression ratios for individual genes when comparing the iPSC to the differentiated neuronal cultures. For example, while GABRA2 was significantly overexpressed in

Table 1. Gene ontology enrichment using DAVID (https://david.ncifcrf.gov/).

Up in trisomic iPSCs				Down in trisomic iPSCs			
Functional cluster #1	**Fold Enrichment**	**PValue**	**FDR**	**Functional cluster #1**	**Fold Enrichment**	**P Value**	**FDR**
transmission of nerve impulse	3.96	8.90E-16	1.54E-12	cell adhesion	2.75	2.03E-12	3.55E-09
synaptic transmission	3.88	6.27E-13	1.09E-09	biological adhesion	2.74	2.12E-12	3.71E-09
cell-cell signaling	2.84	8.56E-13	1.49E-09	homophilic celladhesion	5.38	6.15E-10	1.08E-06
neurological system process	1.74	2.92E-06	0.005092592	cell-cell adhesion	3.14	5.10E-07	8.92E-04
Functinal cluster #2				**Functinal cluster #2**			
regulation of system process	3.37	5.80E-10	1.01E-06	gastrulation	6.87	2.54E-08	4.44E-05
regulation o ftransmission of nerve impulse	4.52	6.31E-09	1.10E-05	formation of primary germ layer	8.82	2.81E-07	4.92E-04
regulation of neurological system process	4.35	1.35E-08	2.34E-05	embryonic morphogenesis	3.03	3.60E-07	6.29E-04
regulation of synaptic transmission	4.46	4.20E-08	7.32E-05	pattern specification process	3.12	9.48E-07	0.001657046
regulation of synaptic plasticity	3.61	0.006351579	10.5118869	mesoderm morphogenesis	8.44	1.84E-06	0.003213325
				mesoderm development	5.63	2.70E-06	0.0047194
Functinal cluster #3				mesoderm formation	8.01	1.16E-05	0.02024587
neuron differentiation	3.17	3.89E-12	6.77E-09	anterior/posterior pattern formation	3.66	3.04E-05	0.053073992
neuron development	3.15	1.93E-09	3.37E-06	regionalization	3.09	4.39E-05	0.07681769
cell morphogenesis involved in differentiation	3.55	6.04E-09	1.05E-05	tissue morphogenesis	3.03	1.60E-04	0.278911765
cell morphogenesis involved in neuron differentiation	3.60	6.05E-08	1.06E-04	tube morphogenesis	3.03	0.002034975	3.499102339
cell morphogenesis	2.84	7.80E-08	1.36E-04	embryonic development ending in birth or egg hatching	1.92	0.008574202	13.97796428
neuron projection morphogenesis	3.53	8.80E-08	1.53E-04	chordate embryonic development	1.84	0.015827991	24.3434896
cell projection morphogenesis	3.30	9.94E-08	1.73E-04				
cell part morphogenesis	3.16	2.47E-07	4.31E-04	**Functional cluster #3**			
neuron projection development	3.16	2.47E-07	4.31E-04	smooth muscle contraction	8.74	5.77E-06	0.010082818
axonogenesis	3.44	8.72E-07	0.001520019	muscle contraction	3.14	3.01E-04	0.525109932
cellular component morphogenesis	2.55	1.02E-06	0.001771102	muscle system process	2.86	7.74E-04	1.344891904
cell projection organization	2.51	4.34E-06	0.007560515				
axon guidance	3.51	3.15E-04	0.547811961				
cell motion	1.95	5.13E-04	0.891166229				
Up in trisomic neuronal culture				**Down in trisomic neuronal culture**			
Functional cluster #1	**Fold Enrichment**	**PValue**	**FDR**	**Functional cluster #1**	**Fold Enrichment**	**PValue**	**FDR**
gamma-aminobutyric acid signaling pathway	29.41	3.09E-04	0.481161011	celadhesion	2.56	6.45E-18	1.17E-14
neurological system process	1.94	0.014631542	20.55054889	biologicaadhesion	2.56	7.10E-18	1.29E-14
synaptic transmission	3.45	0.01518596	21.24538095	cell-celadhesion	2.34	5.88E-06	0.010659657
transmission of nerve impulse	2.94	0.030529405	38.36333371				
cell-cell signaling	2.21	0.048634045	54.07387315	**Functional cluster #2**			
				bloovessedevelopment	3.24	2.11E-11	3.82E-08
Functional cluster #2				vasculaturdevelopment	3.17	4.71E-11	8.53E-08
cell adhesion	2.31	0.018959843	25.82606353	angiogenesis	4.00	5.25E-11	9.52E-08

(Continued)

Table 1. (Continued)

biological adhesion	2.31	0.019129339	26.02582494	bloovessemorphogenesis	3.24	6.71E-10	1.22E-06
cell-cell adhesion	2.66	0.115339039	85.23209769				
				Functional cluster #3			
Functional cluster #3				extracellular matrix organization	4.44	1.01E-09	1.82E-06
mechanical stimulus involved in sensory perception	33.93	0.003326548	5.067655192	collagen fibril organization	7.65	1.50E-07	2.72E-04
detection of mechanical stimulus	27.57	0.005051111	7.599250716	extracellular structure organization	3.06	6.24E-07	0.001130422

the trisomic iPSCs compared to disomic iPSCs, there was not a statistically significant difference in GABRA2 levels between the trisomic and disomic neuronal cultures where the overall expression of GABRA2 is much higher than in the iPSCs. Nevertheless, gene ontology analysis indicated that, overall, transcripts involved in the GABA signaling pathway had increased levels in the trisomic differentiated neuronal cultures (Table 1).

An important caveat of our analysis is that although transcripts for general neuronal markers (e.g., MAP2, βIII tubulin, and SNAP25) were not increased in the trisomic neuronal cultures compared to the disomic neuronal cultures, cortical neuronal differentiation leads to the production of multiple glial and neuronal cell types, and we do not have sufficient data to demonstrate that all cultures were equivalently differentiated. Thus, our RNA-seq data from the differentiated cultures can best be interpreted as demonstrating that the trisomic iPSCs were not limited in their ability to differentiate into cortical neuronal cultures.

In both the iPSC and neuronal trisomic cultures, chromosome 21 transcripts were increased overall, at close to the 1.5 fold increase predicted from a simple gene dosage effect (mean fold change for chromosome 21 genes, 1.47, median fold change, 1.38). The specific chromosome 21 genes over-expressed were dependent on whether iPSCs or cortical neuronal cultures were being analyzed (Fig 4).

Adenosine-to-inosine (A-to-I) editing in trisomic cells

ADARB1 encodes one of the two enzymes responsible for post-transcriptional A-to-I editing of RNA. This gene is encoded on chromosome 21, leading to suggestions that excessive RNA editing may occur in trisomy 21 cells. To test this hypothesis, we used SPRINT software to globally identify RNA editing sites in transcriptome data from trisomic and disomic differentiated cortical neuronal cultures. (We did not score editing in the iPSCs, as A-to-I editing happens preferentially in neuronal cells, and in fact the iPSCs had very low levels of ADARB1 transcripts.) We identified and scored >250,000 editing sites in both the trisomic and disomic transcriptomes, and quantified fraction of editing across all these sites. Although ADARB1 transcripts appear to be increased in the trisomic neuronal cultures, we did not observe a significant increase in editing levels in the trisomic cultures (Fig 5). We conclude that increased levels of ADARB1 transcripts do not lead to increases in global RNA editing. Our dataset does not have sufficient read depth to determine rigorously if there were significant changes in site-specific editing in the trisomic neuronal cultures.

Splicing alterations in trisomic cells

To identify differences in splicing between trisomic and disomic cells, we tested multiple splicing algorithms, and chose JunctionSeq (see Methods). Using this algorithm we identified 117 annotated genes with splicing changes when comparing trisomic and disomic iPSCs, and 36
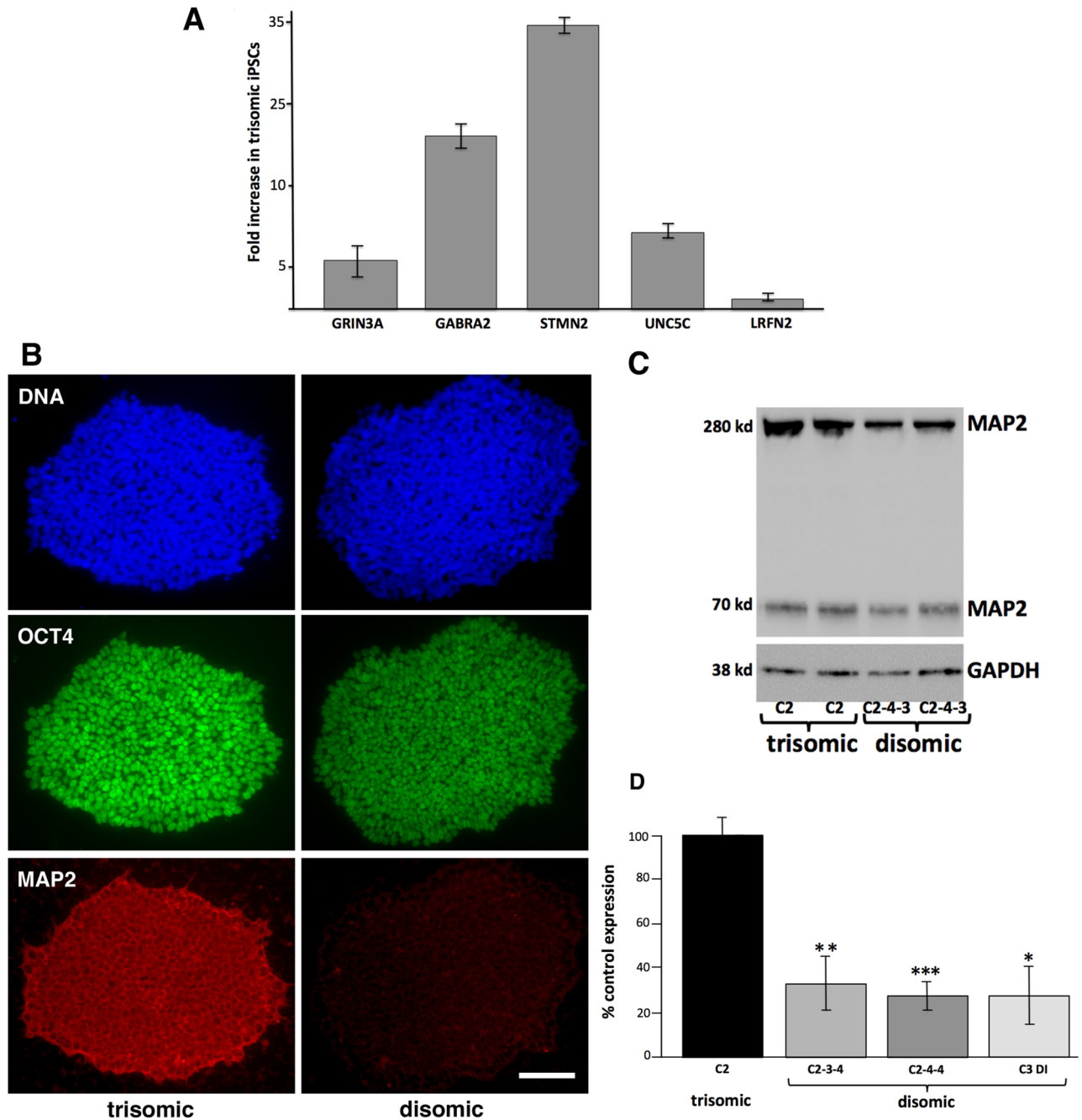
Fig 3. Trisomic iPSCs overexpress neuronal markers. A. RT-PCR quantitation of transcripts overexpressed in trisomic iPSCs relative to disomic iPSCs. cDNA was prepared from the same RNA used for the original sequencing libraries, specifically early passage trisomic iPSC culture (C2B) and an early passage disomic clone (C2-4-3). Similar expression levels were observed using RNA from independent trisomic (C2C) and disomic (C2-4-3) clones. **B**. iPSC colonies (trisomic, C2B; disomic C2-4-4) fixed and stained for MAP2 and OCT4. **C**. Representative immunoblot from trisomic and disomic iPSCs. Note two MAP2 isoforms are detected. **D**. Quantitation of MAP2 immunoblots. (* = P<0.02, ** = P< 0.01, ***, P< 0.001, paired T-tests).

https://doi.org/10.1371/journal.pone.0194581.g003

such genes when comparing the derived cortical neuron cultures (using a conservative adjusted P value < 0.01). (See S3 Table) Only one gene, SLC38A2, appeared to have altered splicing in both the iPSCs and cortical neuronal cultures, perhaps not surprising given the large transcriptional and splicing differences between stem cells and neurons (e.g., note large
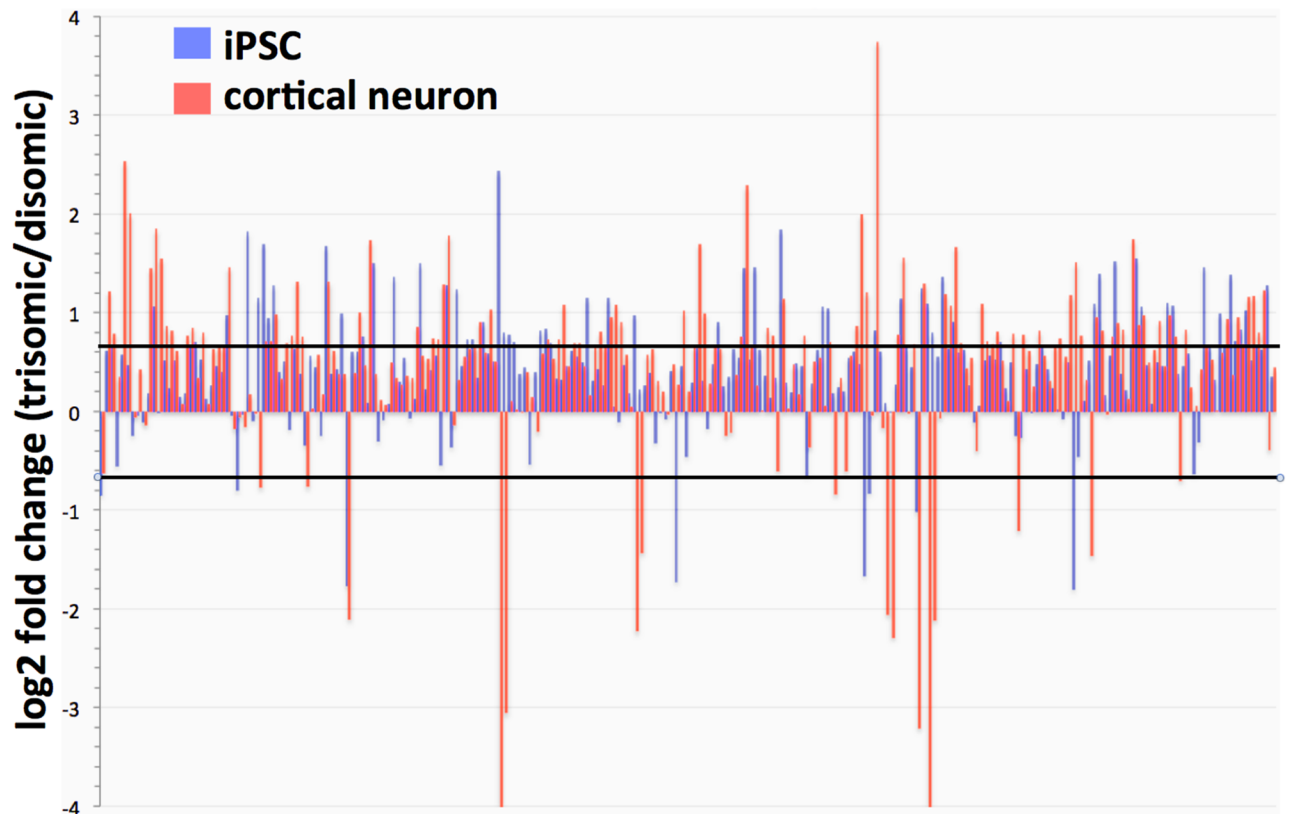
Fig 4. Expression of Chr 21 genes in iPSCs and cortical neuronal cultures (trisomic/disomic ratios). Lines represent individual gene ratios across chromosome 21. Note strong general trend for Chr 21 genes to be over-expressed, although specific genes may be strongly up-or-down regulated depending on cell type. (Black lines indicate 1.5 fold increases or decreases in trisomic cells).

https://doi.org/10.1371/journal.pone.0194581.g004

separation of iPSCs and neurons in Principal Component 1 in Fig 2). To verify this bioinformatic analysis, we performed semi-quantitative RT-PCR on Apolipoprotein O gene (APOO), a gene identified as having altered splicing between the trisomic and disomic iPSCs (Fig 6). This analysis confirmed that trisomic cells show increased exclusion of exon 14 of APOO.

Accumulation of repetitive element transcripts in trisomic cells

Retrotransposons, one class of repetitive elements, appear to play an important role in the maintenance of stem cell pluripotency [30, 31]. To determine whether chromosome 21 trisomy might impact the accumulation of repetitive element (RE) transcripts globally, we used the RepEnrich algorithm [27] to compare RE transcript levels in both the iPSC and differentiated neuronal trisomic/disomic culture pairs (see S4 Table). In the iPSC comparison, there was a trend towards decreased RE transcript accumulation in the trisomic cells, with 26/38 differentially expressed RE transcripts (FDR <10%) lower in the trisomic cells. Interestingly, the most highly expressed RE transcript in the iPSCs was HERVH-int, which also showed the most (statistically) significant difference between the trisomic and disomic iPSCs: a > 3 fold reduction in the trisomic cells. There was no significant difference in HERVH-int expression in cortical neuronal cultures, and the overall trend in RE expression appeared reversed, with 18/21 significantly different RE transcripts increased in the trisomic cells.
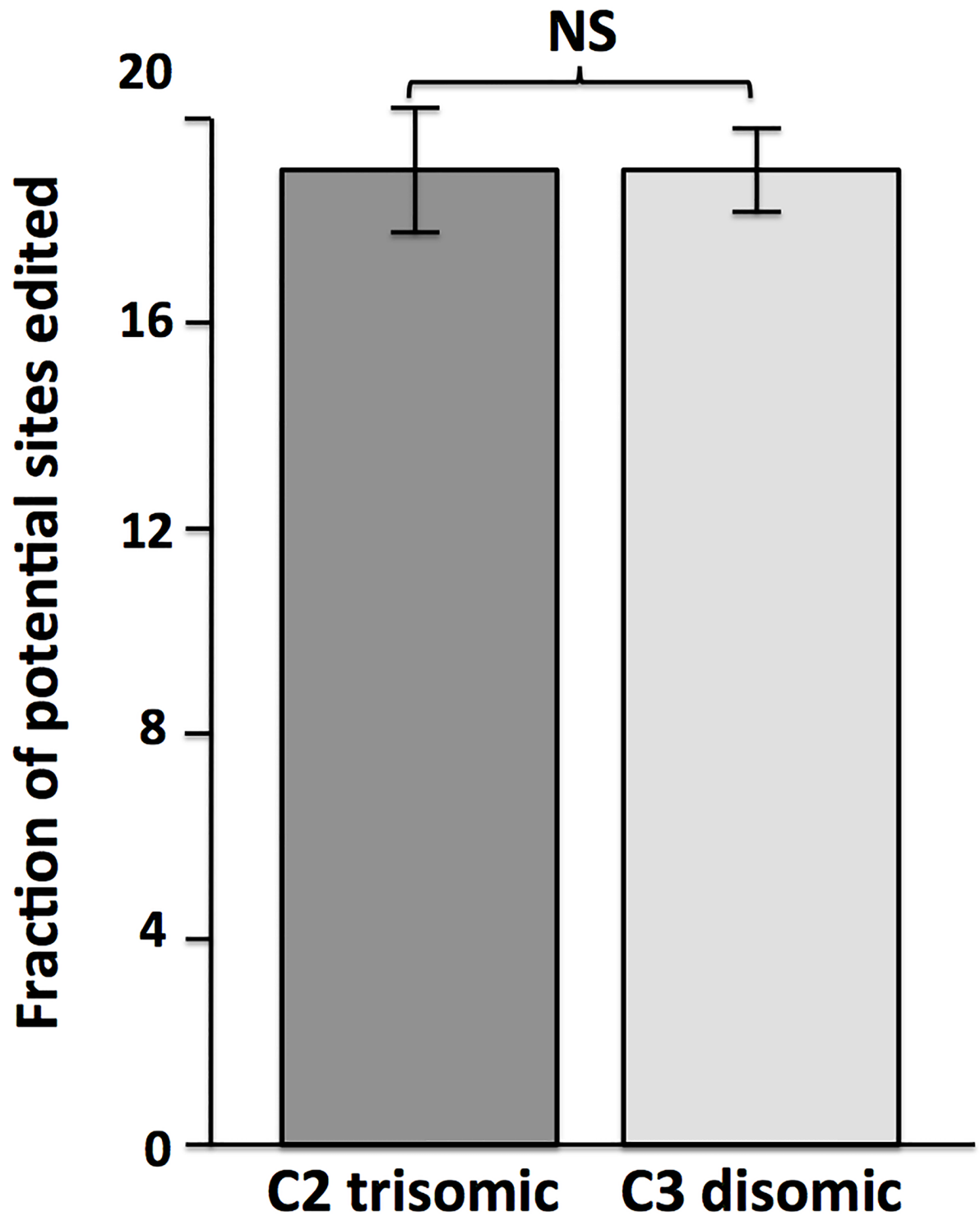
Fig 5. RNA editing levels do not differ between trisomic and disomic cortical neuronal cultures. SPRINT software [25], which employs clustering of editing sites to distinguish *bona fide* RNA editing from DNA polymorphisms and sequencing errors, was used to quantify overall RNA editing levels. Shown are percent overall editing (reads containing edits/potential editing sites) calculated from transcriptome data from biological replicates of trisomic (C2) and disomic (C3 Di) cortical neuronal cultures.
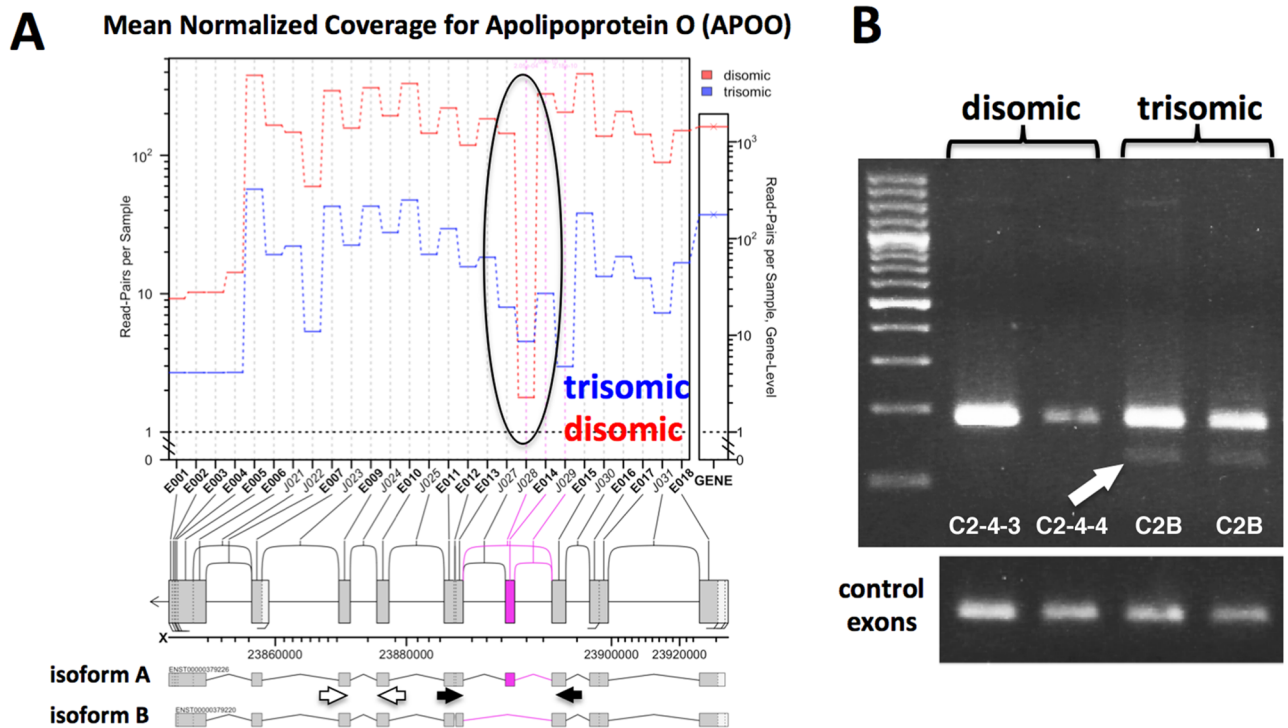
https://doi.org/10.1371/journal.pone.0194581.g005

Fig 6. Altered splicing pattern of APOO in trisomic cells. A. JunctionSeq output supporting the skipping of exon 4 in trisomic iPSCs. Black arrows indicate primers designed to amplify alternatively spliced exon; white arrows indicate control exons 6 and 7. **B**. RT-PCR targeting exon 4 exclusion. Arrow indicates exon exclusion band (band "B") only recovered in trisomic cells. Quantitative RT-PCR was used to show that the included exon is ~ 10 fold enriched in the trisomic samples relative to the unchanged control exons (C2B: C243, 12.7 +/- 1.3 SEM fold, C2B:C244, 9.4 +/- 1.6 SEM fold).

https://doi.org/10.1371/journal.pone.0194581.g006

Comparison to other transcriptome data sets

The iPSCs and induced neuronal cultures that we have characterized were derived from a single individual, thereby intrinsically limiting the generalization of our results. However, three datasets have been published that have attempted a similar transcriptome characterization using trisomic/disomic iPSC pairs that have the same genetic background within each study (Table 2). Weick et al [13] derived iPSCs from fibroblasts from a DS individual who was mosaic for chromosome 21 trisomy, which allowed the derivation of both trisomic and disomic iPSCs. Jiang et al [14] used an elegant, inducible XIST-based system to inhibit expression of one chromosome 21 in trisomic iPSC clones, which also allowed a trisomic/disomic comparison in the same iPSC clone. Letourneau et al [15] identified a rare set of monozygotic twins discordant for chromosome 21 trisomy, and that enabled them to derive a trisomic/disomic iPSC pair. As we have done, these studies all used statistical criteria to generate lists of genes with differential expression in trisomic vs. disomic iPSCs. Comparison of these lists

Table 2. Comparable studies assaying transcriptome differences in isogenic disomic/trisomic iPSCs pairs.

study	isogenic iPSC approach	sex	transforming factors	culture	transcriptome analysis
Li et al, 2012 [12] (used in this study)	genetic selection for Chr21 loss	M	LIN28, NANOG, SOX2, OCT4	feeder-free	RNA-seq
Weick et al, 2013 [13]	fibroblasts from mosaic individual	M	OCT4, SOX2, KLF4, MYC	MEFs	microarray
Jiang et al, 2013 [14]	inducible XIST inactivation of one Chr 21	M	OCT4, SOX2, KLF4, MYC	MEFs	microarray
Letourneau et al, 2014 [15]	fibroblasts from dischordant twin pair	F	OCT4, SOX2, KLF4, MYC	Human foreskin	RNA-seq

https://doi.org/10.1371/journal.pone.0194581.t002

identified no differentially expressed gene common to all the studies, except for genes expressed on chromosome 21. There are multiple, non-mutually exclusive, possible explanations for this observation: inherent variability in independently-generated iPSCs, large effects of the experimental conditions used in individual labs (including differences between chromosome 21 loss and inactivation by XIST targeting), gender differences in donors, insufficient replication to identify all transcriptome differences accurately, strong effects of genetic background, one or more study with outlying data, or a fundamental absence of a non-chromosome 21 transcriptome signature in trisomic iPSCs. To sort through these possibilities, we undertook a cluster analysis to determine how the datasets grouped. In particular, we sought to determine if our dataset was truly an outlier in comparison to the previous studies.

The relevant data from previous studies were downloaded from online repositories, and trisomic/disomic ratios were calculated for all the annotated genes assayed. This approach enabled us to use a dimensionless measure that allowed comparison of RNA-seq and microarray data. Our study and the twin study by Letourneau et al [15] produced single sets of trisomic/disomic ratios. The study by Weick et al [13] generated two sets of ratios, using the data from two trisomic clones (DS1 and DS4) and one disomic clone (DS2). (We calculated the DS1/DS2 and DS4/DS2 ratios to maximize information, while appreciating that these ratios are not independent because they use the same denominator.) The study by Jiang et al [14] examined 3 independent XIST chromosome 21 insertion subclones derived from the same parental trisomic iPSC clone. These subclones were subjected to doxycyline-induced XIST silencing, and trisomic/disomic ratios could be calculated by comparing uninduced/induced transcriptome measurements, thus generating 3 independent ratios (clones 1, 2, 3). The cluster analysis was based on the trisomic/disomic ratios of our set of differentially expressed genes and the calculated ratios for the same genes in the other datasets. After filtering for genes with values in at least 5/6 of the other datasets, 942 genes were used for the clustering analysis (see Materials and Methods).

Fig 7A shows the dendrogram of the calculated dataset relationships. Note that our dataset does not appear to be an outlier, but instead clusters with clone 3 of the study by Jiang et al [14] and with the twin study of Letourneau et al [15]. Fig 7B shows a portion of the gene expression heat map for the clustered datasets, capturing genes that had increased expression (red color) in our dataset (second column in heat map display). Note that these genes include neuron-associated transcripts such as neurexin (NRXN1), neurocan (NCAN), and α-2A adrenoreceptor (ADRA2A), all of which also show increased expression in the trisomic iPSCs characterized in the twin study of Letourneau et al [14] (first column in heat map display). Repeating this analysis after excluding chromosome 21 genes gives exactly the same clustering of datasets, indicating the relationship between datasets is being driven by global gene expression.

Discussion

Transcript accumulation differences between cells trisomic or disomic for chromosome 21

An inherent limitation o human transcriptome studies is accounting for the differing genetic background of individuals. In the context of Down syndrome, this limitation has hindered determining the degree to which trisomy of chromosome 21 alters expression of genes on this chromosome and in the genome as a whole. However, recent technical advances have allowed the derivation of iPSCs trisomic or disomic for chromosome 21 from the same individual. We have used RNA-seq to characterize in depth the transcriptome of one such trisomic/disomic iPSC pair, originally generated by Li and colleagues [12]. We have also differentiated these
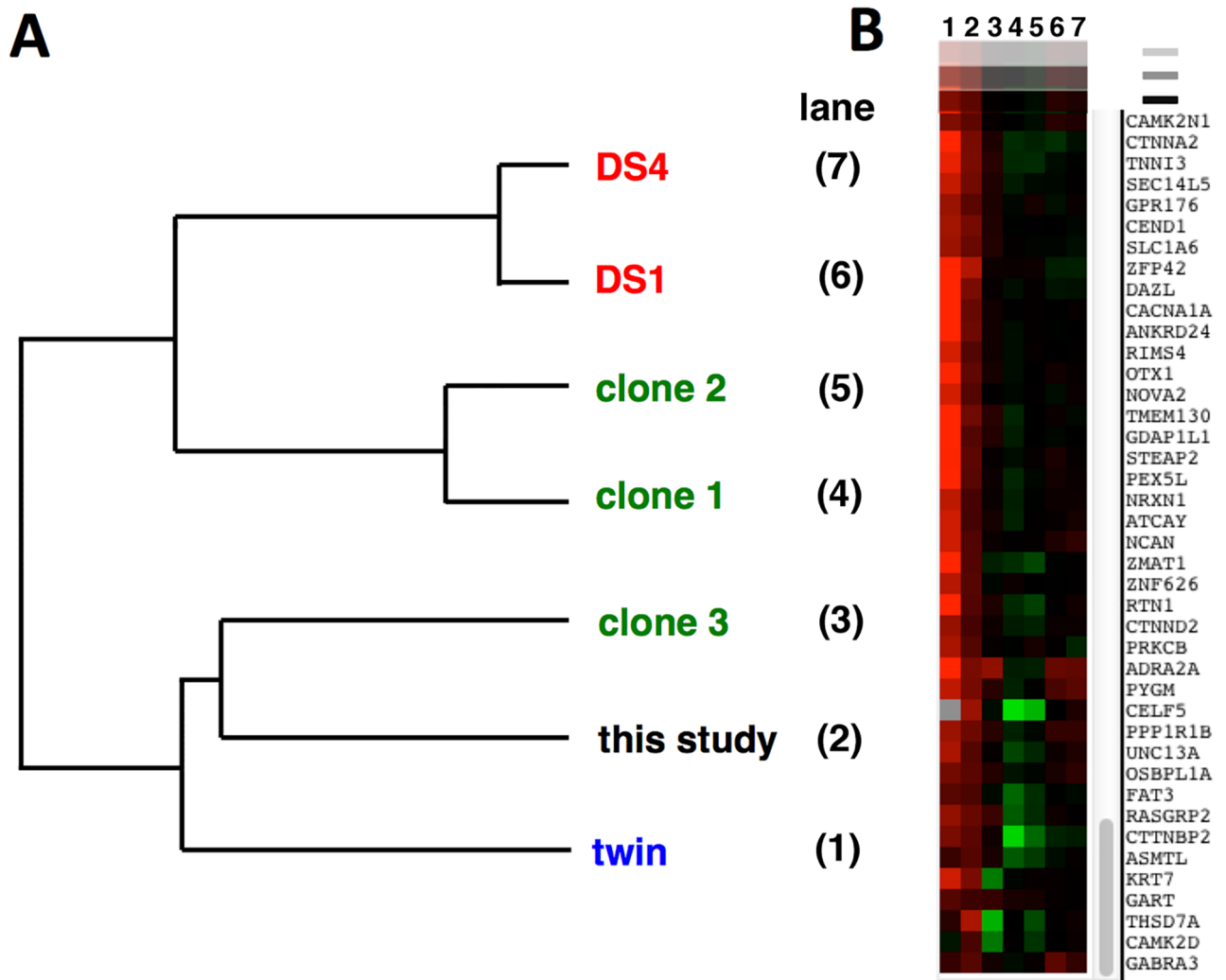
Fig 7. Cluster analysis of comparable studies using isogenic disomic/trisomic cell pairs. A. Dendrogram representation of relatedness of datasets listed in Table 2. DS1 and DS4, from Weick et al. [13]; clones 1,2,3 from Jiang et al [14]; "twin" from Letourneau et al [15]. **B**. Section of heat map display showing expression ratios for individual genes calculated for datasets shown in (**A**). Note overall similarity between "twin" data (leftmost column 1) and the data from this study (column 2).

https://doi.org/10.1371/journal.pone.0194581.g007

iPSCs into cortical neuronal cultures to investigate how trisomy of chromosome 21 impacts the transcriptome in differentiated neurons. Our results indicate that an increase in chromosome 21 dosage does generally lead to increased transcript accumulation for chromosome 21 genes. However, not all chromosome 21 genes follow this pattern, and which genes show increased expression depends on the cell type assayed (iPSCs vs. differentiated cortical neurons). This observation may not be surprising, given how regulatory loops, compensatory changes, and cell-specific transcriptional programs may modulate or override the dosage effect on any given chromosome 21 gene. In this regard, it has recently been reported [32] that the interferon response is up-regulated in trisomy 21 fibroblasts and lymphoblast lines, likely due to increased expression in these cells of the four interferon receptors (IFNAR1, IFNAR2, IFNGR2, and IL10RB) located on chromosome 21. We find that these receptors also have increased transcript accumulation in the trisomic iPSCs we analyzed, but this is not apparent in the derived neuronal cultures.

As observed in other studies, the majority of significant transcriptome changes we identified in both iPSCs and derived neuronal cultures occur in non-chromosome 21 genes. Gene ontology analysis of genes with transcript abundance changes in trisomic iPSCs unexpectedly revealed increases in transcripts from neuronal-associated genes and decreases in transcripts from genes involved in cell adhesion function and germ layer/mesoderm development. These differences did not result from visible spontaneous differentiation in the trisomic iPSCs, which were monitored daily.

One significant consideration in comparing the trisomic and disomic iPSCs used in this study is that the disomic clones have gone through two cycles of selection: first with G418 to select for the insertion of a dual selection cassette into exon 3 of the *APP* gene, and then with gancyclovir to select for loss of the chromosome 21 bearing the inserted selection cassette. It is possible that these rounds of selection resulted in epigenetic (or, less likely, genetic) changes that added to the transcriptional differences between the trisomic and disomic iPSCs. In fact, we were unable to efficiently differentiate into cortical neuronal cultures any of the 4 disomic clones that had gone through these selection procedures. As the differentiation protocol we used has been employed by other research groups to produce cortical neuron cultures from iPSCs with normal karyotypes, we suspect that the derivation of disomic clones may have resulted in a subtle intrinsic change in the iPSC state. Fortunately, we were able to derive a spontaneous disomic line from the C3 trisomic iPSC clone, which did readily differentiate into cortical neurons. The C3 disomic iPSCs (C3-D21) were therefore used to generate the disomic cortical neuronal cultures. Importantly, C3-D21 disomic iPSCs expressed less MAP2 than trisomic iPSCs (Fig 4C), and we still observed increased expression of neuronal-associated genes in the C2 trisomic neuronal cultures compared to the disomic neuronal cultures derived from C3-D21. These observations argue that our finding that C2 trisomic cells have a more neuronal character than disomic cells from the same individual cannot be attributed to an artifact of the genetically-selected disomic clones.

The neuronal differentiation protocol we used [16] has been reported to result in the production of layer V and VI cortical neurons after 40 days in culture, which we confirmed by immunofluorescence microscopy using antibodies against layer V and VI—specific transcription factors CTIP2 and TBR1, as well as antibodies to glutamate receptor GRIK2 and glutamate transporter vGLUT1 (Fig 1). While the neuronal cultures derived from the trisomic and disomic iPSCs were not readily distinguishable using any of these markers, we cannot exclude the possibility that the transcriptional differences we observed were due in part to different distributions of glial and neuronal cell populations in the differentiated neuronal cultures.

Transcriptome characteristics not affected by chromosome 21 trisomy

The increased chromosome 21 gene dosage in trisomic cells could conceivably affect global RNA metabolism, as chromosome 21 encodes the RNA editing gene ADARB1 and multiple splicing factors (including U2AF1, SON and SCAF4). Global chromatin effects of chromosome 21 trisomy were also proposed by Letourneau et al [15], who identified gene expression dysregulation domains (GEDDs) in trisomic iPSCs and fibroblasts. We have analyzed our transcriptome data for trisomy-dependent changes in adenosine-to-inosine RNA editing, alternative splicing, repetitive element expression, and chromosomal domains of altered expression. Examining more than 250,000 RNA editing sites we did not find any evidence for global changes in RNA editing in trisomic iPSC-derived neuronal cultures. This analysis does not preclude changes in site-specific editing, which would require significantly deeper sequencing depth to assay accurately. Nevertheless, our data indicate that global editing levels do not correlate with ADARB1 transcript levels, likely due to the multiple factors that control

the accumulation of edited transcripts. Our data also did not reproduce the gene expression domains identified by Letourneau et al [15], and we note that an independent analysis of the Letourneau data has questioned the existence of GEDDs [33].

Altered splicing in trisomic cells

We did observe trisomy-dependent splicing changes in both the iPSCs and neuronal cultures. Our identification of splicing alterations in trisomic cells is not completely novel, as splicing differences in a selected set of genes has been observed in fetal DS tissue [34]. RNA-seq has also been used previously to support altered splicing in DS endothelial progenitor cells, although no confirmatory studies were done [35]. The large majority of splicing changes we identified did not occur in genes located on chromosome 21, suggesting that they were not a direct effect of gene dosage (i.e., higher gene dosage leading to competition for splicing factors and subsequent altered splicing). More likely is the possibility that altered expression or altered activity of splicing factors caused the splicing dysregulation observed in trisomic cells. We did observe increased transcript accumulation for three known or suspected splicing factors (U2AF1, SON, and SCAF4) located on chromosome 21 in the trisomic iPSC and/or neuronal cultures. Alternatively, increased expression of chromosome 21 gene DYRK1A, a kinase known to regulate splicing factors [36], could contribute to the splicing changes we identified. Over-expression of DYRK1A in mice has been reported to mimic Down syndrome splicing aberrations [34]. Direct demonstration of a role of DYRK1A or specific splicing factors in the trisomy-dependent splicing changes we have observed will require additional studies in which the gene dosage or expression levels of these genes are normalized in trisomic backgrounds.

Altered HERVH expression in trisomic iPSCs

Examination of the accumulation of repetitive element transcripts revealed a highly significant decrease in transcripts from the endogenous retrovirus HERVH in the trisomic iPSCs. HERVH transcripts have been specifically identified as a pluripotency marker [37], and disruption of HERVH transcription compromises self-renewal in human embryonic stem cells [38]. The reduced HERVH expression we observe in trisomic iPSCs is consistent with these cells having an altered pluripotent state in comparison to the disomic iPSCs.

Comparison with other datasets

As described above, consideration of the possible effects of the selection procedures used to generate the disomic derivative clones suggests a possible technical reason why we have observed transcriptional differences not reported by other investigators. However, cluster analysis of our dataset with other comparable studies suggests that our results are not anomalous, but may reflect the range of states assumed by human stem cells trisomic for chromosome 21. The dendrogram shown in Fig 7A could be interpreted as showing that there are two classes of trisomic stem cells: one class with a more neuronal character (cells from this study, the twin study of Letourneau et al, and clone 3 from the Jiang study), and one class with a more mesodermal character (clones 1 and 2 from the Jiang study, and both trisomic lines from the Wieck et al study). This is likely an oversimplification; more extensive data might readily reveal a spectrum of trisomic stem cell fates. However, this analysis does suggest that there is unlikely to be a single transcriptome signature for trisomic stem cells, and the observed transcriptome variation may not simply be a result of the varying experimental conditions used by different investigators. This latter point is supported by the observation that clone 3 of the Jiang et al study clustered more closely with our data set than with the other XIST clones. The Jiang study was particularly well controlled, as the iPSCs were grown under the same

conditions (using mouse embryonic fibroblasts as feeder cells) with the same length of XIST-based knockdown of chromosome 21 expression (3 weeks). Nevertheless, clone 3 trisomic iPSCs appeared more similar to ours, which were grown in feeder-free media.

An alternative interpretation of our cluster analysis is that due to an inherent instability/variability of iPSCs in general, the grouping of trisomic transcriptome profiles we have observed reflects a divergence of individual iPSC clones in culture, rather than true trisomic/disomic transcriptional differences. While we cannot exclude this possibility, this seems unlikely to explain the grouping differences in the Jiang study, where the iPSCs were only grown for 3 weeks after XIST-based knockdown of chromosome 21 gene expression.

Chromosome 21 trisomy and stem cell pluripotency

Why might trisomic iPSCs show a range of transcriptome signatures? Given our observation that the trisomic iPSCs we characterized had enhanced accumulation of transcripts associated with differentiated neurons, and reduced expression of the HERVH pluripotency marker, we speculate that trisomic iPSCs may have an inherent inability to maintain a full pluripotent state. Following this assumption, trisomic iPSCs would have a more constrained differentiation capacity than corresponding disomic iPSCs, which could explain their reported defects in neuronal differentiation [14]. On the other hand, in the absence of full pluripotency, trisomic iPSCs might alternatively be biased towards ectodermal fates (e.g., the trisomic iPSCs we have characterized). The specific developmental bias of a given trisomic iPSC might depend on a variety of factors: genetic background, initial culture conditions, pluripotency factor expression, or stochastic variation. If the gene dosage burden imposed by chromosome 21 trisomy generally interferes with the maintenance of stem cell fate, this might account for some Down syndrome clinical phenotypes, including the hematopoietic developmental defects observed in DS [39]. In this regard, it has been reported that the Ts65Dn mice have defects in the self-renewal of hematopoietic stem cells as well as in the expansion of mammary epithelial cells, neural progenitors and fibroblasts, specifically due to the triplication of Usp16 [40].

Supporting information

S1 Fig. Comparison of putative trisomy-dependent gene expression domains identified on chromosome 1 by Letourneau et al (2014) and this study. Lines show regions of increased or decreased gene expression in cells trisomic for chromosome 21 as calculated using the smoothing algorithm employed by Letourneau et al. The red (iPSCs) and black (fibroblast) traces are taken from the Letourneau study; the green trace is from the iPSCs used in this study. Note that there are significantly different expression domains calculated from the different studies (arrows).
(TIF)

S1 Table. Transcript accumulation in iPSCs.
(XLSX)

S2 Table. Transcript accumulation in cortical neuronal cultures.
(XLSX)

S3 Table. Genes with altered splicing in iPSC and cortical neuronal cultures.
(XLSX)

S4 Table. Repetitive element transcripts in iPSCs and differentiated cortical neurons.
(XLSX)

Acknowledgments

We would like to thank Drs L. Li and D. Russell for providing the iPSC cultures.

Author Contributions

Conceptualization: Gretchen H. Stein, Christopher D. Link.

Formal analysis: Patrick K. Gonzales.

Funding acquisition: Gretchen H. Stein, Christopher D. Link.

Investigation: Patrick K. Gonzales, Christine M. Roberts, Virginia Fonte, Connor Jacobsen, Gretchen H. Stein.

Methodology: Patrick K. Gonzales.

Supervision: Gretchen H. Stein, Christopher D. Link.

Writing – original draft: Gretchen H. Stein, Christopher D. Link.

Writing – review & editing: Patrick K. Gonzales, Gretchen H. Stein, Christopher D. Link.

References

1. Antonarakis SE, Lyle R, Chrast R, Scott HS. Differential gene expression studies to explore the molecular pathophysiology of Down syndrome. Brain Res Brain Res Rev. 2001; 36(2–3):265–74. PMID: 11690624.

2. Mittaz L, Scott HS, Rossier C, Seeburg PH, Higuchi M, Antonarakis SE. Cloning of a human RNA editing deaminase (ADARB1) of glutamate receptors that maps to chromosome 21q22.3. Genomics. 1997; 41(2):210–7. https://doi.org/10.1006/geno.1997.4655 PMID: 9143496.

3. Lalioti MD, Gos A, Green MR, Rossier C, Morris MA, Antonarakis SE. The gene for human U2 snRNP auxiliary factor small 35-kDa subunit (U2AF1) maps to the progressive myoclonus epilepsy (EPM1) critical region on chromosome 21q22.3. Genomics. 1996; 33(2):298–300. https://doi.org/10.1006/geno.1996.0196 PMID: 8660980.

4. Shi J, Zhang T, Zhou C, Chohan MO, Gu X, Wegiel J, et al. Increased dosage of Dyrk1A alters alternative splicing factor (ASF)-regulated alternative splicing of tau in Down syndrome. J Biol Chem. 2008; 283(42):28660–9. https://doi.org/10.1074/jbc.M802645200 PMID: 18658135.

5. Gardiner K, Herault Y, Lott IT, Antonarakis SE, Reeves RH, Dierssen M. Down syndrome: from understanding the neurobiology to therapy. J Neurosci. 2010; 30(45):14943–5. https://doi.org/10.1523/JNEUROSCI.3728-10.2010 PMID: 21068296.

6. Liu C, Belichenko PV, Zhang L, Fu D, Kleschevnikov AM, Baldini A, et al. Mouse models for Down syndrome-associated developmental cognitive disabilities. Dev Neurosci. 2011; 33(5):404–13. https://doi.org/10.1159/000329422 PMID: 21865664.

7. Saran NG, Pletcher MT, Natale JE, Cheng Y, Reeves RH. Global disruption of the cerebellar transcriptome in a Down syndrome mouse model. Hum Mol Genet. 2003; 12(16):2013–9. PMID: 12913072.

8. Potier MC, Rivals I, Mercier G, Ettwiller L, Moldrich RX, Laffaire J, et al. Transcriptional disruptions in Down syndrome: a case study in the Ts1Cje mouse cerebellum during post-natal development. J Neurochem. 2006; 97 Suppl 1:104–9. https://doi.org/10.1111/j.1471-4159.2005.03624.x PMID: 16635258.

9. Ling KH, Hewitt CA, Tan KL, Cheah PS, Vidyadaran S, Lai MI, et al. Functional transcriptome analysis of the postnatal brain of the Ts1Cje mouse model for Down syndrome reveals global disruption of interferon-related molecular networks. BMC Genomics. 2014; 15:624. https://doi.org/10.1186/1471-2164-15-624 PMID: 25052193.

10. Guedj F, Pennings JL, Ferres MA, Graham LC, Wick HC, Miczek KA, et al. The fetal brain transcriptome and neonatal behavioral phenotype in the Ts1Cje mouse model of Down syndrome. Am J Med Genet A. 2015; 167A(9):1993–2008. https://doi.org/10.1002/ajmg.a.37156 PMID: 25975229.

11. Olmos-Serrano JL, Kang HJ, Tyler WA, Silbereis JC, Cheng F, Zhu Y, et al. Down Syndrome Developmental Brain Transcriptome Reveals Defective Oligodendrocyte Differentiation and Myelination. Neuron. 2016; 89(6):1208–22. https://doi.org/10.1016/j.neuron.2016.01.042 PMID: 26924435.

12. Li LB, Chang KH, Wang PR, Hirata RK, Papayannopoulou T, Russell DW. Trisomy correction in Down syndrome induced pluripotent stem cells. Cell Stem Cell. 2012; 11(5):615–9. https://doi.org/10.1016/j.stem.2012.08.004 PMID: 23084023.

13. Weick JP, Held DL, Bonadurer GF 3rd, Doers ME, Liu Y, Maguire C, et al. Deficits in human trisomy 21 iPSCs and neurons. Proc Natl Acad Sci U S A. 2013; 110(24):9962–7. https://doi.org/10.1073/pnas.1216575110 PMID: 23716668.

14. Jiang J, Jing Y, Cost GJ, Chiang JC, Kolpa HJ, Cotton AM, et al. Translating dosage compensation to trisomy 21. Nature. 2013; 500(7462):296–300. https://doi.org/10.1038/nature12394 PMID: 23863942.

15. Letourneau A, Santoni FA, Bonilla X, Sailani MR, Gonzalez D, Kind J, et al. Domains of genome-wide gene expression dysregulation in Down's syndrome. Nature. 2014; 508(7496):345–50. https://doi.org/10.1038/nature13200 PMID: 24740065.

16. Espuny-Camacho I, Michelsen KA, Gall D, Linaro D, Hasche A, Bonnefont J, et al. Pyramidal neurons derived from human pluripotent stem cells integrate efficiently into mouse brain circuits in vivo. Neuron. 2013; 77(3):440–56. https://doi.org/10.1016/j.neuron.2012.12.011 PMID: 23395372.

17. Gaspard N, Gaillard A, Vanderhaeghen P. Making cortex in a dish: in vitro corticopoiesis from embryonic stem cells. Cell Cycle. 2009; 8(16):2491–6. https://doi.org/10.4161/cc.8.16.9276 PMID: 19597331.

18. Shi Y, Kirwan P, Smith J, Robinson HP, Livesey FJ. Human cerebral cortex development from pluripotent stem cells to functional excitatory synapses. Nat Neurosci. 2012; 15(3):477–86, S1. https://doi.org/10.1038/nn.3041 PMID: 22306606.

19. Andrews S. FastQC: a quality control tool for high throughput sequence data. (2010). http://www.bioinformatics.babraham.ac.uk/projects/fastqc

20. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014; 30(15):2114–20. https://doi.org/10.1093/bioinformatics/btu170 PMID: 24695404.

21. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013; 14(4):R36. https://doi.org/10.1186/gb-2013-14-4-r36 PMID: 23618408.

22. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25(16):2078–9. https://doi.org/10.1093/bioinformatics/btp352 PMID: 19505943.

23. Liao Y, Smyth GK, Shi W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Res. 2013; 41(10):e108. https://doi.org/10.1093/nar/gkt214 PMID: 23558742.

24. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol. 2010; 11 (10):R106. https://doi.org/10.1186/gb-2010-11-10-r106 PMID: 20979621.

25. Zhang F, Lu Y, Yan S, Xing Q, Tian W. SPRINT: an SNP-free toolkit for identifying RNA editing sites. Bioinformatics. 2017. https://doi.org/10.1093/bioinformatics/btx473 PMID: 29036410.

26. Hartley SW, Mullikin JC. Detection and visualization of differential splicing in RNA-Seq data with JunctionSeq. Nucleic Acids Res. 2016; 44(15):e127. https://doi.org/10.1093/nar/gkw501 PMID: 27257077.

27. Criscione SW, Zhang Y, Thompson W, Sedivy JM, Neretti N. Transcriptional landscape of repetitive elements in normal and cancer human cells. BMC Genomics. 2014; 15:583. https://doi.org/10.1186/1471-2164-15-583 PMID: 25012247.

28. de Hoon MJ, Imoto S, Nolan J, Miyano S. Open source clustering software. Bioinformatics. 2004; 20 (9):1453–4. https://doi.org/10.1093/bioinformatics/bth078 PMID: 14871861.

29. Saldanha AJ. Java Treeview—extensible visualization of microarray data. Bioinformatics. 2004; 20 (17):3246–8. https://doi.org/10.1093/bioinformatics/bth349 PMID: 15180930.

30. Macfarlan TS, Gifford WD, Driscoll S, Lettieri K, Rowe HM, Bonanomi D, et al. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. Nature. 2012; 487(7405):57–63. https://doi.org/10.1038/nature11244 PMID: 22722858.

31. Fort A, Hashimoto K, Yamada D, Salimullah M, Keya CA, Saxena A, et al. Deep transcriptome profiling of mammalian stem cells supports a regulatory role for retrotransposons in pluripotency maintenance. Nat Genet. 2014; 46(6):558–66. https://doi.org/10.1038/ng.2965 PMID: 24777452.

32. Sullivan KD, Lewis HC, Hill AA, Pandey A, Jackson LP, Cabral JM, et al. Trisomy 21 consistently activates the interferon response. Elife. 2016; 5. https://doi.org/10.7554/eLife.16220 PMID: 27472900.

33. Do LH, Mobley WC, Singhal N. Questioned validity of Gene Expression Dysregulated Domains in Down's Syndrome. F1000Res. 2015; 4:269. https://doi.org/10.12688/f1000research.6735.1 PMID: 26664707.

34. Toiber D, Azkona G, Ben-Ari S, Toran N, Soreq H, Dierssen M. Engineering DYRK1A overdosage yields Down syndrome-characteristic cortical splicing aberrations. Neurobiol Dis. 2010; 40(1):348–59. https://doi.org/10.1016/j.nbd.2010.06.011 PMID: 20600907.

35. Costa V, Angelini C, D'Apice L, Mutarelli M, Casamassimi A, Sommese L, et al. Massive-scale RNA-Seq analysis of non ribosomal transcriptome in human trisomy 21. PLoS One. 2011; 6(4):e18493. https://doi.org/10.1371/journal.pone.0018493 PMID: 21533138.

36. Qian W, Liang H, Shi J, Jin N, Grundke-Iqbal I, Iqbal K, et al. Regulation of the alternative splicing of tau exon 10 by SC35 and Dyrk1A. Nucleic Acids Res. 2011; 39(14):6161–71. https://doi.org/10.1093/nar/gkr195 PMID: 21470964.

37. Santoni FA, Guerra J, Luban J. HERV-H RNA is abundant in human embryonic stem cells and a precise marker for pluripotency. Retrovirology. 2012; 9:111. https://doi.org/10.1186/1742-4690-9-111 PMID: 23253934.

38. Wang J, Xie G, Singh M, Ghanbarian AT, Rasko T, Szvetnik A, et al. Primate-specific endogenous ret-rovirus-driven transcription defines naive-like stem cells. Nature. 2014; 516(7531):405–9. https://doi.org/10.1038/nature13804 PMID: 25317556.

39. Roberts I, Izraeli S. Haematopoietic development and leukaemia in Down syndrome. Br J Haematol. 2014; 167(5):587–99. https://doi.org/10.1111/bjh.13096 PMID: 25155832.

40. Adorno M, Sikandar S, Mitra SS, Kuo A, Nicolis Di Robilant B, Haro-Acosta V, et al. Usp16 contributes to somatic stem-cell defects in Down's syndrome. Nature. 2013; 501(7467):380–4. https://doi.org/10.1038/nature12530 PMID: 24025767