

ARTICLE OPEN



TCF7L2 lncRNA: a link between bipolar disorder and body mass index through glucocorticoid signaling

Duan Liu^{1,8}, Thanh Thanh Le Nguyen^{1,2,8}, Huanyao Gao¹, Huaizhi Huang^{1,2}, Daniel C. Kim¹, Brenna Sharp¹, Zhenqing Ye³, Jeong-Heon Lee⁴, Brandon J. Coombes³, Tamas Ordog^{5,6}, Liewei Wang¹, Joanna M. Biernacka^{3,7}, Mark A. Frye⁷✉ and Richard M. Weinshilboum¹✉

© The Author(s) 2021

Bipolar disorder (BD) and obesity are highly comorbid. We previously performed a genome-wide association study (GWAS) for BD risk accounting for the effect of body mass index (BMI), which identified a genome-wide significant single-nucleotide polymorphism (SNP) in the gene encoding the transcription factor 7 like 2 (*TCF7L2*). However, the molecular function of *TCF7L2* in the central nervous system (CNS) and its possible role in the BD and BMI interaction remained unclear. In the present study, we demonstrated by studying human induced pluripotent stem cell (hiPSC)-derived astrocytes, cells that highly express *TCF7L2* in the CNS, that the BD-BMI GWAS risk SNP is associated with glucocorticoid-dependent repression of the expression of a previously uncharacterized *TCF7L2* transcript variant. That transcript is a long non-coding RNA (lncRNA-*TCF7L2*) that is highly expressed in the CNS but not in peripheral tissues such as the liver and pancreas that are involved in metabolism. In astrocytes, knockdown of the lncRNA-*TCF7L2* resulted in decreased expression of the parent gene, *TCF7L2*, as well as alterations in the expression of a series of genes involved in insulin signaling and diabetes. We also studied the function of *TCF7L2* in hiPSC-derived astrocytes by integrating RNA sequencing data after *TCF7L2* knockdown with *TCF7L2* chromatin-immunoprecipitation sequencing (ChIP-seq) data. Those studies showed that *TCF7L2* directly regulated a series of BD risk genes. In summary, these results support the existence of a CNS-based mechanism underlying BD-BMI genetic risk, a mechanism based on a glucocorticoid-dependent expression quantitative trait locus that regulates the expression of a novel *TCF7L2* non-coding transcript.

Molecular Psychiatry (2021) 26:7454–7464; <https://doi.org/10.1038/s41380-021-01274-z>

INTRODUCTION

Bipolar disorder (BD) is a psychiatric disease with significant morbidity and mortality [1]. BD is often comorbid with other disorders such as alcohol use disorder [2–4] and binge eating disorder [5]. The overall heritability of BD is between 60% and 85%, indicating that genetic factors contribute substantially to disease risk [6]. Genome-wide association studies (GWAS) have identified >60 genome-wide significant loci associated with BD risk [6–13]. Pathway enrichment analyses of BD risk genes have revealed several significant gene sets, including those that regulate insulin secretion, energy metabolism [6, 7], and corticotropin-releasing hormone (CRH) signaling [14]. Those results are consistent with clinical observations that BD is often associated with binge eating disorder [5], addiction disorders [2–4, 15] (i.e., alcohol and food), and with metabolic syndromes [16]. The prevalence of obesity in BD patients is more than twofold greater than that for the general population [16, 17]. Furthermore, obesity in BD is associated with more severe mood symptoms and worse treatment outcomes [18, 19]. The pathophysiology underlying the comorbidity of BD and obesity remains unclear.

In an attempt to gain a greater understanding of the potential contribution of genetic factors to the comorbidity of BD and obesity, we performed a GWAS to identify variants associated with body mass index (BMI)-dependent BD [20]. Even though BMI is an imperfect measure of excessive fat accumulation, it is a widely used clinical variable in studies of obesity [21, 22]. That GWAS identified a genome-wide significant single-nucleotide polymorphism (SNP) (rs12772424, $P = 2.85E-08$) that mapped to an intronic region of the *TCF7L2* gene [20]. Specifically, the minor allele for the rs12772424 SNP was more strongly associated with BD as BMI increased [20]. However, this SNP was not associated with risk for either BD or BMI alone, but only showed a significant association with BD when accounting for the effect of BMI [20]. The association of this *TCF7L2* SNP with BMI-dependent BD risk was later replicated in an independent cohort of BD patients [23]. *TCF7L2* is a well-known risk gene for type 2 diabetes (T2D). Specifically, a different *TCF7L2* SNP (rs7903146)—a SNP that is not in linkage disequilibrium (LD) with rs12772424 ($r^2 < 0.035$, $D' < 0.282$) and which maps 122.2 kb distant from rs12772424—has been associated with risk for T2D in a series of studies [24–27].

¹Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic, Rochester, MN, USA. ²Graduate School of Biomedical Sciences, Mayo Clinic, Rochester, MN, USA. ³Division of Computational Biology, Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN, USA. ⁴Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA. ⁵Department of Physiology and Biomedical Engineering, Mayo Clinic, Rochester, MN, USA. ⁶Division of Gastroenterology and Hepatology, Department of Internal Medicine, Mayo Clinic, Rochester, MN, USA. ⁷Department of Psychiatry and Psychology, Mayo Clinic, Rochester, MN, USA. ⁸These authors contributed equally: Duan Liu, Thanh Thanh Le Nguyen. ✉email: MFrye@mayo.edu; Weinshilboum.Richard@mayo.edu

Received: 8 May 2021 Revised: 21 July 2021 Accepted: 19 August 2021

Published online: 17 September 2021

The existence of two independent *TCF7L2* SNPs, each of which is associated with different diseases, suggests that *TCF7L2* might have different functions in T2D and in BD associated with BMI.

TCF7L2 encodes the transcription factor 7 like 2, which plays an important role in the Wnt/ β -catenin signaling pathway [28], a pathway known to be involved in the pharmacotherapy of BD [29]. *TCF7L2* is widely expressed across many human organs and tissues, including the brain [30]. As *TCF7L2* is well recognized as a T2D risk gene, its function in T2D and glucose metabolism has been studied extensively in animal models with results that suggest a complex molecular role for this protein [31]. Specifically, *Tcf7l2* dysfunction in hepatic [32–34] and pancreatic [35, 36] tissue, as well as involvement in the gut-brain axis [37] have all been thought to contribute to T2D risk. *Tcf7l2* function in the central nervous system (CNS) has also been studied in rodents [38–40]. In addition to disrupted systemic glucose homeostasis, abnormalities in nervous system morphology and neurologically based behaviors such as polyphagia and hypoactivity were observed in *Tcf7l2* knockout mice [41, 42]. Despite mounting evidence that *TCF7L2* may link psychiatric disorders and metabolic dysregulation [20, 43], its function in human CNS cell line models, to our knowledge, has not been reported.

In the present study, we set out to characterize the molecular function of *TCF7L2* in the CNS with the goal of understanding its possible role in the comorbidity of BD and obesity. These studies began with a review of human brain single-nucleus and single-cell RNA-seq data for *TCF7L2*, results which showed that it is most highly expressed in astrocytes. Furthermore, an examination of the genomic architecture surrounding the BD-BMI interaction GWAS identified SNP, rs12772424, showed that the variant SNP genotype created a half-palindrome glucocorticoid response element (GRE)—raising the possibility that *TCF7L2* transcription might be regulated by glucocorticoid signaling. After treatment of hiPSC-derived astrocytes with dexamethasone (DEX), a potent synthetic glucocorticoid, expression levels of two *TCF7L2* long non-coding transcripts (lncRNA-TCF7L2) were found to be significantly repressed. One of the lncRNA-TCF7L2 transcripts, namely “T-3”, was predominantly expressed in the human brain but not in peripheral tissues such as the liver or pancreas. As a result, after starting with an SNP that was associated with a BD subtype, the series of experiments described subsequently showed that the SNP was a pharmacogenomic expression quantitative trait locus (PGx-eQTL), i.e., an eQTL that is functional only in the presence of a drug (e.g., DEX) or endogenous hormones such as glucocorticoids [44–46]. Finally, we found that the “T-3” lncRNA-TCF7L2 influenced expression in astrocytes of both the parent gene and scores of other genes involved in both energy metabolism and BD risk. This series of studies identified and functionally characterized a *TCF7L2* lncRNA and its SNP-dependent expression as a molecular mechanism related to BD-BMI interaction and, as a result, illustrates the potential importance of lncRNAs for neuropsychiatry.

MATERIALS AND METHODS

TCF7L2 transcriptional analysis in the human brain at the single-cell level

A single-nucleus RNA sequencing (snRNA-seq) data set generated by the Allen Institute that included samples across multiple human cortical areas [47] was consulted to obtain information with regard to *TCF7L2* expression in CNS cells. In addition, single-cell RNA sequencing (scRNA-seq) data generated using surgically removed human cerebral cortical samples from 12 subjects [48, 49] were downloaded and combined for analysis of *TCF7L2* expression. Finally, chromatin accessibility information for the *TCF7L2* gene in human brain cells was obtained from a single-cell assay for transposase-accessible chromatin using sequencing (scATAC-seq) data set [50] to evaluate the transcriptional activity of *TCF7L2* in clustered CNS cell types. The sources of these data sets and other key resources used in this study

are listed in Supplementary Table S1. Please see the Supplementary Text for details.

Identification of the *TCF7L2* transcript variants regulated by glucocorticoid signaling

The *TCF7L2* rs12772424 SNP was annotated using HaploReg v4.1 [51], which indicated that the rs12772424 variant allele (A) created a half-palindrome GRE. To test the hypothesis that *TCF7L2* transcription might be regulated by glucocorticoid signaling, quantitative reverse transcription-polymerase chain reaction (qRT-PCR) was used to quantitate RNA levels for all known *TCF7L2* transcript variants in a variety of cell lines before and after treatment with the glucocorticoid receptor (GR) agonist DEX. The rs12772424 SNP effect on *TCF7L2* transcriptional activity was also determined by reporter gene assay with the comparison of luciferase activities for rs12772424 SNP wild-type and variant cDNA constructs. See the Supplementary Text for details.

Characterization of *TCF7L2* long non-coding RNA transcripts

Expression levels of *TCF7L2* long non-coding RNA transcripts (lncRNA-TCF7L2) in the human brain and peripheral tissues as well as in human cell lines were quantified by qRT-PCR using primers targeted to their unique exon junctions. The existence of lncRNA-TCF7L2 “T-3” was confirmed by cDNA amplification of its “full length” sequence using total RNA samples prepared from human brain tissues and from hiPSC-derived astrocytes. RNA-seq was also performed with samples from hiPSC-derived astrocytes that have been transfected with ASOs to knockdown (KD) the lncRNA-TCF7L2 or with non-targeting control ASOs. Differentially expressed genes (DEGs) after KD of the lncRNA-TCF7L2 “T-3” were identified by comparison of transcriptomes of control and KD samples. Those DEGs were then used for pathway enrichment analysis using EnrichR [52] to annotate the function of the lncRNA-TCF7L2 “T-3”. See the Supplementary Text for additional details.

Identification of *TCF7L2* target genes in hiPSC-derived astrocytes

RNA-seq and ChIP-seq data were generated to identify *TCF7L2* target genes. Specifically, *TCF7L2* protein expression was knocked down (KD) by siRNA transfection. After the KD efficiency had been validated by both qRT-PCR and western blot assays, RNA-seq was performed with control and KD samples to identify transcriptome-wide DEGs. ChIP-seq was also performed to identify *TCF7L2* genome-wide DNA-binding sites. RNA-seq and ChIP-seq data were integrated as described previously [53] for the identification of *TCF7L2* target genes. See the Supplementary Text for details.

TCF7L2 and BD risk genes

To annotate *TCF7L2* protein function, the RNA-seq and ChIP-seq identified *TCF7L2* target genes in hiPSC-derived astrocytes were subjected to pathway enrichment analysis. Those *TCF7L2* target genes enriched BD as the top disease pathway. Expression levels of the BD risk genes quantified by RNA-seq were plotted, and the *TCF7L2*-binding sites that mapped to or near the BD risk genes identified by ChIP-seq were visualized by using the Integrative Genomics Viewer [54]. See the Supplementary Text for details.

RESULTS

TCF7L2 gene transcription in human brain cells

To determine which CNS cell type(s) might express *TCF7L2*, we first consulted human brain snRNA-seq data generated by the Allen Institute [47]. Their data set included single-nucleus transcriptomes of 49,495 nuclei across multiple human brain cortical areas. Based on marker gene expression, cells were “clustered” into seven non-neuronal and two major neuronal cell types (Fig. 1a). *TCF7L2* was more highly expressed in astrocytes than in any of the other “clustered” cell types (Fig. 1b, c). Previously we had also analyzed scRNA-seq data [55] from two data sets generated from surgically removed human cerebral cortical tissue [48, 49] (Supplementary Fig. S1). In those studies, *TCF7L2* was also expressed predominantly in astrocytes, e.g., 75.2% of astrocytes were *TCF7L2* positive, a percentage nearly twice that of any other CNS cell type (Fig. 1d). The high levels of *TCF7L2* expression in astrocytes were supported by

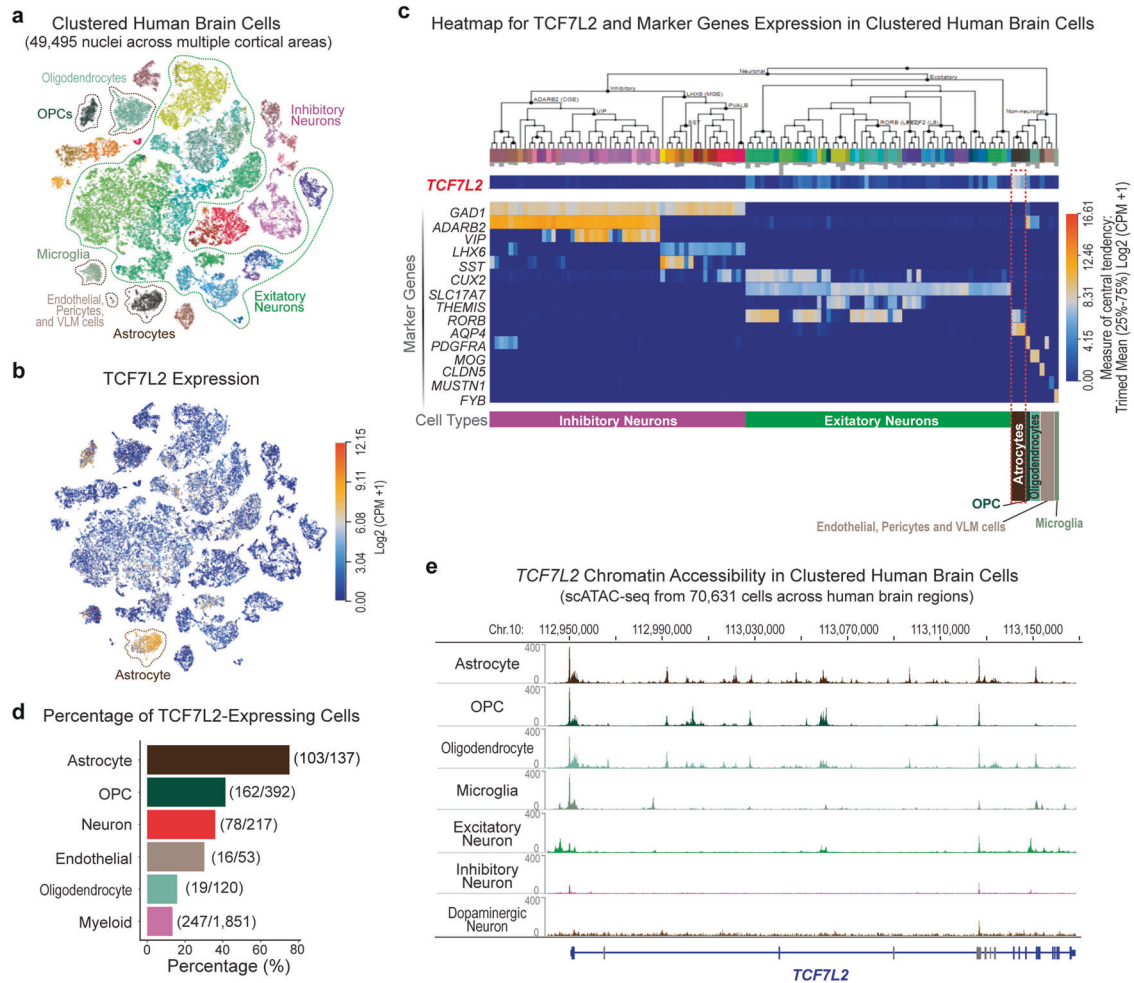


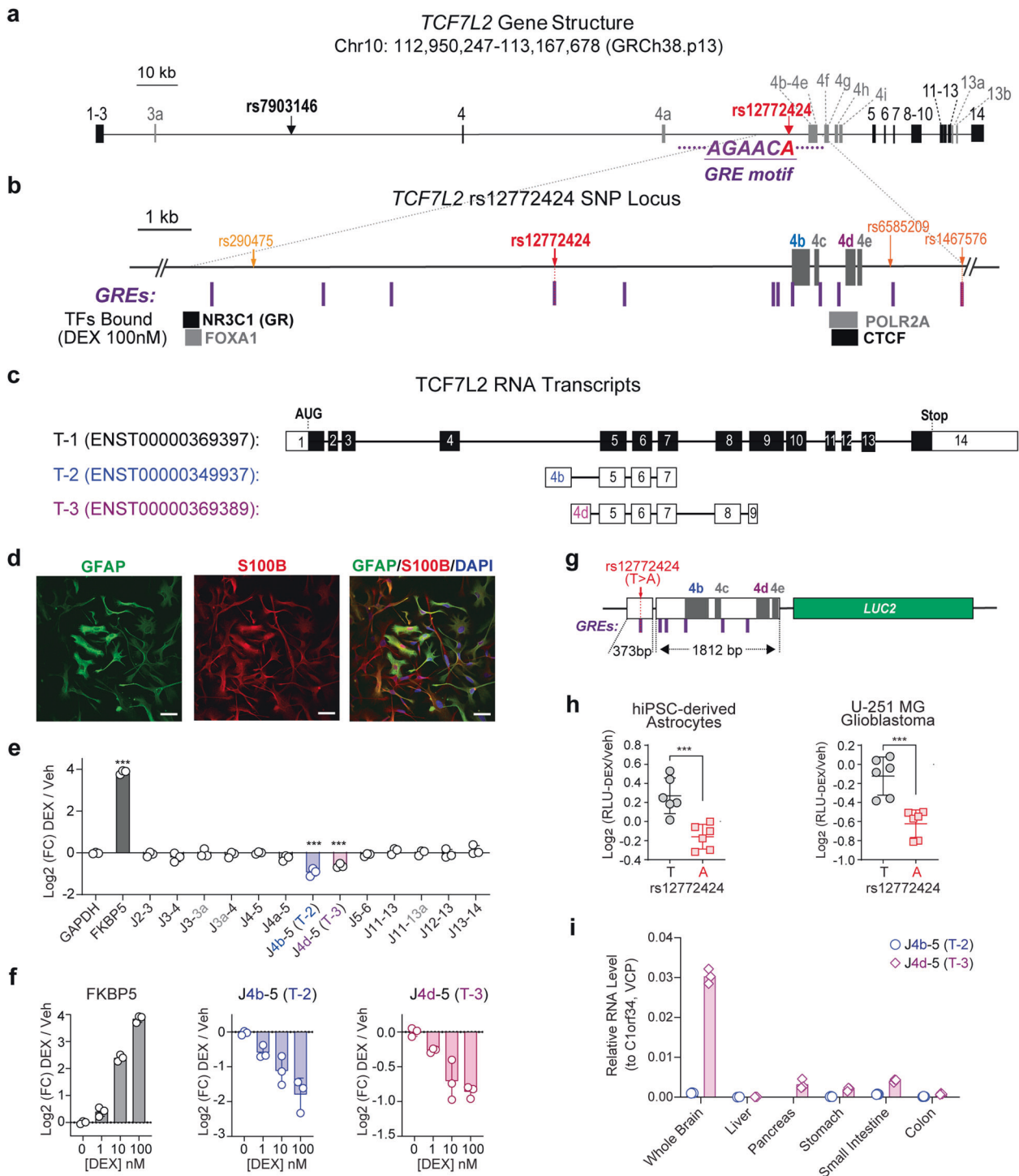
Fig. 1 *TCF7L2* transcription in the human brain. **a** Clusters of human brain cells that were assigned cell types based on the expression of cell type-specific marker genes quantified by single-nucleus RNA-seq in 49,495 single nuclei from multiple cortical areas. Each dot represents a single cell. Cell types have been color-coded and labeled. OPCs: oligodendrocyte progenitor cells. **b** *TCF7L2* expression in each of these brain cells was color-coded based on *TCF7L2* expression levels. **c** Heatmap for the bulk expression level of *TCF7L2* and marker genes in the clustered brain cell. Expression level in astrocytes is highlighted in the dashed box. **a–c** show data generated by the Allen Institute (<https://portal.brain-map.org/atlas-and-data/rnaseq>) [47]. **d** Percentages of *TCF7L2*-expressing cells in each sub-group of cell types are shown graphically. Numbers shown in parentheses are the numbers of *TCF7L2*-expressing cells/total number of sequenced cells. Single-cell RNA-seq data were generated from surgically removed human cerebral cortical tissues in two studies [48, 49]. The two data sets were combined and re-analyzed as reported in a previous study [55]. **e** *TCF7L2* chromatin accessibility was quantified by single-cell ATAC-seq in 70,631 cells across human brain regions. ATAC-seq peaks across the *TCF7L2* gene in each clustered cell type are shown separately from top to bottom. Physical positions on chromosome 10 (Chr.10) are labeled across the top. Based on its physical location on Chr.10 (GRCh38/p13), the *TCF7L2* gene is depicted at the bottom with exons shown as bars and introns as lines. The scATAC-seq data were generated by Corces et al. [50] and visualized by using the WashU Epigenome Browser [81].

scATAC-seq data generated from 70,631 human brain cells [50]. The human brain scATAC-seq data showed that *TCF7L2* chromatin accessibility in astrocytes was higher than that for other clustered CNS cell types (Fig. 1e), compatible with highly active *TCF7L2* transcription in astrocytes. Since *TCF7L2* is most highly expressed in astrocytes, and since insulin signaling in astrocytes is known to co-regulate CNS glucose sensing and systemic glucose metabolism [56, 57], we next set out to characterize *TCF7L2* function in hiPSC-derived astrocytes in an effort to understand its possible role in the comorbidity of BD and obesity.

The BMI-dependent BD risk SNP regulates brain-specific *TCF7L2* transcript variant expression through glucocorticoid signaling

We began this series of experiments by attempting to understand the *TCF7L2* gene structure and the SNP signal that had initially suggested a possible relationship of *TCF7L2* to BD risk. The *Ensembl* [58] human genome assembly (GRCh38/p13) annotated 26 human *TCF7L2* exons

(numbered in Fig. 2a) which gave rise to 30 possible transcript variants (see Supplementary Tables S2 and S3). The most highly expressed *TCF7L2* transcript across human tissues is a protein-coding transcript (ENST00000369397) consisting of 14 exons [30]. We designated that transcript as the *TCF7L2* reference transcript (Fig. 2a). Alternative exons that mapped to intronic regions of the reference transcript were named on the basis of intron number, followed by a letter of the alphabet (see Fig. 2a). The *TCF7L2* rs12772424 SNP that was associated with the BD-BMI interaction is independent of the *TCF7L2* rs7903146 T2D risk SNP, e.g., they are not in LD in any racial or ethnic groups ($r^2 < 0.035$, $D' < 0.282$). Unlike the T2D risk rs7903146 SNP, which is an eQTL for *TCF7L2* RNA levels in multiple tissues [30], the rs12772424 SNP is not an eQTL for *TCF7L2* in any tissue or organ [30]. However, we noticed that the variant allele for the rs12772424 SNP created a half-palindrome GRE (see Fig. 2a), a DNA motif that binds to the ligand-activated GR, which could then either enhance or repress gene transcription [59]. We also noted that eleven additional GREs were present in the “LD block” that includes rs12772424



(Fig. 2b). To determine whether the GR or other GR-related transcription factors (TFs) might bind to this “LD block”, we consulted the ENCODE [60] database that contains ChIP-seq data generated using A549 lung cancer cells after treatment with the GR agonist, DEX. Those ChIP-seq data showed that GR (encoded by *NR3C1*), FOXA1, POLR2A, and CTCF were all bound to this haplotype block (Fig. 2b). These observations raised the possibility that *TCF7L2* transcription might be influenced by glucocorticoid signaling.

To determine whether glucocorticoids might regulate *TCF7L2* transcription, we designed qRT-PCR primers that targeted all known *TCF7L2* exon junctions (J) (see Supplementary Table S4) to

make it possible to quantitate all *TCF7L2* RNA transcripts after DEX treatment. The reference *TCF7L2* transcript variant is designated “T-1” in Fig. 2c. More than half of the 30 *Ensembl*-annotated *TCF7L2* transcript variants failed to include exons 1–4 of “T-1”. Furthermore, as we will demonstrate subsequently, the expression of two “short” variant transcripts, namely “T-2” and “T-3” (see Fig. 2c), were significantly repressed by DEX treatment.

As a first step, we measured RNA levels for all annotated *TCF7L2* transcripts in A549 cells since ChIP-seq showed that GR and other TFs bound near the rs12772424 SNP in A549 cells after treatment with 100 nM DEX (Fig. 2b). The RNA level of *FKBP5*, a prototypic

Fig. 2 *TCF7L2* SNPs and glucocorticoid-regulated transcript variants. **a** *TCF7L2* gene structure. *TCF7L2* exons are numbered from 1 to 14 based on the reference RNA transcript (ID: ENST00000369397). Alternative exons that map to introns of the reference transcript were named using the intron number, followed by a letter of the alphabet. Positions of the rs7903146 SNP and the rs12772424 SNP are indicated and they map 122.2 kb distant from each other. The rs12772424 SNP variant allele (A) created a half-palindrome glucocorticoid response element (GRE). **b** The rs12772424 SNP locus. Positions of rs12772424 and its “linked” SNPs ($r^2 > 0.35$, $D' > 0.79$ in European population) are labeled. *TCF7L2* exons are depicted as boxes with exon numbers labeled above. Putative GREs are indicated below. Transcription factor (TF) binding sites for the glucocorticoid receptor (NR3C1), FOXA1, POLR2A, and CTCF are shown as boxes below the GREs. Darker colors represent stronger binding. **c** *TCF7L2* transcript variants (T-1 to T-3) annotated by *Ensembl* are shown with their IDs shown in parentheses. **d** Immunofluorescent staining of astrocyte markers, GFAP and S100B, for hiPSC-derived astrocytes. Scale bars represent 50 μ m. **e** Fold changes in RNA levels for *TCF7L2* transcript variants in hiPSC-derived astrocytes after treatment with 100 nM DEX. Unique exon junctions (“J”) for specific transcript variants were quantified by qRT-PCR. *GAPDH* and *FKBP5* expression was measured as internal and treatment controls, respectively. Data are mean \pm s.d. ($n = 3$). $***P < 0.001$ by Dunnett’s test. **f** DEX-dose-dependent induction of *FKBP5* and repression of “T-2” and “T-3” in hiPSC-derived astrocytes. **g** Construction of reporter gene plasmid. DNA fragments that included the *TCF7L2* rs12772424 SNP (373 bp) and exons 4b–4d (1812 bp) were cloned into the 5’-region of the *LUC2* reporter gene. **h** Luciferase assays after cells were transfected with plasmids containing rs12772424 wild type (T) or variant (A) alleles and were treated with DEX or vehicle control. Data are presented as relative luciferase units (RLUs) after dexamethasone/vehicle (DEX/Veh) normalization. $***P < 0.001$ by *t* test. **i** Relative RNA levels of *TCF7L2* “T-2” and “T-3” compared with those of the housekeeping genes, *VCP* and *C1orf43*, in total RNA samples prepared from human tissue lysates. Data are mean \pm s.d. ($n = 3$).

GR-inducible gene, was also assayed as a positive control for response to DEX treatment. As expected, *FKBP5* RNA expression increased significantly after DEX treatment (Supplementary Fig. S2a). However, the expression of three of the possible *TCF7L2* transcript variants, including “T-2” and “T-3”, was significantly downregulated, whereas the expression of other *TCF7L2* potential transcript variants was not changed significantly (Supplementary Fig. S2a) nor were they detectable by qRT-PCR (data not shown). Furthermore, repression of the “T-2” and “T-3” transcripts was DEX-dose dependent (Supplementary Fig. S2b). In an attempt to validate these observations for A549 cells by using CNS cell lines, we cultured human iPSC-derived astrocytes (Fig. 2d) and once again showed that “T-2” and “T-3” expression was repressed by DEX treatment (Fig. 2e) in a dose-dependent fashion (Fig. 2f). These observations were also reproduced in U251-MG human glioblastoma cells (Supplementary Fig. S2c, d), indicating that the regulation of those *TCF7L2* transcript variants by glucocorticoid signaling may be a general phenotype across cell lines.

To study the possible role of the rs12772424 SNP in GR-dependent repression of “T-2” and “T-3” in the presence of DEX, DNA fragments containing the rs12772424 SNP and exon 4b–4d sequences were cloned into luciferase reporter gene constructs (see Fig. 2g). Constructs containing the rs12772424 SNP variant allele (A), which creates a GRE motif, were associated with decreased luciferase activity in both iPSC-derived astrocytes and U251-MG glioblastoma cells after DEX treatment (Fig. 2h), suggesting that the transcriptional activity of the cloned DNA fragments was GR and rs12772424 SNP-dependent in both types of cells.

Finally, “T-3” RNA was much more abundant in total RNA samples obtained from the human brain than in total RNA isolated from liver, pancreas, or gut tissue (Fig. 2i), tissues in which the function of *TCF7L2* may be linked to T2D [32–37]. Compared with “T-3”, RNA levels of “T-2” were very low or undetectable in all tissues that were tested (Fig. 2i). In summary, this series of experiments suggested that the *TCF7L2* transcript variant “T-3”, which is highly expressed in the human brain but not in the other tissues that we tested, might be linked to the BD-BMI risk SNP rs12772424 through GR signaling—a possibility supported by the fact that the rs12772424 SNP creates a GRE.

***TCF7L2* “T-3” is a long non-coding RNA transcript that functions in the regulation of *TCF7L2* mRNA and of genes involved in metabolic pathways**

Transcript variant “T-3” was distinguished from other *TCF7L2* transcripts by its first exon, exon-4d (Fig. 3a). In addition to quantification of the “T-3” unique exon junction, “J4d-5”, by the use of qRT-PCR (Fig. 2), we were also able to amplify all six “T-3” exons by RT-PCR using primers mapping to “T-3” (see Supplementary

Table S5) and total RNA prepared from human brain tissue as a template (Fig. 3a–c). This result confirmed the existence of “T-3” in the human brain. We should note that “T-3” has been annotated as a protein-coding transcript without a stop codon or a 3’ untranslated region (UTR) but—based on our results—it appears to be a *TCF7L2* long non-coding RNA transcript (lncRNA-*TCF7L2*) in the human brain. That conclusion is supported by the western blot assays shown subsequently, which failed to detect the putative protein band (~18 kD) that would be encoded by “T-3” and by the fact that PCR performed with primers targeting “T-3” failed to amplify a product when cDNA synthesized from human brain poly (A) + RNA was used as template (Fig. 3c). This result also explains why, when RNA-seq is performed with poly(A)+ selected cDNA libraries which capture mainly protein-coding RNA, including the libraries used by GTEX [30] and the scRNA-seq data that we analyzed [48, 49], “T-3” could not be detected.

In an effort to study the possible function of “T-3”, we first quantified “T-3” expression in hiPSC-derived astrocytes, hepatocytes, and pancreatic β cells (cell characterization shown in Supplementary Fig. S3), which showed that “T-3” was highly expressed in hiPSC-derived astrocytes but not in two peripheral cell models (Fig. 3d). The presence of “T-3” in hiPSC-derived astrocytes was further confirmed by cDNA amplification using total RNA as a template (Fig. 3e). “T-3” was then KD in hiPSC-derived astrocytes by using antisense oligonucleotides (ASO) targeting exon-4d. The “T-3” KD efficiency was >80% in astrocytes transfected with pooled ASOs (Fig. 3f). That RNA sample, together with a non-targeting ASO control, was then used to perform RNA-seq. “T-3” KD in astrocytes resulted in a total of 1363 DEGs with fold change (FC) ≥ 2 (FDR < 0.05), which included *TCF7L2* poly(A)+/mRNA (see Fig. 3g and Supplementary Table S6). Western blot analysis further confirmed the decrease in *TCF7L2* protein after “T-3” KD in hiPSC-derived astrocytes (Fig. 3h), an observation that was reproduced in two additional cell lines that express “T-3” (Supplementary Fig. S4). When the DEGs after T-3 KD were used to perform pathway enrichment analysis using Enrichr [52], the “top” enriched pathway in the BioPlanet database [61] was “unfolded protein response” (Fig. 3i), an endoplasmic reticulum stress response pathway which is well-known to play a role in diabetes, metabolic syndrome and neurodegenerative disorders [62]. Diabetes and insulin signaling pathways were also identified by the pathway enrichment analysis (Fig. 3i). Those DEGs also identified “waist circumference adjusted for BMI” as the “top” associated phenotype in the GWAS Catalog [63] (Fig. 3j), a phenotype that is a clinical measure related to obesity and related disease risk [64]. In summary, we experimentally validated the existence of the *TCF7L2* long non-coding transcript “T-3” in the human brain and, particularly, in hiPSC-derived astrocytes. The RNA-seq analysis after “T-3” KD indicated that this lncRNA might contribute to the regulation of the expression of genes related to

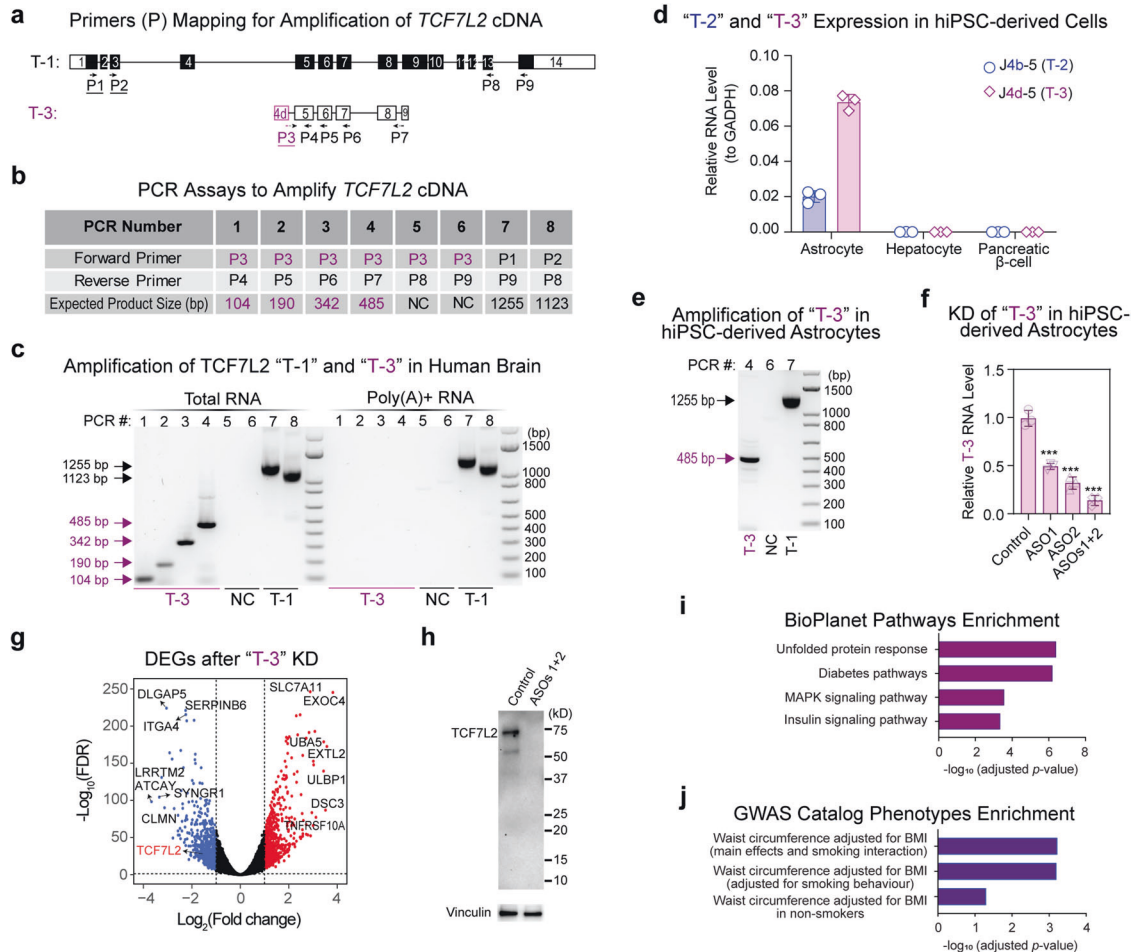


Fig. 3 Characterization of *TCF7L2* Transcript "T-3". **a** *TCF7L2* transcripts "T-1" and "T-3" and mapping sites of primers (P) that were used to amplify *TCF7L2* cDNA. Primer mapping sites are labeled by arrows below the exons depicted as boxes in which exon numbers are labeled. Primers (P) are numbered from 5' to 3' of the gene. Forward primers are underlined. Primer sequences are listed in Supplementary Table S5. **b** The table lists eight polymerase chain reactions (PCR) used to amplify portions of the *TCF7L2* cDNA. The combination of primers for each PCR reaction and the estimated base pair (bp) length of the PCR product are listed. NC = negative control reaction, which was not expected to yield a PCR product. **c** RT-PCR products using human brain total RNA or poly(A) + RNA as templates are shown after separation on a 1.2% agarose gel. PCR numbers (#) match the PCR assays listed in **b**. PCR assays that were designed to amplify "T-3" (from #1 to #4) yielded products of the expected sizes when total RNA was used as the PCR template but failed to amplify target sequences when poly(A) + RNA was the template. NC reactions (#5 and #6) and positive control reactions (#7 and #8) that amplify *TCF7L2* mRNA transcript "T-1" produced the expected results. **d** Relative RNA levels of *TCF7L2* "T-2" and "T-3" compared with GAPDH in hiPSC-derived astrocytes, hepatocytes, and pancreatic β -cells. Data are mean \pm s.d. **e** RT-PCR products amplified *TCF7L2* cDNA in hiPSC-derived astrocytes. **f** qRT-PCR determined "T-3" levels in hiPSC-derived astrocytes after transfection with antisense oligonucleotides (ASOs) targeting exon-4d. The Y axis represents "T-3" expression relative to control (non-targeting ASO). The combination of ASO 1 and 2 (ASOs 1+2) resulted in the best knockdown (KD) efficiency. **g** Volcano plot for RNA-seq identified differentially expressed genes (DEGs) in hiPSC-derived astrocytes. *TCF7L2* mRNA expression was downregulated after "T-3" KD. **h** Western blot assays for *TCF7L2* after knockdown of lncRNA-*TCF7L2* "T-3" in hiPSC-derived astrocytes. **i** Top pathways enriched by DEGs of "T-3" KD (FC > 2.0; FDR < 0.05) in BioPlanet data set and **j** GWAS Catalog phenotypes.

metabolic phenotypes, a link to the BD-BMI phenotype identified during our original GWAS in which rs12772424 was the "top hit" SNP.

Integration of *TCF7L2* RNA-seq and ChIP-seq identified genes that play a role in CNS development

TCF7L2 mRNA was more highly expressed in astrocytes than in other CNS cell types (Fig. 1). However, the function of *TCF7L2* as a TF in astrocytes remained unclear. Since *TCF7L2* mRNA and protein levels were significantly decreased after "T-3" KD (Fig. 3g, h), we next performed *TCF7L2* RNA-seq and ChIP-seq to identify *TCF7L2* target genes in hiPSC-derived astrocytes. *TCF7L2* mRNA was efficiently (>80%) KD by siRNA (Fig. 4a), which mainly targets cytoplasmic mRNA [65]. KD of *TCF7L2* protein was confirmed by western blot assay (Fig. 4b). DEGs were identified by the comparison of RNA-seq data obtained from astrocytes transfected

with non-targeting control and those transfected with *TCF7L2* siRNAs (Fig. 4c). We also performed *TCF7L2* ChIP-seq to identify the DNA-binding sites for this TF across the genome of hiPSC-derived astrocytes (Fig. 4d). A total of 1858 *TCF7L2* ChIP peaks were identified, approximately one-third of which mapped to promoters and transcription start and termination sites (TSS and TTS) (Fig. 4e). As expected, DNA motif analysis identified the prototypic *TCF7L2*-binding motif as the "top hit", supporting the specificity of the ChIP-seq results (Fig. 4f). The ChIP-seq peaks were then integrated with RNA-seq-identified DEGs [53] to identify genes directly targeted by *TCF7L2* (Fig. 4g). A total of 186 "top" DEGs after *TCF7L2* KD that bound *TCF7L2* were identified based on the ChIP-seq data (Fig. 4g, Supplementary Table S7). Those 186 genes included *TCF7L2* itself, which was previously known to be bound by *TCF7L2* protein [66], the *AXIN2* gene that is a prototypic Wnt target gene [67], and other genes that are known to be

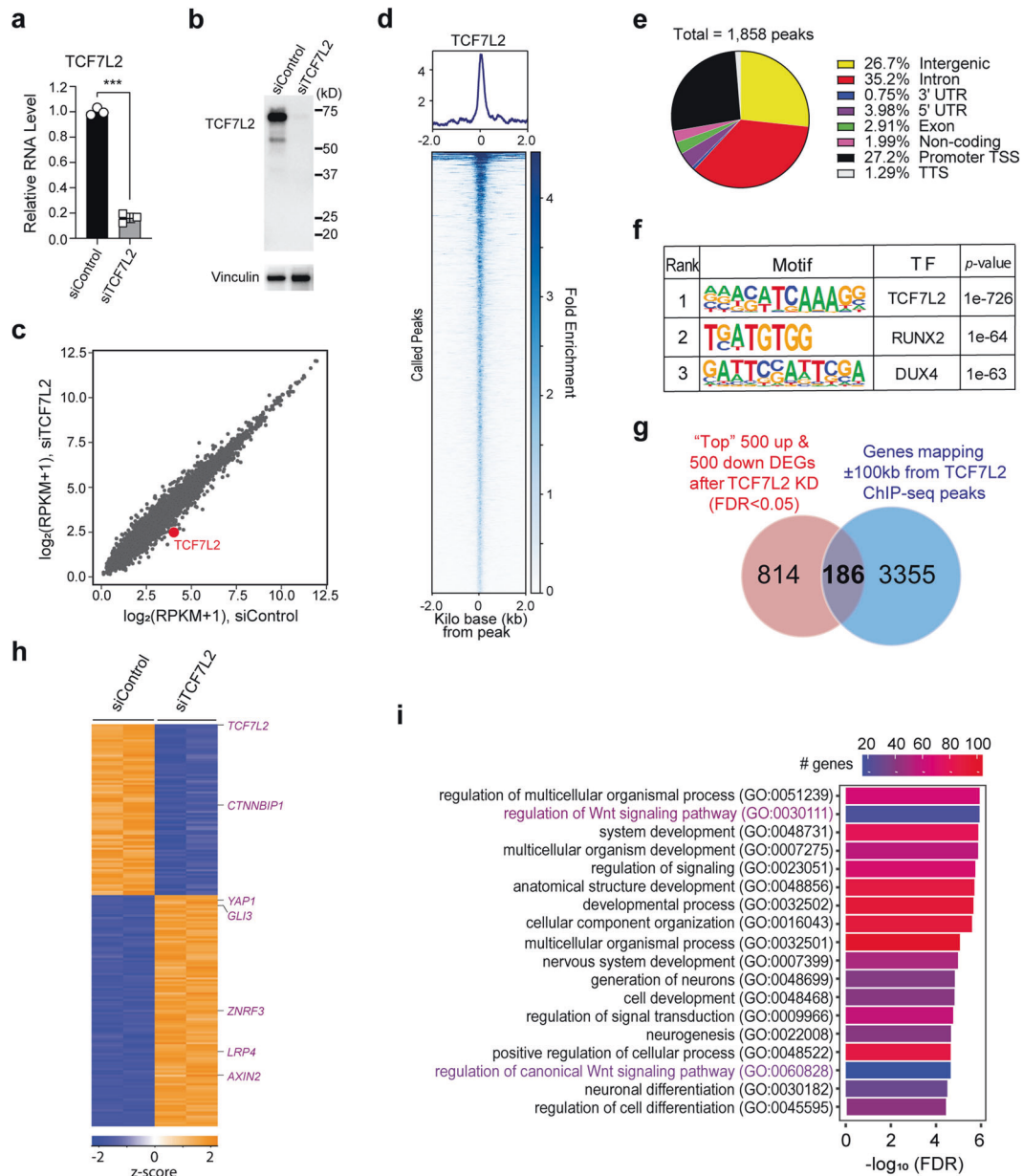


Fig. 4 TCF7L2 protein function in hiPSC-derived astrocytes. **a** TCF7L2 RNA levels in hiPSC-derived astrocytes after transfection with non-targeting siRNA control (siControl) or "pooled" TCF7L2-targeting siRNAs (siTCF7L2). **b** Western blot assay for TCF7L2. Vinculin was blotted as an internal control. **c** Expression levels as \log_2 (reads per kilobase of transcript per million reads mapped [RPKM] + 1) values for all detected genes in the RNA-seq libraries for siTCF7L2 (Y axis) compared with siControl (X axis). $n = 2$ independent replicates for both groups. **d** Heatmap of TCF7L2 ChIP-seq results, in which regions covering the called peaks ± 2 kb were scaled to the same size relative to the center point of the peaks. The heatmap intensity represents fold enrichment of the TCF7L2 ChIP signal over input, with each row depicting each called peak, and with the X axis representing position relative to the center point of the peaks. The top panel is the average profile of the ChIP-seq peaks, with the Y axis representing fold enrichment of the TCF7L2 ChIP signal over input. **e** Distribution of 1858 identified TCF7L2 peaks (FDR < 0.01) across the genome. *UTR* untranslated region, *TSS* transcription start site, *TTS* transcription termination site. **f** The top three binding motifs and their binding transcription factors (TFs) were identified from HOMER de novo analysis based on p values. **g** Integration of RNA-seq and ChIP-seq data identified 186 genes that were both bound by TCF7L2 and were differentially expressed after TCF7L2 KD. Based on false discovery rate (FDR < 0.05), the "top" 500 up- and 500 downregulated DEGs and genes mapping ± 100 kb from the TCF7L2 ChIP peaks (3541 genes) were used to identify "overlapping" TCF7L2 target genes. **h** Heatmap of expression for the 186 TCF7L2 target genes in control and TCF7L2 KD samples. Genes involved in the Wnt signaling pathways have been labeled. **i** Gene Ontology enrichment analysis for the 186 TCF7L2 target genes identified by integration of the RNA-seq and ChIP-seq data. Colors represent the number of genes enriched in each pathway. The X axis represents $-\log_{10}$ of the FDR. Wnt signaling pathways have been highlighted.

involved in Wnt signaling (Fig. 4h). When the 186 TCF7L2 target genes in hiPSC-derived astrocytes were subjected to pathway enrichment analysis, as expected, the Wnt signaling pathway was identified as one of the "top" biological pathways (Fig. 4i). In

addition, developmental pathways, including nervous system development and neurogenesis pathways, were enriched by the 186 TCF7L2 target genes (Fig. 4i), a result consistent with the known role of TCF7L2 in Wnt signaling and in cell development

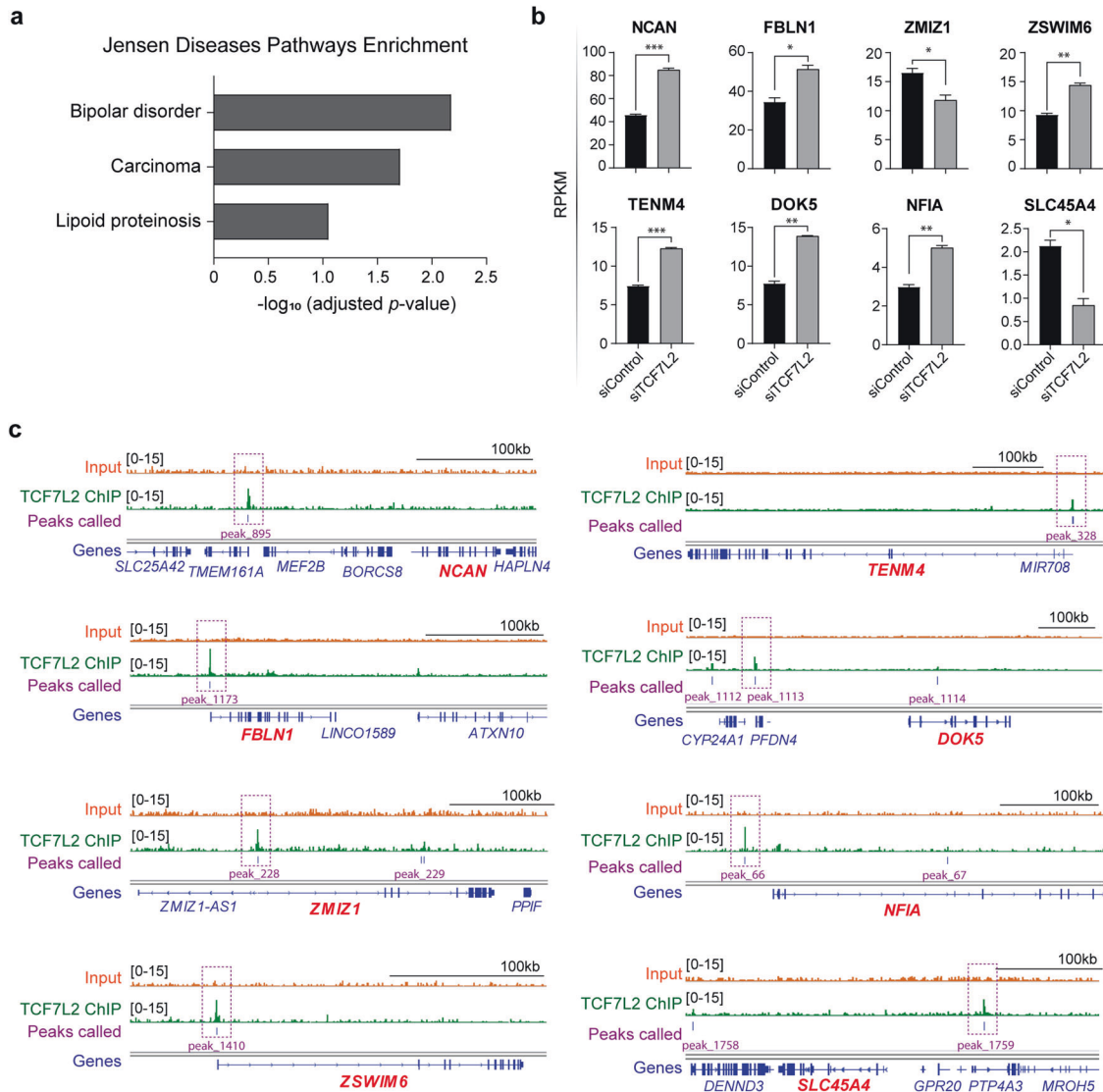


Fig. 5 TCF7L2 directly regulates BD risk genes. **a** Disease pathway enrichment of the 186 TCF7L2 target genes that showed bipolar disorder as the only significant enrichment in the Jensen's Lab Disease database. No other pathway was significantly ($FDR < 0.01$) enriched in databases on the Enrichr website. **b** Expression levels of BD risk genes enriched in the disease pathway analysis were significantly changed after TCF7L2 KD in hiPSC-derived astrocytes. **c** Integrative Genomic Views of TCF7L2 ChIP peaks mapped to the BD risk genes shown in **b**. The orange panel represents the input control. The green panel represents TCF7L2 ChIP-seq reads. The purple panel represents TCF7L2 ChIP peaks that were called based on statistical significance ($FDR < 0.01$). Major ChIP peaks have been highlighted in dotted boxes. Genes mapping to the TCF7L2 ChIP peaks are shown in the blue panels at the bottom of each panel. The names of BD risk genes for which expression was changed after TCF7L2 KD are highlighted in red.

[28]. These results provided evidence which suggested that TCF7L2 may play a role in human CNS development, a process that has been linked to BD pathophysiology [68].

TCF7L2 directly regulates BD risk genes

To further investigate possible disease pathway(s) that might be related to the 186 TCF7L2 target genes in astrocytes, disease pathway enrichment was performed. BD was the most significantly enriched disease for these genes ($FDR < 0.05$) based on the Jensen disease database [69] (Fig. 5a). Eight of the enriched BD genes were *NCAN*, *TENM4* (previously named *ODZ4*), *FBLN1*, *ZMIZ1*, *NFIA*, *ZSWIM6*, *DOK5*, and *SLC45A4*. Expression of all eight of these BD risk genes was significantly changed in astrocytes after TCF7L2 KD (Fig. 5b). Finally, ChIP-seq demonstrated that TCF7L2 bound directly to sequences in or near all of these genes, suggesting direct regulation of their expression (Fig. 5c). Integration of RNA-seq and ChIP-seq improved the accuracy of the identification of

TCF7L2 target genes, e.g., TCF7L2 ChIP peaks mapped ~100 kb away from *NCAN* and *SCL45A4* (Fig. 5c) but based on the RNA-seq data, no other genes that mapped nearer to those peaks were differentially expressed in hiPSC-derived astrocytes after TCF7L2 KD (Supplementary Table S7), suggesting possible direct regulation of *NCAN* and *SCL45A4* expression. Seven of the TCF7L2-regulated BD risk genes shown in Fig. 5b, with the exception of *SCL45A4*, which encodes a sucrose transporter, are known to function in CNS cellular differentiation or neurodevelopment (Supplementary Table S8). Furthermore, seven of those eight genes (all except *TENM4*) contain genetic polymorphisms that are genome-wide significantly associated with either BMI or metabolic phenotypes (see Supplementary Table S8). These results suggested that TCF7L2 may directly regulate genes associated with BD risk and BMI or metabolic phenotypes in astrocytes, a possible molecular mechanism related to the role of TCF7L2 in the comorbidity of BD and obesity.

DISCUSSION

Metabolic comorbidities are highly prevalent in BD and are associated with more severe mood symptoms and worse treatment outcomes [16–19]. Disrupted insulin signaling in the brain and the hypothalamic–pituitary–adrenal axis has been proposed as one explanation for the comorbidity of mood disorders and obesity [70]. Recent large GWA studies of BD risk [6, 7] and obesity [71] have provided additional evidence for genetic correlations between BD and metabolic syndromes. Specifically, pathway enrichment of GWAS-identified BD risk genes often identified metabolic pathways such as insulin signaling and energy metabolism [6, 7]. Moreover, pathway analyses of BMI-related genes have strongly supported a role for the CNS in obesity susceptibility [71]. Obesity management by pharmacologic treatment in patients with psychiatric disorders is presently being tested as a novel therapeutic intervention [72]. In an effort to understand the possible contribution of genetics to the comorbidity of BD and obesity, we previously performed a GWAS for BD risk incorporating interaction with BMI, a study in which we identified a genome-wide significant SNP (rs12772424) associated with BMI-related BD that mapped to the *TCF7L2* gene [20]. In the present study, we functionally characterized *TCF7L2* in astrocytes, CNS cells that highly express *TCF7L2* (Fig. 1), in an attempt to understand its possible role in the comorbidity of BD and obesity. Our results suggest that the *TCF7L2* gene might play a role in both CNS development and in the regulation of metabolism in astrocytes, at least in part through the actions of different *TCF7L2* variant transcripts.

TCF7L2 has been studied extensively as a TF in peripheral tissues with a role in glucose metabolism and T2D risk. The *TCF7L2* protein has also been well characterized and includes a β -catenin-binding domain, suggesting a role in the canonical Wnt pathway (or Wnt/ β -catenin pathway) [73]. Our studies have confirmed its role in Wnt signaling in hiPSC-derived astrocytes (Fig. 4), a function that might be related to BD risk (see Fig. 5). However, *TCF7L2* transcription is complicated, e.g., with a total of 30 transcript variants listed in the *Ensembl* database. More than half of those transcript variants do not include exons 1–4 (Fig. 2a), which encode the β -catenin-binding domain [73], supporting Wnt-independent *TCF7L2* functions that have been observed in the mouse brain [38]. The *TCF7L2* rs12772424 SNP that we reported led us to a series of lncRNA-*TCF7L2* transcripts that do not include exons 1–4, and specifically to “T-3”, which is highly expressed in the human brain and with an expression that is regulated by glucocorticoid signaling (Fig. 2). lncRNAs are known to affect the expression of a large number of genes by both *cis*- and *trans*-regulation [74] and to play key roles in various brain disorders [75]. Our functional characterization of “T-3” in hiPSC-derived astrocytes revealed a role in the regulation of the expression of genes involved in diabetes and insulin signaling pathways (Fig. 3), results which suggest that in the CNS *TCF7L2*, in addition to its role in the liver [32–34], pancreas [35, 36], and gut [37], might also contribute to metabolic regulation.

Abnormal cortisol levels in BD patients have been observed during DEX/CRH testing [76] and the GR is expressed in astrocytes [77]. Astrocytes have recently been recognized as key players in neuropsychiatric disorders, including BD [78, 79], and they have been proposed as cellular targets for BD pharmacotherapy [80]. Recent studies have demonstrated that dysfunction of insulin signaling in hypothalamic astrocytes can influence systemic glucose levels in mice [56, 57]. Of course, these results with experimental animals will also have to be pursued in human subjects. Our identification of the “T-3” glucocorticoid-responsive lncRNA-*TCF7L2* (Fig. 2), a non-coding transcript that regulates the expression of both *TCF7L2* mRNA and of genes involved in insulin signaling pathways in hiPSC-derived astrocytes (Fig. 3) may help to explain, in part, the correlation of glucocorticoid signaling, insulin signaling, and BMI in BD.

Finally, we should point out the limitations of our studies, beginning with our use of cell line models. We used hiPSC-derived astrocytes to study *TCF7L2* molecular function because *TCF7L2* is more highly expressed in astrocytes than in other CNS cell types and, to our knowledge, its function in human CNS cell model(s) has not been reported previously. However, *TCF7L2* is also expressed in neurons and oligodendrocyte progenitor cells (see Fig. 1), and it may regulate oligodendrocyte differentiation in mice [38, 39]. Whether *TCF7L2* in other CNS cell types might contribute to BD and/or BMI is unknown and should be the subject of future studies. However, our data strongly suggest that *TCF7L2* may play dual roles in the regulation of astrocytic insulin signaling and CNS development, perhaps through the effects of different RNA transcripts, a possibility that should also be pursued in the future. Obviously, molecular mechanisms based on cell line model systems are not able to provide a comprehensive view of physiological processes underlying clinical phenotypes. Therefore, genetically modified animal models and/or three-dimensional human brain organoids represent logical next steps in pursuit of the observations reported here. We should also point out that, although BMI was the best obesity-related phenotype that we had for all of our GWAS patients, we appreciate that it is not an ideal measure of obesity [21] and that other phenotypes such as waist circumference could provide complementary information on obesity-related phenotypes. In this context, it is of interest that “waist circumference adjusted for BMI” emerged as a top GWAS Catalog phenotype in our pathway enrichment analysis using DEGs after knockdown of the lncRNA-*TCF7L2* transcript in astrocytes (see Fig. 3j).

In summary, by setting out to define the molecular function of *TCF7L2* in hiPSC-derived astrocytes, we have demonstrated that glucocorticoid regulation of the transcription of a *TCF7L2* encoded non-coding transcript in a SNP-dependent fashion might influence astrocytic insulin signaling pathways, an effect which could also influence systemic glucose metabolism. We also obtained evidence that astrocytic *TCF7L2* might contribute to BD pathophysiology by regulating the expression of BD risk genes. These results strongly support the possibility that the function of *TCF7L2*, and specifically the function of the lncRNA-*TCF7L2* T-3 variant transcript in the CNS might represent an important link between BD, BMI, and metabolism.

DATA AVAILABILITY

The RNA-seq and ChIP-seq data sets generated during the current study are available in the Gene Expression Omnibus with accession ID: GSE179922.

REFERENCES

1. Frye MA. Bipolar disorder — a focus on depression. *N Engl J Med*. 2011;364:51–9.
2. Kessler RC, Crum RM, Warner LA, Nelson CB, Schulenberg J, Anthony JC. Lifetime co-occurrence of DSM-III-R alcohol abuse and dependence with other psychiatric disorders in the National Comorbidity Survey. *Arch Gen Psychiatry*. 1997;54:313–21.
3. Oquendo MA, Currier D, Liu SM, Hasin DS, Grant BF, Blanco C. Increased risk for suicidal behavior in comorbid bipolar disorder and alcohol use disorders: results from the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). *J Clin Psychiatry*. 2010;71:902–9.
4. Frye MA, Altshuler LL, McElroy SL, Suppes T, Keck PE, Denicoff K, et al. Gender differences in prevalence, risk, and clinical correlates of alcoholism comorbidity in bipolar disorder. *Am J Psychiatry*. 2003;160:883–9.
5. McElroy SL, Frye MA, Helleman G, Altshuler L, Leverich GS, Suppes T, et al. Prevalence and correlates of eating disorders in 875 patients with bipolar disorder. *J Affect Disord*. 2011;128:191–8.
6. Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetskov V, et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet*. 2019;51:793–803.
7. Li HJ, Zhang C, Hui L, Zhou DS, Li Y, Zhang CY, et al. Novel risk loci associated with genetic risk for bipolar disorder among han chinese individuals: a genome-wide association study and meta-analysis. *JAMA Psychiatry*. 2021;78:320–30.

8. Cichon S, Mühleisen TW, Degenhardt FA, Mattheisen M, Miró X, Strohmaier J, et al. Genome-wide association study identifies genetic variation in neurocan as a susceptibility factor for bipolar disorder. *Am J Hum Genet.* 2011;88:372–81.
9. PsychiatricGWASConsortiumBipolarDisorderWorkingGroup. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet.* 2011;43:977–83.
10. Mühleisen TW, Leber M, Schulze TG, Strohmaier J, Degenhardt F, Treutlein J, et al. Genome-wide association study reveals two new risk loci for bipolar disorder. *Nat Commun.* 2014;5:3339.
11. Ikeda M, Takahashi A, Kamatani Y, Okahisa Y, Kunugi H, Mori N, et al. A genome-wide association study identifies two novel susceptibility loci and trans population polygenicity associated with bipolar disorder. *Mol Psychiatry.* 2018;23:639–47.
12. Ferreira MA, O'donovan MC, Meng YA, Jones IR, Ruderfer DM, Jones L, et al. Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. *Nat Genet.* 2008;40:1056–8.
13. Mullins N, Forstner AJ, O'connell KS, Coombes B, Coleman JRI, Qiao Z, et al. Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nat Genet.* 2021;53:817–29.
14. Nurnberger Jr., Koller DL, Jung J, Edenberg HJ, Foroud T, Guella I, et al. Identification of pathways for bipolar disorder: a meta-analysis. *JAMA Psychiatry.* 2014;71:657–64.
15. Holgerson AA, Clark MM, Frye MA, Kellogg TA, Mundi MS, Veldic M, et al. Symptoms of bipolar disorder are associated with lower bariatric surgery completion rates and higher food addiction. *Eat Behav.* 2021;40:101462.
16. Vancampfort D, Vansteelandt K, Correll CU, Mitchell AJ, De Herdt A, Sienaert P, et al. Metabolic syndrome and metabolic abnormalities in bipolar disorder: a meta-analysis of prevalence rates and moderators. *Am J Psychiatry.* 2013;170:265–74.
17. Frye MA, McElroy SL, Fuentes M, Sutor B, Schack KM, Galardy CW, et al. Development of a bipolar disorder biobank: differential phenotyping for subsequent biomarker analyses. *Int J Bipolar Disord.* 2015;3:30.
18. Fagioli A, Kupfer DJ, Houck PR, Novick DM, Frank E. Obesity as a correlate of outcome in patients with bipolar I disorder. *Am J Psychiatry.* 2003;160:112–7.
19. McElroy SL, Crow S, Biernacka JM, Winham S, Geske J, Cuellar Barboza AB, et al. Clinical phenotype of bipolar disorder with comorbid binge eating disorder. *J Affect Disord.* 2013;150:981–6.
20. Winham SJ, Cuellar-Barboza AB, Oliveros A, McElroy SL, Crow S, Colby C, et al. Genome-wide association study of bipolar disorder accounting for effect of body mass index identifies a new risk allele in TCF7L2. *Mol Psychiatry.* 2014;19:1010–16.
21. Nimptsch K, Konigorski S, Pischon T. Diagnosis of obesity and use of obesity biomarkers in science and clinical medicine. *Metabolism.* 2019;92:61–70.
22. Adams KF, Schatzkin A, Harris TB, Kipnis V, Mouw T, Ballard-Barbash R, et al. Overweight, obesity, and mortality in a large prospective cohort of persons 50 to 71 years old. *N Engl J Med.* 2006;355:763–78.
23. Cuellar-Barboza AB, Winham SJ, McElroy SL, Geske JR, Jenkins GD, Colby CL, et al. Accumulating evidence for a role of TCF7L2 variants in bipolar disorder with elevated body mass index. *Bipolar Disord.* 2016;18:124–35.
24. Grant SF, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, Sainz J, et al. Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet.* 2006;38:320–3.
25. Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, Novartis Institutes of BioMedical Research, Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, and Novartis Institutes of BioMedical R, Saxena R, Voight BF, Lyssenko V, Burt NP, de Bakker PI, et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science.* 2007;316:1331–6.
26. Xue A, Wu Y, Zhu Z, Zhang F, Kemper KE, Zheng Z, et al. Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. *Nat Commun.* 2018;9:2941.
27. Vujkovic M, Keaton JM, Lynch JA, Miller DR, Zhou J, Tcheandjieu C, et al. Discovery of 318 new risk loci for type 2 diabetes and related vascular outcomes among 1.4 million participants in a multi-ancestry meta-analysis. *Nat Genet.* 2020;52:680–91.
28. Jin T, Liu L. The Wnt signaling pathway effector TCF7L2 and type 2 diabetes mellitus. *Mol Endocrinol.* 2008;22:2383–92.
29. Nestler EJ, Gould E, Manji H, Buncan M, Duman RS, Greshenfeld HK, et al. Pre-clinical models: status of basic research in depression. *Biol Psychiatry.* 2002;52:503–28.
30. GTExConsortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science.* 2020;369:1318–30.
31. Nobrega MA. TCF7L2 and glucose metabolism: time to look beyond the pancreas. *Diabetes.* 2013;62:706–8.
32. Boj SF, van Es JH, Huch M, Li VS, José A, Hatzis P, et al. Diabetes risk gene and Wnt effector Tcf7l2/TCF4 controls hepatic response to perinatal and adult metabolic demand. *Cell.* 2012;151:1595–607.
33. Ip W, Shao W, Song Z, Chen Z, Wheeler MB, Jin T. Liver-specific expression of dominant-negative transcription factor 7-like 2 causes progressive impairment in glucose homeostasis. *Diabetes.* 2015;64:1923–32.
34. Oh KJ, Park J, Kim SS, Oh H, Choi CS, Koo SH. TCF7L2 modulates glucose homeostasis by regulating CREB- and FoxO1-dependent transcriptional pathway in the liver. *PLoS Genet.* 2012;8:e1002986.
35. Mitchell RK, Mondragon A, Chen L, McGinty JA, French PM, Ferrer J, et al. Selective disruption of Tcf7l2 in the pancreatic beta cell impairs secretory function and lowers beta cell mass. *Hum Mol Genet.* 2015;24:1390–99.
36. da Silva Xavier G, Mondragon A, Mourougavelou V, Cruciani-Guglielmacci C, Denom J, Herrera PL, et al. Pancreatic alpha cell-selective deletion of Tcf7l2 impairs glucagon secretion and counter-regulatory responses to hypoglycaemia in mice. *Diabetologia.* 2017;60:1043–50.
37. Shao W, Wang D, Chiang YT, Ip W, Zhu L, Xu F, et al. The Wnt signaling pathway effector TCF7L2 controls gut and brain proglucagon gene expression and glucose homeostasis. *Diabetes.* 2013;62:789–800.
38. Hammond E, Lang J, Maeda Y, Pleasure D, Angus-Hill M, Xu J, et al. The Wnt effector transcription factor 7-like 2 positively regulates oligodendrocyte differentiation in a manner independent of Wnt/beta-catenin signaling. *J Neurosci.* 2015;35:5007–22.
39. Zhao C, Deng Y, Liu L, Yu K, Zhang L, Wang H, et al. Dual regulatory switch through interactions of Tcf7l2/Tcf4 with stage-specific partners propels oligodendroglial maturation. *Nat Commun.* 2016;7:10883.
40. Duncan A, Heyer MP, Ishikawa M, Caligiuri SPB, Liu XA, Chen Z, et al. Habenular TCF7L2 links nicotine addiction to diabetes. *Nature.* 2019;574:372–7.
41. Savic D, Ye H, Aneas I, Park SY, Bell GI, Nobrega MA. Alterations in TCF7L2 expression define its role as a key regulator of glucose metabolism. *Genome Res.* 2011;21:1417–25.
42. Brinkmeier ML, Potok MA, Cha KB, Gridley T, Stifani S, Meeldijk J, et al. TCF and Groucho-related genes influence pituitary growth and development. *Mol Endocrinol.* 2003;17:2152–61.
43. Hansen T, Ingason A, Djurovic S, Melle I, Fenger M, Gustafsson O, et al. At-risk variant in TCF7L2 for type II diabetes increases risk of schizophrenia. *Biol Psychiatry.* 2011;70:59–63.
44. Neavin DR, Lee JH, Liu D, Ye Z, Li H, Wang L, et al. Single nucleotide polymorphisms at a distance from aryl hydrocarbon receptor (AHR) binding sites influence AHR ligand-dependent gene expression. *Drug Metab Dispos.* 2019;47:983–94.
45. Liu D, Qin S, Ray B, Kalari KR, Wang L, Weinshilboum RM. Single nucleotide polymorphisms (SNPs) distant from xenobiotic response elements can modulate aryl hydrocarbon receptor function: SNP-dependent CYP1A1 induction. *Drug Metab Dispos.* 2018;46:1372–81.
46. Ingle JN, Liu M, Wickerham DL, Schaid DJ, Wang L, Mushiroda T, et al. Selective estrogen receptor modulators and pharmacogenomic variation in ZNF423 regulation of BRCA1 expression: individualized breast cancer prevention. *Cancer Discov.* 2013;3:812–25.
47. Hodge RD, Bakken TE, Miller JA, Smith KA, Barkan ER, Grayback LT, et al. Conserved cell types with divergent features in human versus mouse cortex. *Nature.* 2019;573:61–8.
48. Darmanis S, Sloan SA, Croote D, Mignardi M, Chernikova S, Samghababi P, et al. Single-cell RNA-seq analysis of infiltrating neoplastic cells at the migrating front of human glioblastoma. *Cell Rep.* 2017;21:1399–410.
49. Darmanis S, Sloan SA, Zhang Y, Enge M, Caneda C, Shuer LM, et al. A survey of human brain transcriptome diversity at the single cell level. *Proc Natl Acad Sci USA.* 2015;112:7285–90.
50. Corces MR, Shcherbina A, Kundu S, Gloudemans MJ, Frésard L, Granja JM, et al. Single-cell epigenomic analyses implicate candidate causal variants at inherited risk loci for Alzheimer's and Parkinson's diseases. *Nat Genet.* 2020;52:1158–68.
51. Ward LD, Kellis M. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res.* 2016;44:D877–81.
52. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* 2016;44:W90–97.
53. Wang S, Sun H, Ma J, Zang C, Wang C, Wang J, et al. Target analysis by integration of transcriptome and ChIP-seq data with BETA. *Nat Protoc.* 2013;8:2502–15.
54. Robinson JT, Thorvaldsdóttir H, Winkler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol.* 2011;29:24–6.
55. Liu D, Zhuang Y, Zhang L, Gao H, Neavin D, Carrillo-Roa T et al. ERICH3: vesicular association and antidepressant treatment response. *Mol Psychiatry.* 2020; <https://doi.org/10.1038/s41380-020-00940-y>.

56. García-Cáceres C, Quarta C, Varela L, Gao Y, Gruber T, Legutko B, et al. Astrocytic insulin signaling couples brain glucose uptake with nutrient availability. *Cell*. 2016;166:867–80.
57. García-Cáceres C, Bolland E, Prevot V, Luquet S, Woods SC, Koch M, et al. Role of astrocytes, microglia, and tanycytes in brain control of systemic metabolism. *Nat Neurosci*. 2019;22:7–14.
58. Yates AD, Achuthan P, Akanni W, Allen J, Allen J, Alvarez-Jarreta J, et al. Ensembl 2020. *Nucleic Acids Res*. 2020;48:D682–D8.
59. Weikum ER, Knuesel MT, Ortlund EA, Yamamoto KR. Glucocorticoid receptor control of transcription: precision and plasticity via allostery. *Nat Rev Mol Cell Biol*. 2017;18:159–74.
60. Davis CA, Hitz BC, Sloan CA, Chan ET, Davidson JM, Gabdank I, et al. The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Res*. 2018;46:D794–801.
61. Huang R, Grishagin I, Wang Y, Zhao T, Greene J, Obenauer JC, et al. The NCATS BioPlanet - an integrated platform for exploring the universe of cellular signaling pathways for toxicology, systems biology, and chemical genomics. *Front Pharm*. 2019;10:445.
62. Wang S, Kaufman RJ. The impact of the unfolded protein response on human disease. *J Cell Biol*. 2012;197:857–67.
63. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res*. 2019;47: D1005–12.
64. Ross R, Neeland IJ, Yamashita S, Shai I, Seidell J, Magni P, et al. Waist circumference as a vital sign in clinical practice: a Consensus Statement from the IAS and ICCR Working Group on Visceral Obesity. *Nat Rev Endocrinol*. 2020;16:177–89.
65. Hannon GJ. RNA interference. *Nature*. 2002;418:244–51.
66. Frietze S, Wang R, Yao L, Tak YG, Ye Z, Gaddis M, et al. Cell type-specific binding patterns reveal that TCF7L2 can be tethered to the genome by association with GATA3. *Genome Biol*. 2012;13:R52.
67. Yan D, Wiesmann M, Rohan M, Chan V, Jefferson AB, Guo L, et al. Elevated expression of axin2 and hnkcd mRNA provides evidence that Wnt/beta-catenin signaling is activated in human colon tumors. *Proc Natl Acad Sci USA*. 2001;98:14973–8.
68. Madison JM, Zhou F, Nigam A, Hussain A, Barker DD, Nehme R, et al. Characterization of bipolar disorder patient-specific induced pluripotent stem cells from a family reveals neurodevelopmental and mRNA expression abnormalities. *Mol Psychiatry*. 2015;20:703–17.
69. Pletscher-Frankild S, Palleja A, Tsafou K, Binder JX, Jensen LJ. DISEASES: text mining and data integration of disease-gene associations. *Methods*. 2015;74:83–9.
70. Nguyen TTL, Chan LC, Borregine K, Kale RP, Hu C, Tye SJ. A review of brain insulin signaling in mood disorders: from biomarker to clinical target. *Neurosci Biobehav Rev*. 2018;92:7–15.
71. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature*. 2015;518:197–206.
72. Whicher CA, Price HC, Phiri P, Rathod S, Barnard-Kelly K, Reidy C, et al. Liraglutide and the management of overweight and obesity in people with schizophrenia, schizoaffective disorder and first-episode psychosis: protocol for a pilot trial. *Trials*. 2019;20:633.
73. Poy F, Lepourcelet M, Shivdasani RA, Eck MJ. Structure of a human Tcf4-beta-catenin complex. *Nat Struct Biol*. 2001;8:1053–7.
74. Statello L, Guo CJ, Chen LL, Huarte M. Gene regulation by long non-coding RNAs and its biological functions. *Nat Rev Mol Cell Biol*. 2021;22:96–118.
75. Yang S, Lim KH, Kim SH, Joo JY. Molecular landscape of long noncoding RNAs in brain disorders. *Mol Psychiatry*. 2021;26:1060–74.
76. Vieta E, Martínez-De-Osaba MJ, Colom F, Martínez-Aran A, Benabarre A, Gasto C. Enhanced corticotropin response to corticotropin-releasing hormone as a predictor of mania in euthymic bipolar patients. *Psychol Med*. 1999;29:971–8.
77. Bohn MC, Howard E, Vielkind U, Krozowski Z. Glial cells express both mineralocorticoid and glucocorticoid receptors. *J Steroid Biochem Mol Biol*. 1991;40:105–11.
78. Seifert G, Schilling K, Steinhauser C. Astrocyte dysfunction in neurological disorders: a molecular perspective. *Nat Rev Neurosci*. 2006;7:194–206.
79. Toker L, Mancarci BO, Tripathy S, Pavlidis P. Transcriptomic evidence for alterations in astrocytes and parvalbumin interneurons in subjects with bipolar disorder and schizophrenia. *Biol Psychiatry*. 2018;84:787–96.
80. Rivera AD, Butt AM. Astrocytes are direct cellular targets of lithium treatment: novel roles for lysyl oxidase and peroxisome-proliferator activated receptor-gamma as astroglial targets of lithium. *Transl Psychiatry*. 2019;9:211.
81. Zhou X, Lowdon RF, Li D, Lawson HA, Madden PA, Costello JF, et al. Exploring long-range genome interactions using the WashU Epigenome Browser. *Nat Methods*. 2013;10:375–6.

ACKNOWLEDGEMENTS

This study was supported in part, by the Mayo Foundation and the Mayo Center for Individualized Medicine and by NIH grants U19 GM61388, R01 AA27486, and R01 GM28157 to Dr. Weinshilboum; R01 DK58185 to Dr. Ordog. We want to thank the Allen Institute, the GTEx Project, and the ENCODE Consortium for generating the particular data sets that have been cited in this study.

AUTHOR CONTRIBUTIONS

D.L., T.T.L.N., J.M.B., M.A.F., and R.M.W. conceived and designed the study; D.L., T.T.L.N., H.H., D.C.K., B.S., and J-H.L. collected the data; D.L., T.T.L.N., H.G., and Z.Y. analyzed the data; L.W. provided resources and interpretive advice; T.O., M.A.F., and R.M.W. supervised the study; D.L., T.T.L.N., and R.M.W. wrote the manuscript; B.J.C., T.O., J.M.B., and M.A.F. edited manuscript. All authors read, revised, and approved the paper.

COMPETING INTERESTS

R.M.W. and L.W. are co-founders and stockholders in OneOme, LLC; M.A.F. receives grant support from Assurex Health, Mayo Foundation, and Medibio. Other authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41380-021-01274-z>.

Correspondence and requests for materials should be addressed to Mark A. Frye or Richard M. Weinshilboum.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021