The landscape of *SETBP1* gene expression and transcription factor activity across human tissues

Jordan H. Whitlock: jwhitlock@uab.edu, Elizabeth J. Wilk: lizzyr@uab.edu, Timothy C. Howton: tchowton@uab.edu, Amanda D. Clark: clarkad@uab.edu, Brittany N. Lasseigne*: bnp0001@uab.edu (*corresponding author)

Keywords

RNA-sequencing, cross-tissue-analysis, Schinzel Giedion, cancer, SETBP1-HD, SETBP1, transcription factor activity

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported in part by the UAB Lasseigne Lab funds, UAB Pilot Center for Precision Animal Modeling (C-PAM)(1U54OD030167)(to BNL), and the UAB Predoctoral Training Grant in Cell, Molecular, and Developmental Biology (CMDB T32)(5T32GM008111-35)(to JHW).

Authors' contributions

Conceptualization: JHW, EJW, TCH, ADC, and BNL. Methodology: JHW and BNL. Software: JHW (including the Shiny app) and EJW. Validation: JHW, EJW, TCH, and ADC. Formal Analysis: JHW and EJW. Investigation: JHW, EJW. Resources: BNL. Data Curation: JHW. Writing - Original Draft: JHW. Writing - Review & Editing: JHW, EJW, TCH, ADC, and BNL. Visualization: JHW and EJW. Supervision: BNL. Project Administration: BNL. Funding Acquisition: BNL.

Abstract

The SET binding protein 1 (*SETBP1*) gene encodes a transcription factor (TF) involved in various cellular processes. Distinct *SETBP1* variants have been linked to three different diseases. Germline variants cause the ultra-rare pediatric Schinzel Giedion Syndrome (SGS) and *SETBP1* haploinsufficiency disorder (*SETBP1*-HD), characterized by severe multisystemic abnormalities with neurodegeneration or a less severe brain phenotype accompanied by hypotonia and strabismus, respectively. Somatic variants in *SETBP1* are associated with hematological malignancies and cancer development in other tissues in adults. To better understand the tissue-specific mechanisms involving *SETBP1*, we analyzed publicly available RNA-sequencing data from the Genotype-Tissue Expression (GTEx) project. We found *SETBP1* and its known target genes were widely expressed across 31 adult human tissues. Hierarchical clustering identified 3 distinct expression patterns of SETBP1 targets across tissues. Functional enrichment analysis revealed gene sets related to transcription regulation, DNA binding, and mitochondrial function. TF activity analysis of SETBP1 and its target TFs revealed tissue-specific TF activity, underscoring the role of tissue context-driven regulation and suggesting its impact in SETBP1-associated disease. In addition to uncovering tissue-specific molecular signatures of *SETBP1* expression and TF activity, we provide a Shiny web application to facilitate the exploration of TF activity across human tissues for 758 TFs.

Introduction

SETBP1 is a gene located on the long (q) arm of chromosome 18 that encodes for the transcription factor (TF) and oncogene SET binding protein 1.(Coccaro et al., 2017) Referred to as a DNA-binding protein, SETBP1 has several motifs, including three nuclear localization signals, a SKI homology region, and a binding region for SET nuclear oncogene. As a protein, it has a role in DNA replication and transcriptional regulation.(Piazza et al., 2018) Different pathogenic variants in *SETBP1* can result in three distinct diseases.(Kohyanagi & Ohama, 2023) Germline variants cause two unique ultra-rare, de novo pediatric diseases: Schinzel Giedion Syndrome (SGS)(Schinzel & Giedion, 1978) and Setbp1 haploinsufficiency disorder (*SETBP1*-HD).(Morgan et al., 2021) These conditions are differentiated by location, phenotypic severity, and accompanying gain or loss of function (GoF, LoF), respectively.(Morgan et al., 2021; Schinzel & Giedion, 1978) SGS is multisystemic, involving gastrointestinal, cardiorespiratory, neurological, musculoskeletal, and urogenital abnormalities, and more severe than *SETBP1*-HD, which is characterized by progressive neurodegeneration and shortened life expectancy.(Liu et al., 2018; 2015) However, SGS and *SETBP1*-HD, as disorders of protein dosage, have overlapping phenotypes, including intellectual disability, developmental delay, language impairment, distinctive facial features, and hypotonia.(Liu et al., 2018; Morgan et al., 2021; Schinzel & Giedion, 1978) In contrast, somatic variants in *SETBP1* are associated with hematological malignancies and exhibit varying evidence for predisposing or promoting cancer in other adult tissue systems (reviewed in(Coccaro et al., 2017)).

There are multiple hypothesized mechanisms for the tissue-specificity of disease (i.e., clinical manifestations in some tissues but not others) related to intrinsic and extrinsic molecular processes spanning epigenetic, genetic, expression, regulation, and network-based mechanisms.(Hekselman & Yeger-Lotem, 2020) Despite genomic advances in variant identification and sequencing technology, gaps remain in translating the role of specific genomic variants to observed phenotypic outcomes. Of these hypothesized mechanisms of tissue-specific manifestation of disease, preferential or exclusive gene expression of SETBP1 targets and their altered regulation remain understudied in SETBP1-associated disorders. Potential mechanisms of SETBP1 dysfunction involving disrupted cell cycle control, DNA damage mechanisms, phosphatase activity, and chromatin remodeling have been hypothesized.(Antonyan & Ernst, 2022) However, how the expression of *SETBP1* and its known TF targets function across non-diseased human tissues requires further study.

Because of this, publicly available non-diseased data, such as the Genotype-Tissue Expression (GTEx) project, provide an opportunity for investigating and generating hypotheses about the underlying function of disease-associated genes in different contexts, including in non-diseased and diseased contexts. Here we investigated the gene expression of SETBP1 and its known targets in each GTEx tissue and then evaluated the functional enrichment of those targets based on how they clustered by expression across tissues. Next, we inferred the TF activity of SETBP1 and other TFs it directly targets by leveraging multivariate linear models to calculate enrichment scores representing activity for all TFs across tissues using GTEx expression and CollecTRI, a curated collection of TFs and their directional regulation on transcriptional targets, as inputs.(**Figure 1**). Collectively, we have mapped gene expression and activity of SETBP1 and its targets across human non-diseased tissues, underscoring the impact of tissue background. Further, we have developed a Shiny web application (https://lasseignelab.shinyapps.io/gtex_tf_activity/) to facilitate the exploration and hypothesis generation of TF activity across human tissues for 758 TFs **(Table_S2)**.

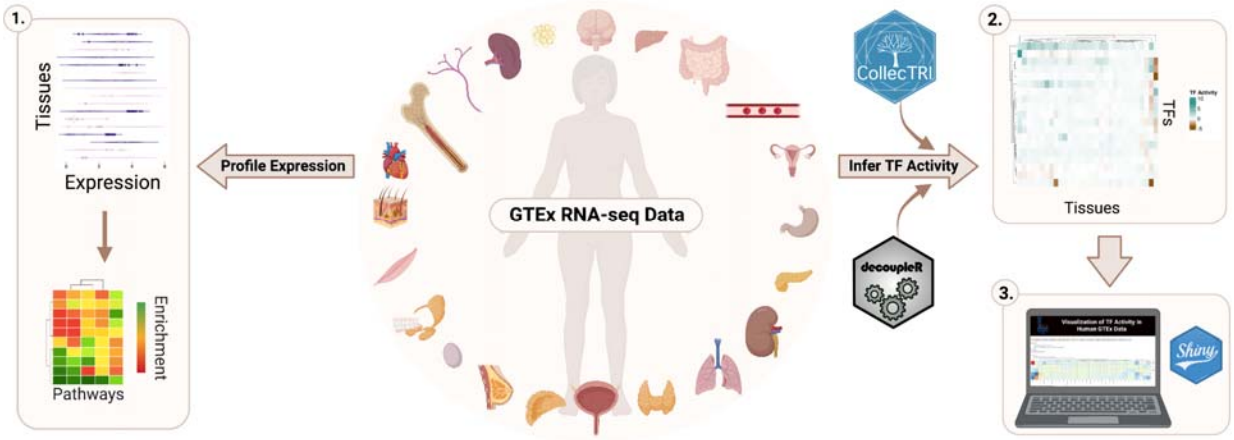

Figure 1 Study Overview. To investigate the tissue-specific expression and transcription factor (TF) activity of *SETBP1*, we analyzed publicly available RNA-sequencing (RNA-seq) data from the Genotype-Tissue Expression (GTEx) project. We profiled the tissue-specific expression of *SETBP1* and its known targets and performed functional enrichment analysis (1). We also inferred TF activity for SETBP1 and TFs it regulates by tissue (2) and developed an interactive web application to enable the exploration of TF activity for 758 TFs across all GTEx tissues (3).

Results and Discussion

SETBP1 is widely expressed across adult human tissues

We first examined the gene expression of *SETBP1* across 31 adult human tissues from the GTEx consortium (n=19,081 samples total). We found that while *SETBP1* was expressed ubiquitously (median transcript per million (TPM) range, 0.364-16.719; **Figure 2A**), it was most highly expressed in the cervix, blood vessel, and uterus (median TPM 16.719, 15.422, and 13.730, respectively) and most lowly expressed in the blood, bone marrow, and the adrenal gland (median TPM 0.135, 0.384, and 0.963, respectively). When we investigated the brain-region-specific expression of *SETBP1*, we found it was similarly expressed across subregions (**Fig. S1**). Further, we examined the gene expression of 209 known SETBP1 targets we previously compiled(Whitlock et al., 2023) from SETBP1 ChIP-seq binding sites(Kolmykov et al., 2021), MSigDB(Liberzon et al., 2015), SIGNOR(Lo Surdo et al., 2023), TRRUST(Han et al., 2018), Piazza et al.(Piazza et al., 2018), and Antonyan et al.(Antonyan & Ernst, 2022) We found that most known SETBP1 targets were also broadly expressed across adult human tissues (62.2% of known SETBP1 targets had a TPM>3 in >90% of tissues; **Figure 2B)**.

Targets of SETBP1 are widely expressed across tissues and cluster by gene expression into functionally distinct pathways

We next assessed the variability of SETBP1 targets' expression across tissues and if particular functions were enriched based on their co-expression across GTEx tissues. Complete linkage hierarchical clustering of all tissues revealed one subclade that included predominantly SETBP1-affected tissues (5 out of 7), including brain, muscle, kidney, heart, and blood **(Figure 2C)**. When clustering *SETBP1* targets by gene expression, we obtained 3 optimal k-means clusters **(Fig. S2, Table S1)** and found a higher correlation of expression between *SETBP1* and its targets in clusters 1 and 2 than in cluster 3 (rho ranges -0.25 to 0.82, -0.28 to 0.84, and -0.46 to 0.69, respectively) **(Table S1)**. The targets whose gene expression was most correlated

to *SETBP1* included *ZEB1* (cluster 2, rho 0.84), *TTC23* (cluster 2, rho 0.82), *RALGAPA1* (cluster 1, rho 0.82), *RAD52* (cluster 2, 0.82), *CZIB* (cluster 2, 0.81), and *PBRM1* (cluster 2, rho 0.77). The most anti-correlated genes were the mitochondrial-associated genes *MT-RNR1* and *MT-TF* (cluster 3, rho -0.462 and -0.459, respectively), *HMGA1* (cluster 2, -0.28), and *MYB*, *CNIH2*, and FKTN-AS1 (cluster 1, rho range -0.25 -0.15). We further analyzed if the SETBP1 gene targets within these clusters had similar functions using over-representation analysis and found that clusters 1, 2, and 3 included genes enriched for the transcription regulator complex, minor groove AT-rich DNA binding, and mitochondrial structure and function, respectively (**Figure 2D**).

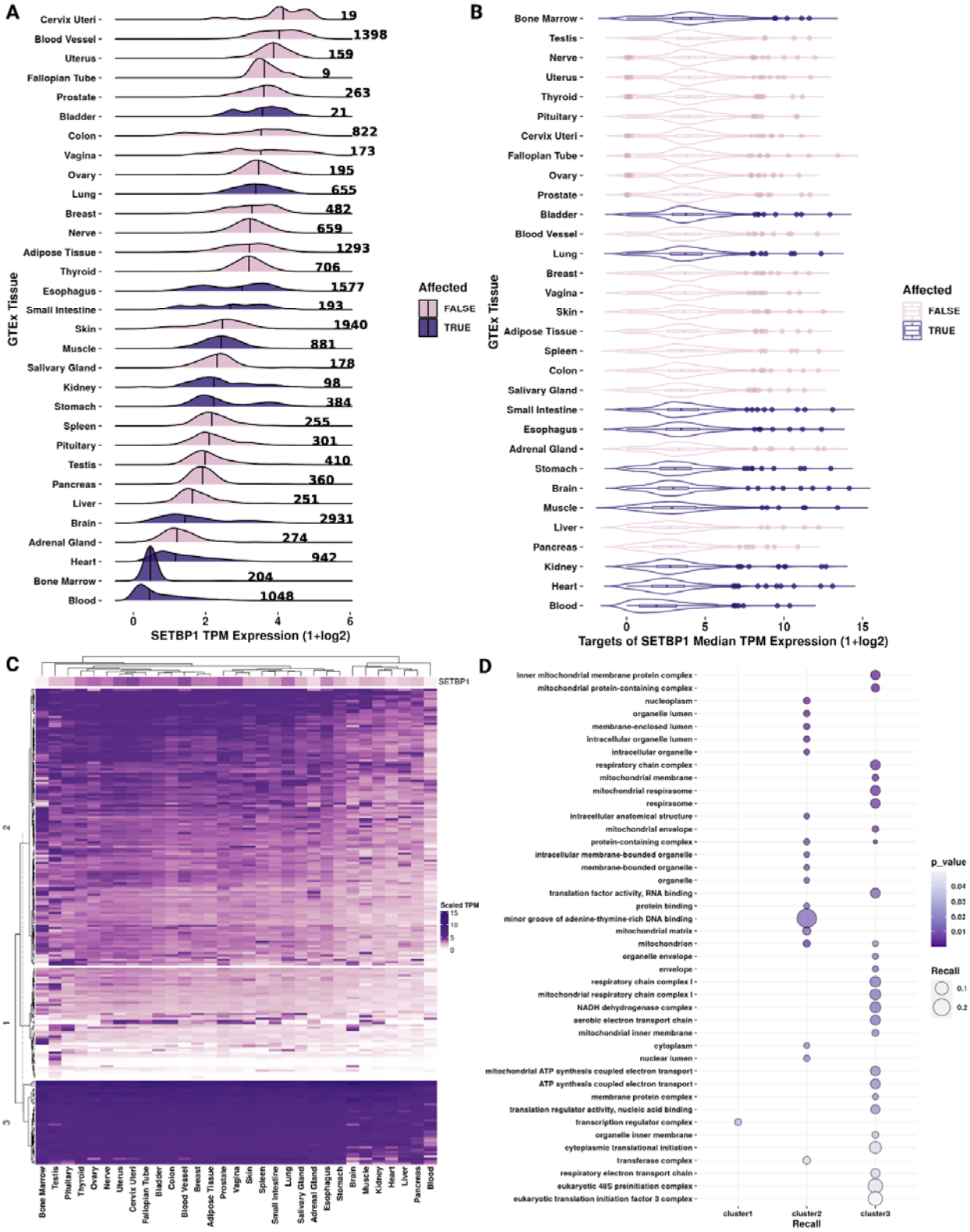

Figure 2 Gene expression of *SETBP1* and its known TF targets across GTEx tissues. Ridgeline and boxplot of the scaled median TPM values across samples for each GTEx tissue for **(A)** *SETBP1* and **(B)** its known targets, respectively. **(C)** Clustering of *SETBP1* and its known transcription factor (TF) targets' gene expression by tissue. **(D)** Dot plot representing the

functional enrichment analysis (FEA) results for genes from clusters 1, 2, and 3 identified in C. Here dot size indicates recall, the proportion of functionally annotated input genes to each term's full geneset size (y-axis), and color indicates significance (darker purple; more significant).

SETBP1 TF activity is decreased in blood and increased in the pituitary

Due to its role as a TF, we next inferred the TF activity of SETBP1 across GTEx tissues. We did this by extracting TF activity for each tissue based on prior knowledge in CollecTRI, a comprehensive, species-specific, curated database of TFs and their transcriptional targets. Using decoupleR, we then built a multivariate linear model taking into account the weight and sign of TF-target interactions and the GTEx gene expression values, where for each tissue, we interpreted positive scores as indicating SETBP1 acting as an activator, negative scores as indicating SETBP1 acting as a repressor, and zero as indicating a lack of coordinated regulation by SETBP1 **(Figure 3)**. We found that across tissues, SETBP1 TF activity was mostly near zero (**Table S2**; range: -0.739 to 3.760 with median = -0.024 and variance = 0.53), with the exception of pituitary (TF activity = 3.760 indicating high activator SETBP1 activity) and blood (TF activity = -0.739 indicating repressive SETBP1 activity). With the exception of blood, we did not find that tissues known to be affected by SETBP1-associated diseases had strong evidence of SETBP1 activation or repression in non-diseased adult tissues.

SETBP1 TF targets demonstrate tissue-specific TF activity

In addition to SETBP1, we also inferred TF activity for targets of SETBP1 that act as TFs **(Figure 3).** We found that these TFs, in addition to SETBP1, had a lack of coordinated TF activity in brain, heart, and kidney, all of which are SETBP1-affected tissues (median = -0.004, variance = 0.008 for brain; median = -0.034, variance = 0.014 for heart; and median = -0.026, variance = 0.022 for kidney). However, in the blood, SETBP1 targets associated with cancer and DNA damage had variable TF activity. For example, we predicted BCL11A (linked to multiple blood cancers as reviewed in(Bao et al., 2019; Yin et al., 2019)) acts as a repressor in both blood and bone marrow under non-diseased conditions (TF activity -4.124 and -3.287, respectively) and MEF2C (often deregulated and associated with recurrence in leukemia(Brown et al., 2018)), ZEB1 (shown to modulate hematopoietic stem cell fates(Almotiri et al., 2021; Li et al., 2021)), and others also act as repressors in blood (TF activity -0.882 and -1.532, respectively).

Furthermore, prior research showed that oncogenic apoptosis resistance and unresolved DNA damage signatures persist as a result of *SETBP1* variants in SGS.(Banfi et al., 2021; Whitlock et al., 2023) Here, we found that in non-diseased adult tissue, TCF4, a known apoptosis regulator(Forrest et al., 2013; Xie et al., 2012), and APEX1, part of the SET complex in DNA damage response(Antonyan & Ernst, 2022) also had variable TF activity across SGS-affected tissues. We predicted TCF4 to have repressor activity in muscle (TF activity = -0.553) and an activator role in small intestine (TF activity = 2.181), while we predicted APEX1 to be repressive in blood (TF activity = -1.810) and an activator in bladder (TF activity = 0.694) and esophagus (TF activity = 0.590) with a lack of coordinated activity in kidney, blood, bone marrow, esophagus and muscle with respect to TCF4 and blood, muscle, and heart for APEX1 **(Figure 3)**. In summary, we have provided a map of SETBP1 and TF targets of SETBP1 activity across non-diseased adult human tissues.

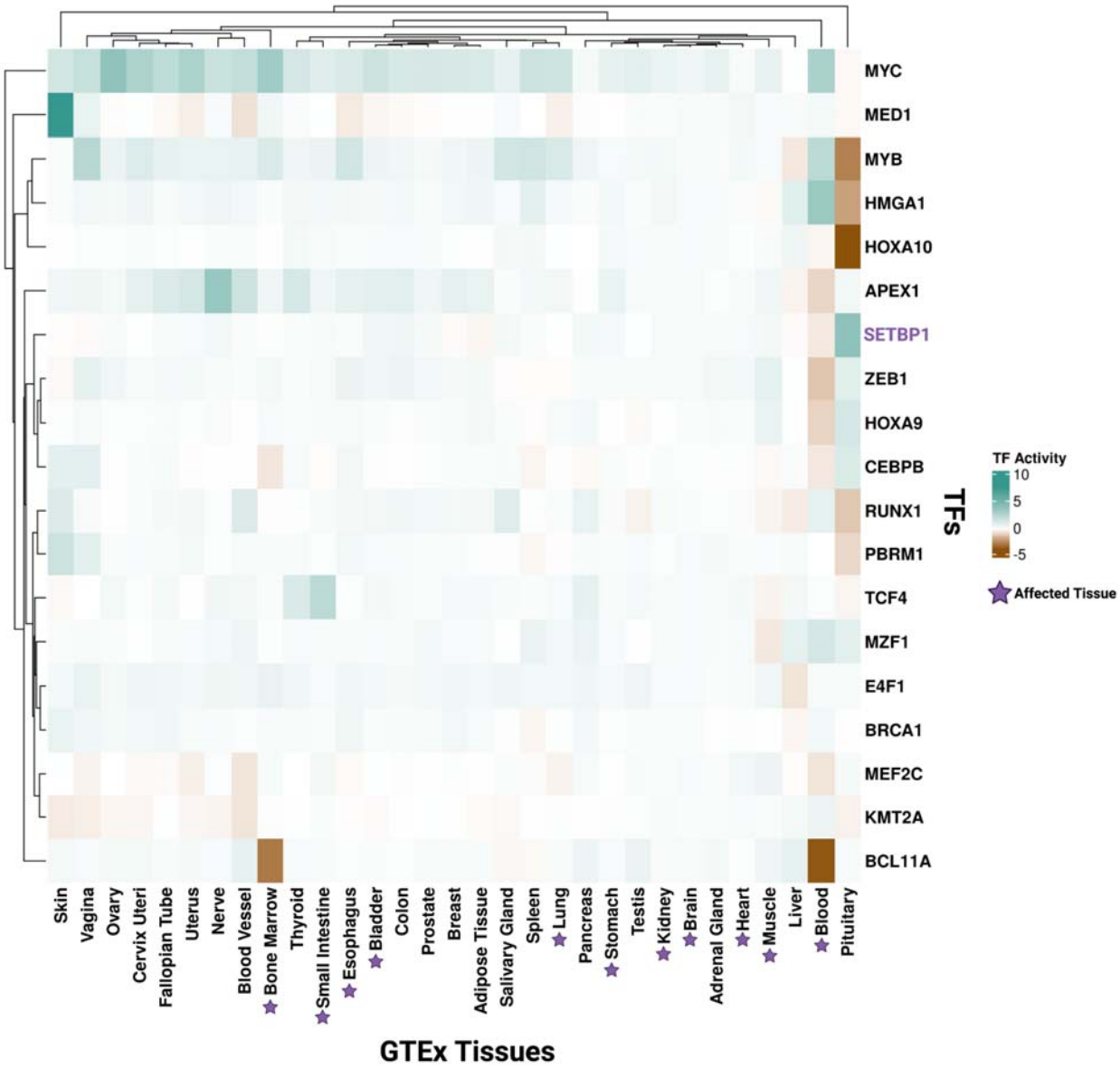

Figure 3 Tissue-specific TF activity of SETBP1 and TFs it targets. Heatmap representing transcription factor (TF) activity scores of SETBP1 (purple) and its known TF targets (black, y-axis) across GTEx tissues (x-axis). Teal and brown represent TF activity as an activator and repressor, respectively.

Discussion

Our study maps tissue-specific molecular signatures associated with SETBP1 across 31 human non-diseased adult tissues. We found that *SETBP1* and its known targets were widely expressed across those tissues, and functional enrichment analysis revealed gene sets related to transcriptional regulation, DNA binding, and mitochondrial function. Further, we uncovered tissue-specific TF activity through TF activity analysis of SETBP1 and its TF targets, underscoring the role of tissue context-driven regulation. We also provide a framework for investigating tissue-specific gene expression and TF activity for other genes, particularly those associated with multiple diseases or multi-syndromic diseases with pathophysiological impacts

across multiple tissues. As many genes associated with developmental disorders are also associated with a predisposition or increased risk of cancer (reviewed in(Balachandran & Narendran, 2023; Nussinov et al., 2022)), applying this framework to those particular genes may be fruitful. To that end, we developed a Shiny web application with pre-computed TF activity scores for 758 TFs that we inferred activity for each GTEx tissue. This app allows users to search for TFs and compare their activity across all 31 GTEx tissues (https://lasseignelab.shinyapps.io/gtex_tf_activity/) .

With complete-linkage hierarchical clustering of SETBP1 target gene expression by tissue, we discovered one subclade largely consisted of affected tissues (5 out of 7 of the tissues in the subclade), including brain, kidney, heart, and blood, the most frequently noted tissues impacted by SETBP1 perturbations (**Figure 2C**).(Acuna-Hidalgo et al., 2017; Coccaro et al., 2017; Matsumoto et al., 2005; Schinzel & Giedion, 1978) This suggests that, in non-diseased tissues, there is a similarity in the expression of SETBP1 targets across multiple tissues known to exhibit a phenotype in SETBP1-associated diseases. We further investigated the trends of expression of SETBP1 targets by k-means clustering and correlation of each target to *SETBP1* gene expression (**Figure 2C**, **Table S1**). We highlighted several of the highest correlated and anti-correlated genes due to SETBP1's roles as an activator as well as a repressor.(Coccaro et al., 2017) Of note, some of the highest correlated and anti-correlated targets included genes with known critical roles in development (*ZEB1*,(Drápela et al., 2020), *RALGAPA1*,(Wagner et al., 2020) *HMGA1*,(Chang et al., 2023)), metastasis (*ZEB1*,(Drápela et al., 2020) *HMGA1*,(Chang et al., 2023) *MYB(Coccaro et al., 2017)*) and therapy resistance (*ZEB1*,(Drápela et al., 2020) *HMGA1*,(Chang et al., 2023) *MYB(Cicirò & Sala, 2021)*). To uncover potential functional patterns by expression clusters, we tested each of our 3 k-means clusters of gene targets for over-representation of GO terms. We found cluster 3 to be characterized as ubiquitously and highly expressed across tissues, with genes enriched for terms regarding mitochondrial structure and function (**Figure 2D**), which included many *SETBP1* targets that are mitochondrially-encoded (MT-ND6, MT-RNR1, MT-TE, MT-TP, MTCO3P12, MT-TF) (**Table S1**). Cluster 2 was enriched with the highest recall (proportion of functionally annotated input genes to each term's full gene set size) was for minor groove AT-rich DNA binding, similar to regulation by SETBP1 as well (Piazza et al., 2018). Cluster 1 was enriched for just one GO term, transcription regulator complex, with targets including cancer-associated transcription factors RUNX1, HOXA9, MYB (Bhagwat & Vakoc, 2015). These functional clusters highlighted the varying roles of SETBP1 targets and their patterns of expression across tissues.

Out of all disease-associated tissues for SETBP1 variants, blood was the only tissue exhibiting notable SETBP1 TF repression. Previous studies support that surpassing a higher functional threshold (i.e., more damaging or impacting variants) is required for SETBP1-driven cancer(Acuna-Hidalgo et al., 2017). Our findings here suggest that compared to SETBP1 in non-diseased blood, GoF variants (already known to drive blood-associated cancers such as myeloid leukemia(Acuna-Hidalgo et al., 2017; Albano et al., 2012; Piazza et al., 2013)) may switch SETBP1 TF repression to activation, suggesting future research directions that may shed light on the mechanism behind SETBP1-associated blood cancers. For example, SETBP1 has been shown to activate *MYC(Carratt et al., 2022)*, which already has a TF activity score of 2.77 in blood, in addition to having positive TF activity scores in 27 out of 30 other tissues **(Figure 3)**. Likewise, we and others(Antonyan & Ernst, 2022; Banfi et al., 2021; Whitlock et al., 2023) have hypothesized that in the presence of pathogenic germline variants, alterations in TCF4 or APEX1 activity lead to apoptosis resistance or increased DNA damage in SGS. Our results here suggest both are active across many tissues in non-diseased adult tissues, so disease-associated perturbations may impact SETBP1 TF activity, contributing to SGS.

A major limitation of this study is that our analyses were conducted in adult bulk expression profiles. For disease-associated genes that also have developmental phenotypes like *SETBP1*, temporal expression(Cardoso-Moreira et al., 2019, 2020) and TF activity are likely key to linking disrupted genes to molecular and physiological phenotypes. Comprehensive prenatal and postnatal gene expression atlases in active development (e.g., developmental GTEx, dGTEx) will provide an unprecedented opportunity to repeat and expand the analyses in this study across developmental time points. Additionally, as we conducted the analyses here in bulk profiles, we cannot assess the gene expression and TF activity of particular cell types. However, as we recently reported, SETBP1's role as an epigenetic hub leads to cell-type-specific differences in TF activity, gene targeting, and regulatory rewiring in the mouse cerebral cortex and kidney(Whitlock et al., 2023). This underscores the importance of future studies that generate and analyze the necessary data to understand cell-type-specific gene expression and TF activity across human tissues. Additionally, tissues may have subclinical pathologies that have not been previously reported, so care must be taken with the interpretation of non-affected tissues.

In summary, our study highlights the importance of considering tissue-specific expression and regulatory properties in investigating disease-related genes and provides a basis for future investigations of TFs involved in processes across many tissues, including developmental and cancer contexts.

Experimental Procedures

Availability of data and materials

Data and code supporting the shiny app and results of this study are available at Zenodo (DOI:10.5281/zenodo.8222799) and https://github.com/lasseignelab/230323_JW_DiseaseNetworks (DOI:10.5281/zenodo.8225613). The docker image used for these analyses is publicly available on Docker Hub (jordanwhitlock/setbp1_manuscript:1.0.12) and Zenodo (DOI:10.5281/zenodo.8225177). Our interactive web application can be accessed at https://lasseignelab.shinyapps.io/gtex_tf_activity/ (DOI:10.5281/zenodo.8225317).

SETBP1 target gene set construction

We used the SETBP1 target gene set from Whitlock et al. 2023 to obtain a list of known TF targets of SETBP1.(Whitlock et al., 2023) All genes were converted to human Ensembl IDs using gprofiler2(Kolberg et al., 2020) (v.0.2.1), resulting in a final list of 209 genes.

RNA-sequencing expression data

Using recount3 (Wilks et al., 2021) (accessed August 2022), we obtained bulk RNA-sequencing (RNA-seq) data represented as transcripts per million (TPM) counts for human tissues (n=31) from the publicly available Genotype-Tissue Expression (GTEx) project. All tissues were included except those classified as "study_na".

Classification of disease-associated affected tissues

We compiled a list of affected tissues in Schinzel Giedion Syndrome (SGS), SETBP1 haploinsufficiency disorder (SETBP1-HD), and SETBP1-associated cancer from the literature, OMIM (MIM #: 616078, 269150), and UpToDate.(Amberger et al., 2019; Elena et al., 2016; Jansen et al., 2021; Morgan et al., 2021; Patnaik et al., 2014; Piazza et al., 2013; Schinzel & Giedion, 1978; 2015) Affected tissues included the following for each disease:

- SETBP1-HD/SGS: brain and muscle
- SGS-specific: heart, kidney, bladder, lung, small intestine, stomach, esophagus
- Cancer: bone marrow and blood

SETBP1 and target gene expression and pathway enrichment

We calculated the median TPM of *SETBP1* and its targets across samples in each tissue and scaled (log2 + 1 transformed) for visualizations. We performed complete linkage hierarchical clustering of tissues. For target gene expression, we identified 3 optimal k-means clusters by elbow method based on the expression of *SETBP1* and plotted median scaled TPM values using ComplexHeatmap (version 2.10.0). (Gu et al., 2016) We next performed functional enrichment analysis (FEA) with gprofiler2 (version 0.2.1)(Kolberg et al., 2020; Reimand et al., 2007) with GO sources (GO:BP, GO:MF, and GO:CC) to identify the expression of pathways from *SETBP1* target genes for each cluster. We applied the Bonferroni procedure for multiple hypothesis correction and used a p-adjusted threshold of 0.05 and, for the background gene list, included *SETBP1* and its targets.

TF activity

We acquired prior knowledge on the direction of TF regulation from CollecTRI (accessed May 2023)(Müller-Dott et al., 2023) and combined it with GTEx expression TPM to infer TF activity using decoupleR(Badia-I-Mompel et al., 2022) (v.2.6.0). We used a multivariate linear model (run_mlm) with a minimum threshold of 5 targets per TF to calculate activity scores (represented as t-values) for all 758 TFs. We scaled and centered data prior to summarizing the averaged regulator activity for each TF. Positive and negative scores denote TF activity and inactivity, respectively.

Supporting Information

Table_S1.xlsx
Table_S2.xlsx
Doc_S1.docx
"Visualization of TF Activity in Human GTEx Data" Shiny app available at https://lasseignelab.shinyapps.io/gtex_tf_activity/

Acknowledgments

The authors thank the Lasseigne Lab members Vishal Oza, Tabea Soelter, Emma Jones, and Victoria Flanary for their feedback throughout this study. We also thank the UAB Biological Data Science group (RRID:SCR_021766) for providing a script for helping to run containers on the UAB high-performance cluster (https://github.com/U-BDS/training_guides/blob/main/run_rstudio_singularity.sh).

References

Acuna-Hidalgo, R., Deriziotis, P., Steehouwer, M., et al. (2017). Overlapping SETBP1 gain-of-function mutations in Schinzel-Giedion syndrome and hematologic malignancies. *PLoS Genet.* **13**, e1006683.

Albano, F., Anelli, L., Zagaria, A., Coccaro, N., Casieri, P., Minervini, A., & Specchia, G. (2012). SETBP1 and miR_4319 dysregulation in primary myelofibrosis progression to acute myeloid leukemia. *J. Hematol. Oncol.* **5**, 48.

Almotiri, A., Alzahrani, H., Menendez-Gonzalez, J.B., et al. (2021). Zeb1 modulates hematopoietic stem cell fates required for suppressing acute myeloid leukemia. *J. Clin. Invest.* **131**.

Amberger, J.S., Bocchini, C.A., Scott, A.F., & Hamosh, A. (2019). OMIM.org: leveraging knowledge across phenotype-gene relationships. *Nucleic Acids Res.* **47**, D1038–D1043.

Antonyan, L., & Ernst, C. (2022). Putative Roles of SETBP1 Dosage on the SET Oncogene to Affect Brain Development. *Front. Neurosci.* **16**, 813430.

Badia-I-Mompel, P., Vélez Santiago, J., Braunger, J., Geiss, C., Dimitrov, D., Müller-Dott, S., Taus, P., Dugourd, A., Holland, C.H., Ramirez Flores, R.O., & Saez-Rodriguez, J. (2022). decoupleR: ensemble of computational methods to infer biological activities from omics data. *Bioinform Adv* **2**, vbac016.

Balachandran, S., & Narendran, A. (2023). The Developmental Origins of Cancer: A Review of the Genes Expressed in Embryonic Cells with Implications for Tumorigenesis. *Genes* **14**.

Banfi, F., Rubio, A., Zaghi, M., et al. (2021). SETBP1 accumulation induces P53 inhibition and genotoxic stress in neural progenitors underlying neurodegeneration in Schinzel-Giedion syndrome. *Nat. Commun.* **12**, 4050.

Bao, E.L., Cheng, A.N., & Sankaran, V.G. (2019). The genetics of human hematopoiesis and its disruption in disease. *EMBO Mol. Med.* **11**, e10316.

Bhagwat, A.S., & Vakoc, C.R. (2015). Targeting Transcription Factors in Cancer. *Trends Cancer Res.* **1**, 53–65.

Brown, F.C., Still, E., Koche, R.P., et al. (2018). MEF2C Phosphorylation Is Required for Chemotherapy Resistance in Acute Myeloid Leukemia. *Cancer Discov.* **8**, 478–497.

Cardoso-Moreira, M., Halbert, J., Valloton, D., et al. (2019). Gene expression across mammalian organ development. *Nature* **571**, 505–509.

Cardoso-Moreira, M., Sarropoulos, I., Velten, B., Mort, M., Cooper, D.N., Huber, W., & Kaessmann, H. (2020). Developmental Gene Expression Differences between Humans and Mammalian Models. *Cell Rep.* **33**, 108308.

Carratt, S.A., Kong, G.L., Curtiss, B.M., Schonrock, Z., Maloney, L., Maniaci, B.N., Blaylock, H.Z., Baris, A., Druker, B.J., Braun, T.P., & Maxson, J.E. (2022). Mutated SETBP1 activates transcription of Myc programs to accelerate CSF3R-driven myeloproliferative neoplasms. *Blood* **140**, 644–658.

Chang, X., Liu, J., Yang, Q., Gao, Y., Ding, X., Zhao, J., Li, Y., Liu, Z., Li, Z., Wu, Y., & Zuo, D. (2023). Targeting HMGA1 contributes to immunotherapy in aggressive breast cancer while suppressing EMT. *Biochem. Pharmacol.* **212**, 115582.

Cicirò, Y., & Sala, A. (2021). MYB oncoproteins: emerging players and potential therapeutic targets in human cancer. *Oncogenesis* **10**, 19.

Coccaro, N., Tota, G., Zagaria, A., Anelli, L., Specchia, G., & Albano, F. (2017). SETBP1 dysregulation in congenital disorders and myeloid neoplasms. *Oncotarget* **8**, 51920–51935.

Drápela, S., Bouchal, J., Jolly, M.K., Culig, Z., & Souček, K. (2020). ZEB1: A Critical Regulator of Cell Plasticity, DNA Damage Response, and Therapy Resistance. *Front Mol Biosci* **7**, 36.

Elena, C., Gallì, A., Such, E., et al. (2016). Integrating clinical features and genetic lesions in the risk assessment of patients with chronic myelomonocytic leukemia. *Blood* **128**, 1408–1417.

Forrest, M.P., Waite, A.J., Martin-Rendon, E., & Blake, D.J. (2013). Knockdown of human TCF4 affects multiple signaling pathways involved in cell survival, epithelial to mesenchymal transition and neuronal differentiation. *PLoS One* **8**, e73169.

Gu, Z., Eils, R., & Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849.

Han, H., Cho, J.-W., Lee, S., et al. (2018). TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Res.* **46**, D380–D386.

Hekselman, I., & Yeger-Lotem, E. (2020). Mechanisms of tissue and cell-type specificity in heritable traits and diseases. *Nat. Rev. Genet.* **21**, 137–150.

Jansen, N.A., Braden, R.O., Srivastava, S., et al. (2021). Clinical delineation of SETBP1 haploinsufficiency disorder. *Eur. J. Hum. Genet.* **29**, 1198–1205.

Kohyanagi, N., & Ohama, T. (2023). The impact of SETBP1 mutations in neurological diseases and cancer. *Genes Cells* , doi: 10.1111/gtc.13057.

Kolberg, L., Raudvere, U., Kuzmin, I., Vilo, J., & Peterson, H. (2020). gprofiler2 -- an R package for gene list functional enrichment analysis and namespace conversion toolset g:Profiler. *F1000Res.* **9**.

Kolmykov, S., Yevshin, I., Kulyashov, M., Sharipov, R., Kondrakhin, Y., Makeev, V.J., Kulakovskiy, I.V., Kel, A., & Kolpakov, F. (2021). GTRD: an integrated view of transcription regulation. *Nucleic Acids Res.* **49**, D104–D111.

Li, L., Feng, Y., Hu, S., Du, Y., Xu, X., Zhang, M., Peng, X., & Chen, F. (2021). ZEB1 serves as an oncogene in acute myeloid leukaemia via regulating the PTEN/PI3K/AKT signalling pathway by combining with P53. *J. Cell. Mol. Med.* **25**, 5295–5304.

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., & Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* **1**, 417–425.

Liu, W.-L., He, Z.-X., Li, F., Ai, R., & Ma, H.-W. (2018). Schinzel–Giedion syndrome: a novel case, review and revised diagnostic criteria. *J. Genet.* **97**, 35–46.

Lo Surdo, P., Iannuccelli, M., Contino, S., Castagnoli, L., Licata, L., Cesareni, G., & Perfetto, L. (2023). SIGNOR 3.0, the SIGnaling network open resource 3.0: 2022 update. *Nucleic Acids Res.* **51**, D631–D637.

Matsumoto, F., Tohda, A., Shimada, K., & Okamoto, N. (2005). Malignant retroperitoneal tumor arising in a multicystic dysplastic kidney of a girl with Schinzel-Giedion syndrome. *Int. J. Urol.* **12**, 1061–1062.

Morgan, A., Srivastava, S., Duis, J., & van Bon, B. (2021). SETBP1 Haploinsufficiency Disorder. In GeneReviews®, M.P. Adam, H.H. Ardinger, R.A. Pagon, S.E. Wallace, L.J.H. Bean, K.W. Gripp, G.M. Mirzaa, & A. Amemiya, eds. (Seattle (WA): University of Washington, Seattle),.

Müller-Dott, S., Tsirvouli, E., Vázquez, M., Ramirez Flores, R.O., Badia-i-Mompel, P., Fallegger, R., Lægreid, A., & Saez-Rodriguez, J. (2023). Expanding the coverage of regulons from high-confidence prior knowledge for accurate estimation of transcription factor activities.

Nussinov, R., Tsai, C.-J., & Jang, H. (2022). How can same-gene mutations promote both cancer and developmental disorders? *Sci Adv* **8**, eabm2059.

Patnaik, M.M., Itzykson, R., Lasho, T.L., Kosmider, O., Finke, C.M., Hanson, C.A., Knudson, R.A., Ketterling, R.P., Tefferi, A., & Solary, E. (2014). ASXL1 and SETBP1 mutations and their prognostic contribution in chronic myelomonocytic leukemia: a two-center study of 466 patients. *Leukemia* **28**, 2206–2212.

Piazza, R., Valletta, S., Winkelmann, N., et al. (2013). Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. *Nat. Genet.* **45**, 18–24.

Piazza, R., Magistroni, V., Redaelli, S., et al. (2018). SETBP1 induces transcription of a network of development genes by acting as an epigenetic hub. *Nat. Commun.* **9**, 2192.

Reimand, J., Kull, M., Peterson, H., Hansen, J., & Vilo, J. (2007). g:Profiler--a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Res.* **35**, W193–W200.

Schinzel, A., & Giedion, A. (1978). A syndrome of severe midface retraction, multiple skull anomalies, clubfeet, and cardiac and renal malformations in sibs. *Am. J. Med. Genet.* **1**, 361–375.

Wagner, M., Skorobogatko, Y., Pode-Shakked, B., et al. (2020). Bi-allelic Variants in RALGAPA1 Cause Profound Neurodevelopmental Disability, Muscular Hypotonia, Infantile Spasms, and Feeding Abnormalities. *Am. J. Hum. Genet.* **106**, 246–255.

Whitlock, J.H., Soelter, T.M., Howton, T.C., Wilk, E.J., Oza, V.H., & Lasseigne, B.N. (2023). Cell-type-specific gene expression and regulation in the cerebral cortex and kidney of atypical Setbp1S858R Schinzel Giedion Syndrome mice.

Wilks, C., Zheng, S.C., Chen, F.Y., et al. (2021). recount3: summaries and queries for large-scale RNA-seq expression and splicing. *Genome Biol.* **22**, 323.

Xie, J., Xiang, D.-B., Wang, H., Zhao, C., Chen, J., Xiong, F., Li, T.-Y., & Wang, X.-L. (2012). Inhibition of Tcf-4 induces apoptosis and enhances chemosensitivity of colon cancer cells. *PLoS One* **7**, e45617.

Yin, J., Xie, X., Ye, Y., Wang, L., & Che, F. (2019). BCL11A: a potential diagnostic biomarker and therapeutic target in human diseases. *Biosci. Rep.* **39**.

(2015). Schinzel Giedion Syndrome - NORD (National Organization for Rare Disorders) (NORD - National Organization for Rare Disorders).