Transcriptomics Analysis of *ASXL1* truncation, *ASXL1* siRNA, and *VANGL2* Transfected CACO2 Cells

Devon Shao, Isabella Lin, and Valerie A. Arboleda

Affiliation

Department of Pathology & Laboratory medicine, University of California, Los Angeles, California, United States of America

Abstract

ASXL1, located on chromosome 20, is a gene that regulates histone methylation and chromatin modification, and thus gene expression. Bohring-Opitz Syndrome (BOS) is a rare neurodevelopmental disorder caused by de-novo truncating mutations in *ASXL1*. Previous studies have shown that *ASXL1* mutations are associated with broad activation of the canonical Wnt pathway and upregulation of the *VANGL2* gene. VANGL2 is a protein that regulates cell migration and differentiation and tissue morphology during development. This study aimed to explore the relationship between truncating *ASXL1* mutations and *VANGL2*, specifically to elucidate what mechanism causes the previously observed upregulation of *VANGL2* present in BOS patients. We used CACO2 cells transfected with *ASXL1* truncated DNA (n=5), ASXL1 siRNA (n=2), and *VANGL2* DNA and controls (n=7) to replicate cell biology relevant to BOS and thus analyze the transcriptomics effects of exogenous truncated *ASXL1* and overexpression of full-length *VANGL2* in CACO2 cells to provide insights into direct and indirect effects of truncating *ASXL1* mutations using RNA-seq data collection and analysis. The data does not show a clear relationship between the truncated *ASXL1I* cells and the full length *VANGL2* cells, but it does show, through the lack of similarity between *ASXL1* siRNA cells and full length *VANGL2* cells, that upregulation of VANGL2 is not related to a loss of function of *ASXL1*. It also identifies gene targets that may be involved in the BOS phenotype that can be further researched.

Introduction

ASXL1 (additional sex combs-like 1) is a chromatin-binding protein that is critical for normal embryo development in humans(1). The protein affects embryo development by chromatin remodeling and thus regulating the expression of a range of other genes, including Hox genes, which themselves play important roles in embryo development(2). The *Drosophila melanogaster* ortholog of *ASXL1*, *Asx*, is critically involved in anterior-posterior patterning, and embryos with a mutation show incomplete head involution and, in all abdominal segments, show posteriorly directed transformations(3). In mammals, *ASXL1* is necessary in early neural crest, cardiac, and hematopoietic development(4). Additionally, the ASXL1 protein may also be involved in methylation of gene promoters, thus regulating downstream transcription of many genes. Eutherian mammals have 3 homologs of ASX, namely ASXL1, ASXL2, and ASXL3, which share conserved domains, including the C-terminal plant homeodomain that is essential in post translational histone modifications(5). Notably, in the truncating mutations of *ASXL1* found in BOS patients, there is premature truncation of this C-terminal plant homeodomain, which is at the end of the gene in all homologs of ASX.

De novo truncating mutations of *ASXL1* lead to Bohring-Opitz Syndrome, a rare genetic disorder characterized by severe intellectual disability, delayed development, growth failure,

distinctive facial features and posture, and other variable anomalies including heart defects, BOS posture and hypertrichosis anecdotally noted by healthcare providers and parents of BOS patients(6-7). Due to its rarity, prevalence of BOS is unknown, and as of 2018 there were only 46 clinically diagnosed individuals reported, with only 20 of those individuals having their diagnosis molecularly confirmed(7). Truncating mutations leading to BOS in these patients are enriched in exons 12 and 13, the last two exons of the protein(8)(Figure 1A). These last two exons are also the largest exons, together making up about two-thirds of the length of the gene. In this project, the transfected truncated *ASXL1* has been truncated at two different locations creating two different truncations, one of which is in exon 12 and the other which is in exon 13(9). This was done to, in the experimental CACO2 cells, model the cell biology of BOS patients with *ASXL1* truncating.

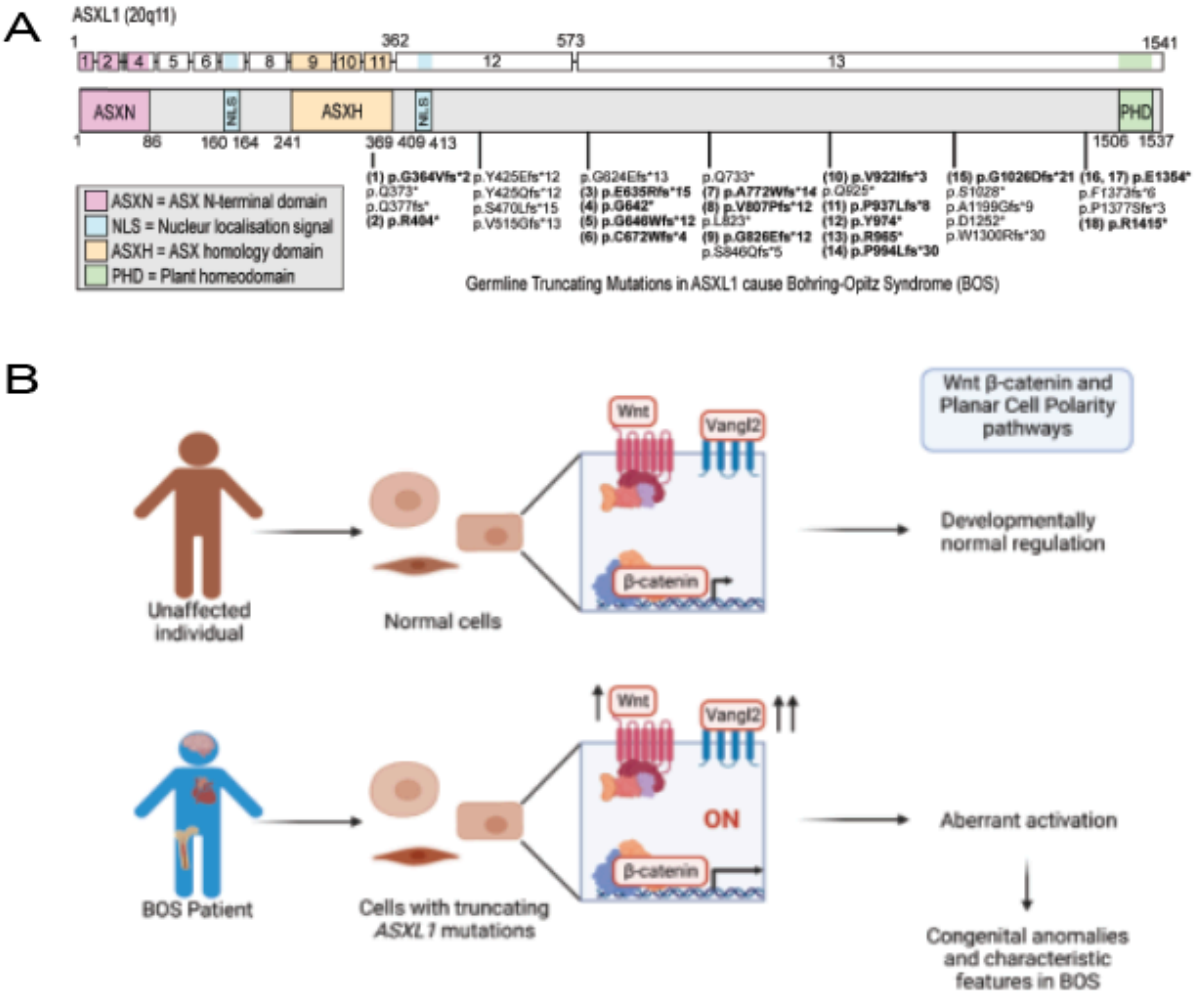


Figure 1

(A) Canonical transcript of *ASXL1*, consists of 13 exons and is located on chromosome 20. Annotations under the chromosome show where truncating mutations occur in patients. Reproduced from Lin et. al., 2023 (4).

(B) Unaffected individuals have normal Wnt and Vangl2 expression levels, which leads to normal developmental regulation, while BOS patients whose cells have truncating *ASXL1* mutations show aberrant activation of Wnt and Vangl2, leading to congenital anomalies and characteristic BOS phenotypes. Reproduced from Lin et. al, 2023 (4).

Somatic mutations in *ASXL1*, that are similar to the germ-line truncating mutations in BOS patients, have been linked to various myeloid malignancies, such as myeloproliferative neoplasms, chronic myelomonocytic leukemia, and acute myeloid leukemia (AML). *ASXL1* mutations have been found to be common in AML and indicate a poor prognosis and therapy outcome(10). Most studies of ASXL1 function in humans are in the context of these myeloid malignancies.

ASXL1 uses polycomb repressive complexes 1 and 2 (PRC1/PRC2) and the polycomb repressive deubiquitinase (PR-DUB) complex to direct histone modifications. These complexes cooperate to regulate signaling pathways such as the Wnt signaling pathway and the Wnt/β-catenin pathway. The canonical Wnt signaling pathway is an essential signaling pathway for cell development, stem cell biology, neural patterning, organogenesis, and regulation of hematopoiesis (11). Wnt proteins are glycoproteins that are involved in signal regulation(12). Major downstream signaling branches to the Wnt pathway have been identified, namely the canonical Wnt/β-catenin dependent pathway and the noncanonical β-catenin-independent pathway, within which exists the Planar Cell Polarity pathway(12).

Van Gogh-like 2 (VANGL2) is a non canonical Wnt and PCP protein involved in these pathways to affect cell migration and tissue patterning. The protein is a membrane protein that is involved in the regulation of planar cell polarity, the development of the neural plate (13), and also regulates Wnt protein distribution(14). The PCP pathway itself is active in the neural system, where it has a role in neuronal maturation and neural complex formation(14) VANGL2 and the noncanonical pathway interacts with the canonical pathway through Dishevelled (DVL) protein activation(4). In a study by Lin et. al. in 2023, in blood and skin fibroblasts of BOS patients with *ASXL1* mutations the Wnt signaling pathway was found to be broadly activated at the transcriptional and protein levels, and VANGL2 was found to be highly upregulated(4). They directly link normal *ASXL1* to repression of noncanonical and canonical Wnt signaling in standard development, and *ASXL1* mutations to increased activation of both the noncanonical and canonical Wnt signaling pathways (Figure 1B).This activity indicates that many of the phenotypes found in BOS patients, such as the neurodevelopmental disorders and heart defects, could be due to the an effect that the *ASXL1* mutation has on VANGL2 levels and the role VANGL2 plays in neural development and general cell migration and tissue patterning.

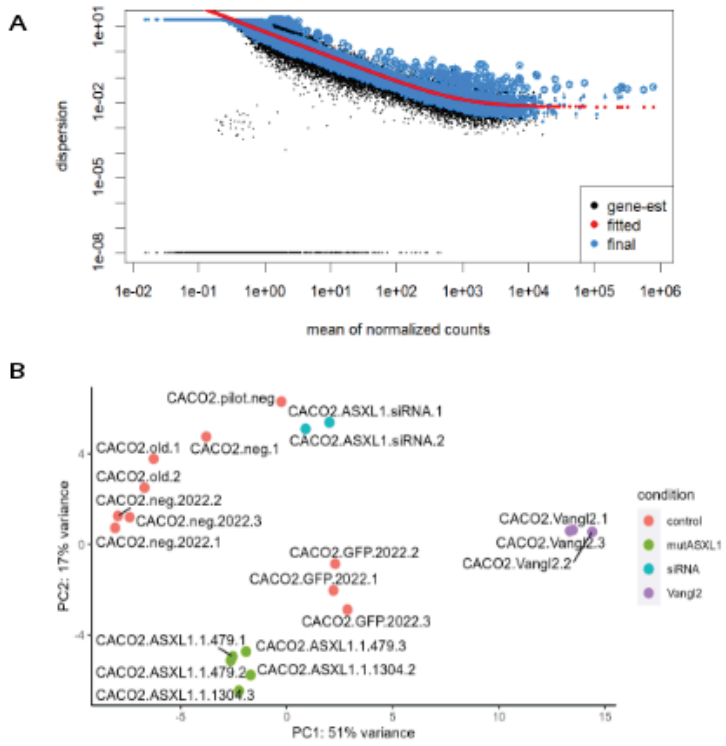
This study seeks to explore the relationship between *ASXL1* truncating mutations and *VANGl2* upregulation in cells in order to understand how both of these genes affect BOS phenotypes and patients. I seek to do this by comparing the transcriptomics of cells transfected with truncated *ASXL1*, with *ASXL1* siRNA, and with *VANGL2* to identify similarities and differences that might elucidate the signaling pathways that ASXL1 and VANGL2 are involved in.

Methods

RNAseq Analysis

Creation of plasmids, transfection of cells, RNA extraction and sequencing and alignment of raw reads against the Gencode human genome was conducted prior to my involvement in the project. Quality control of the cells was conducted using MultiQC. From these processes, I received a list of genes with their expression data for each sample.

Differential gene expression (DGE) analysis was conducted using DESeq2 v1.42.1(15). DGE analysis is a common analysis applied to RNA-sequencing data that allows for quantitative display of differences in gene expression levels between experimental groups/conditions. Genes with an adjusted p-value, using the Benjamini and Hochberg method (BH-adjusted p-values), of less than 0.05 were considered significant. plotDispEsts from DESeq2 was used to plot the per-gene dispersion estimates against the mean of normalized counts. The dispersion of estimates is expected to decrease as the mean of normalized counts increases because the technical variability at low expression levels is made up of mostly technical noise during the measurement process. We see this expected trend in the dispersion estimates graph generated from our data.



C

Sample Name	% Assigned	M Assigned
CACO2-ASXL1-1-1304-1	25.1%	6.5
CACO2-ASXL1-1-1304-2	55.5%	48.2
CACO2-ASXL1-1-1304-3	58.2%	33.6
CACO2-ASXL1-1-479-1	55.0%	41.6
CACO2-ASXL1-1-479-2	56.2%	9.2
CACO2-ASXL1-1-479-3	61.9%	20.6
CACO2-ASXL1-siRNA-1	59.2%	12.3
CACO2-ASXL1-siRNA-2	60.5%	15.9
CACO2-GFP-2022-1	52.0%	28.9
CACO2-GFP-2022-2	55.0%	11.8
CACO2-GFP-2022-3	50.3%	16.6
CACO2-Vangl2-1	52.8%	8.7
CACO2-Vangl2-2	59.4%	15.6
CACO2-Vangl2-3	58.5%	7.1
CACO2-neg-1	50.0%	11.4
CACO2-neg-2	14.7%	2.5
CACO2-neg-2022-1	52.7%	14.5
CACO2-neg-2022-2	55.8%	15.0
CACO2-neg-2022-3	53.8%	9.4
CACO2-neg-3	16.0%	3.6
CACO2-old-1	56.0%	9.8
CACO2-old-2	57.6%	7.0
CACO2-pilot-neg	51.8%	27.7

Figure 2

(A) Dispersion estimates plot, with dispersion plotted against the mean of normalized counts.

(B) PCA of the DESeq2 object normalized counts. Principal component two is plotted against principal component one.

(C) MultiQC report, featurecounts assignment is low for three samples.

Principal component analysis (PCA) was conducted using a customized version of the 'plotPCA' function found in the DESeq2 package. The function was modified to add two options: first, to compute more principal components beyond just the first two provided by the default version of the function, and second, to return a dataframe of values detailing how much each of the genes contributes to the principal components. Both of these modifications use the

provided documentation of the plotPCA in DeSeq2, called with 'DESeq2::plotPCA()'. plotPCA is implemented using 'prcomp', which is the method of conducting PCA in base R contained in the 'stats' package which is by default included in R. The output generated by 'prcomp' has a 'rotation' component, which is a matrix of eigenvectors of every gene. Each of the columns corresponds to a principal component, and each row corresponds to a gene. Each entry in the matrix is a value that indicates how much the gene contributes to the principal component. The 'plotPCA' function modifies the 'prcomp' function to produce the desired output. Our modification of 'plotPCA' utilizes 'prcomp' to expand upon the default capabilities of 'plotPCA'.

Log2 fold change (LFC) shrinkage was conducted on the DESeq object using apeglm. Logarithmic fold change is commonly used to evaluate the effect size of different genes across conditions, despite technical and biological variability in measurements. LFC shrinkage is helpful because commonly there is strong variance of LFC estimates for genes with lower read counts, which leads to unrepresentative results and complicates later data analysis and interpretation. DESeq2 addresses these challenges by shrinking LFC estimates towards zero when the useful information for a gene is low (ie. when gene counts are low, dispersion of the gene is high, or there are less degrees of freedom). The less information is present for a gene, the stronger the shrinkage. The result is that the fold changes from genes with more statistical information and thus less variance are preserved, while the fold changes with more variance are decreased, allowing us to compare LFC more clearly across different trials and experiments(15). Approximate Posterior Estimation for generalized linear model, apeglm, uses heavy-tailed Cauchy prior distribution for effect sizes, which is different from the filter thresholds and pseudocount methods usually employed for LFC shrinkage. The apeglm method, when compared to other shrinkage procedures, has lower bias while accomplishing the same goal of decreasing variance for genes with little statistical information(16).

Boxplots of genes of interest, separated by condition group, were constructed using the ggplot2 package(17). Boxplots were generated based on two different groupings of conditions. The first is based on the transfection groups: mutant *ASXL1*, *VANGL2*, *ASXL1* siRNA, and controls. The second creates groups more specifically by separating the mutant *ASXL1* group into two groups: one group containing the 1-479 truncation cells, and another containing the 1-1394 truncation cells, and also separating the controls into three different groups: a group containing the GFP transfected cells, one containing the standard negative controls, and another containing an older set of control cells. These boxplots are of gene expression of a specific gene, allowing us to see the specific differences between different treatment groups. Genes of interest were chosen based on results from PCA, volcano plots, and previous data, especially RNAseq data from blood and fibroblast cell samples taken from BOS patients. Counts of the genes were taken from normalized, LFC-shrunk counts.

Heatmaps for the three different treatment groups were created from 300 high LFC significant DEGs using normalized counts of the DESeq object. Heatmaps are a way to visualize gene expression data, especially obvious differences between different treatment groups, The pheatmap R package was used to create the heatmaps. Volcano plots for each treatment group were generated using ggplot to visualize genes that pass a certain significance (p-value) and LFC threshold. Gene ontology enrichment of the DEGs was accomplished using the R package clusterProfiler(18). Gene ontology analysis allows identification of which biological processes, molecular functions and/or cellular components the most expressed significant genes are involved in. Specifically, the 'enrichGO' function from clusterProfiler was used to generate the gene ontology enrichment, and redundant gene ontologies were simplified using

clusterProfiler::simplify. Significantly enriched gene ontologies were determined when their p-adjusted value was less than 0.05. Gene ontologies were represented as bar charts and as a graph with nodes and edges, where each node represents a gene ontology and the edges show how related different gene ontologies are to each other.

Results

After QC analysis, three samples, one mutant *ASXL1* sample and two control samples, were excluded due to contamination with COVID reads. The discrepancy was first noticed from PCA plots and the heatmap of the LFC of mutant *ASXL1*, and verified from the multiQC results.

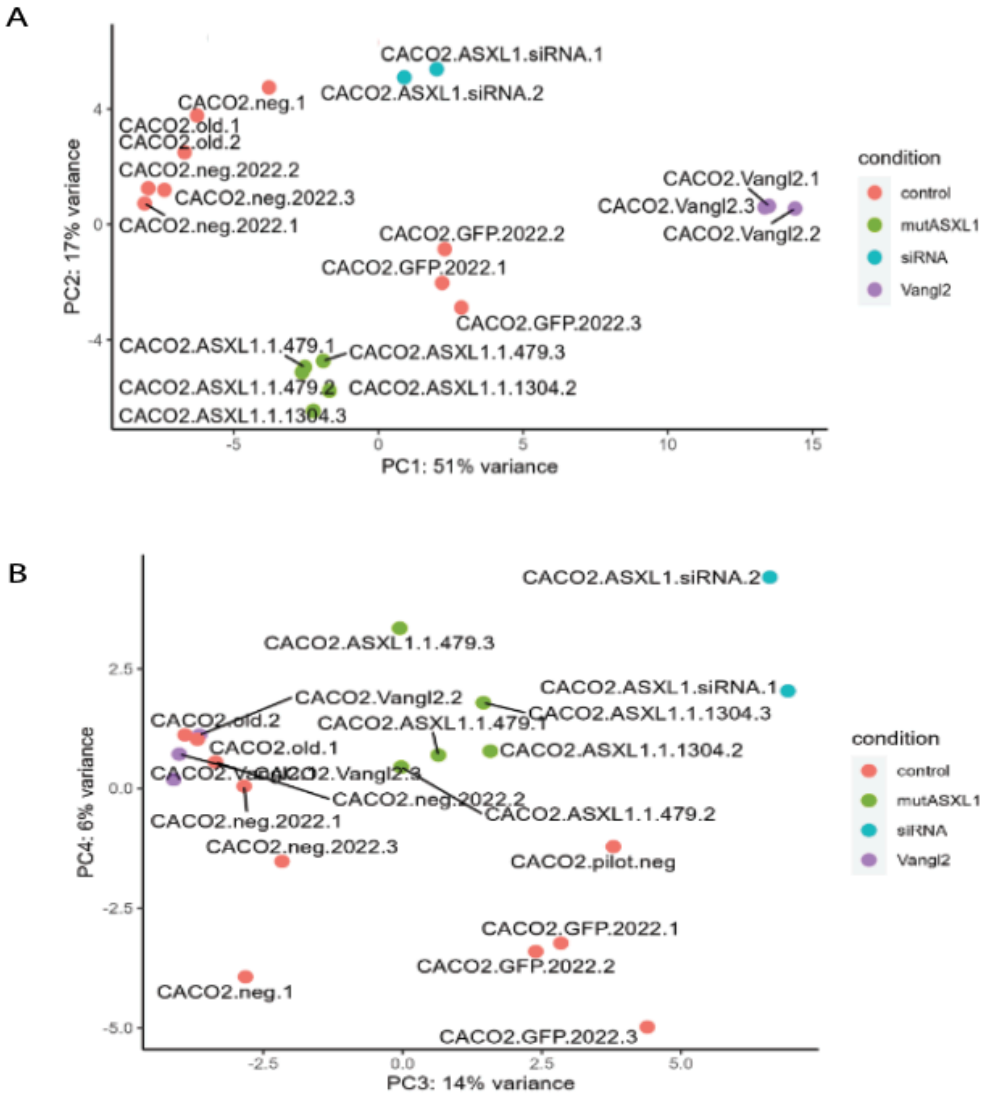


Figure 3

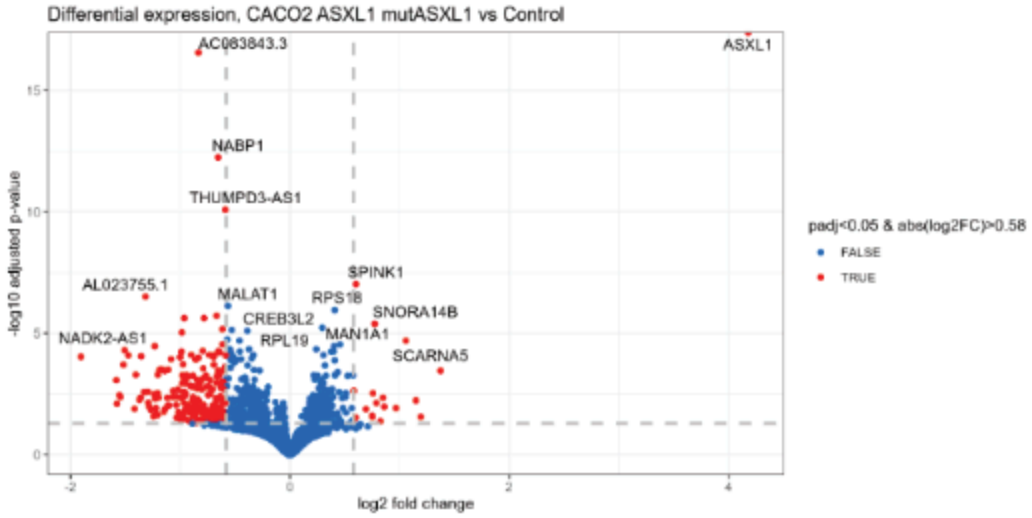
(A) Plot of principal component two against principal component one.
(B) Plot of principal component four against principal component three.

The PCA analysis displays that the majority of the variance is driven by principal component one, which dictates 51 percent of the variance, while principal components two and three drive 17 and 14 percent respectively. Principal component four drives just six percent, and subsequent principal components continue to decrease in contribution to variance. When the samples are plotted with principal component two against principal component one, there is a clear clustering of the different condition groups. The treatment groups separate along both axes. Upon further examination into the genes that contribute most to the variance of each of these principal components, I found that principal component one was mainly driven by VANGL2 expression and principal component two was mainly driven by ASXL1 expression. Because of the separation seen between all of the groups, it is not immediately clear what the relationship between truncated *AXL1* and increased VANGL2 expression is. Interestingly, the GFP controls grouped together and were similar along principal component one as the siRNA *ASXL1* transfection group, indicating that the transfection process itself may have affected the gene expression of the cells in a way that *in vivo* cells would not be affected by.

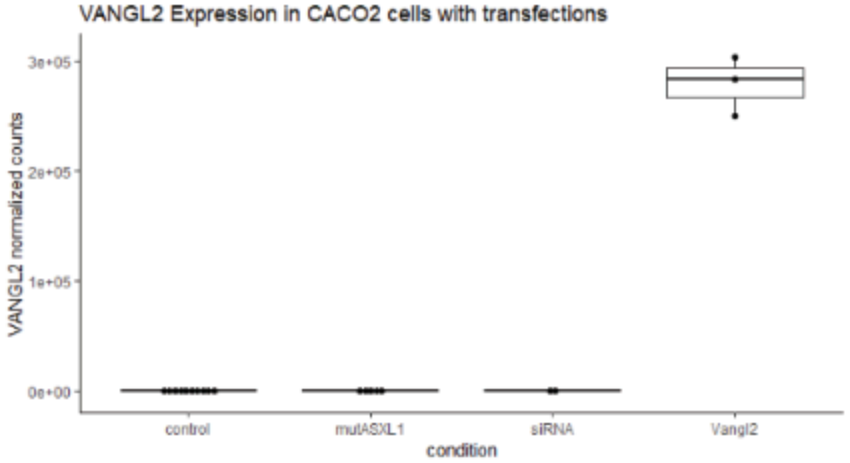
The volcano plot for the data shrunk for mutant *ASXL1* shows, out of the significant genes, the ones with the highest LFC. Notably, ASXL1 is very significant and is highly expressed in the truncated *ASXL1* cells. Also of note is that VANGL2 is not seen here. While Lin et. al. observed high expression of VANGL2 in blood and skin fibroblasts of BOS patients with *ASXL1* truncating mutations, in these *in vitro* CACO2 cells with *ASXL1* truncating mutations, VANGL2 is not observed to be highly expressed. To verify this, I generated a boxplot of VANGL2 normalized accounts across the different treatment groups. The boxplot shows that VANGL2 expression is low in control, mutant *ASXL1*, and *ASXL1* siRNA cells, while it is high in *VANGL2* overexpression cells(Figure 4C). This result does not elucidate the relationship between truncating mutations in *ASXL1* and VANGL2 expression. If anything, it suggests that the upregulation of VANGL2 in cells with truncating *ASXL1* mutations is not due to a gain of function or increased expression of ASXL1.

Besides the relationship between the ASXL1 expression and VANGL2 expression, there are also some notable significant dysregulated genes in these *ASXL1* truncation cells, which may be involved in BOS. Some of these genes are SPINK1, SNORA14B, SCARNA5, NADK2,AS1, AL023755.1, and AC083843.3 (Figure 4A).

A


Differential expression, CACO2 ASXL1 mutASXL1 vs Control
AC083843.3
ASXL1
NABP1
THUMPD3-AS1
SPINK1
AL023755.1
MALAT1
RPS18
SNORA14B
CREB3L2
NADK2-AS1
RPL19
MAN1A1
SCARNA5
-log10 adjusted p-value
log2 fold change
padj<0.05 & abs(log2FC)>0.58
FALSE
TRUE

B


VANGL2 Expression in CACO2 cells with transfections
VANGL2 normalized counts
condition
control
mutASXL1
siRNA
Vangl2
mutASXL1_OE: log2FC=0.03, padj=0.81
Vangl2_OE: log2FC=11.74, padj=0
ASXL1_siRNA: log2FC=0.01, padj=0.98

C

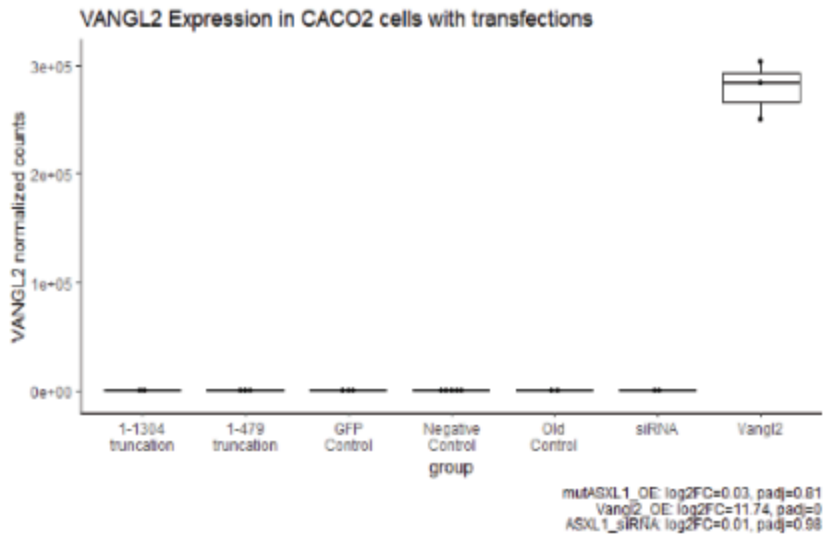

VANGL2 Expression in CACO2 cells with transfections
VANGL2 normalized counts
group
1-1304 truncation
1-479 truncation
GFP Control
Negative Control
Old Control
siRNA
Vangl2
mutASXL1_OE: log2FC=0.03, padj=0.81
Vangl2_OE: log2FC=11.74, padj=0
ASXL1_siRNA: log2FC=0.01, padj=0.98

Figure 4

(A) Volcano plot of p-value plotted against log2 fold change for LFC shrunk mutant *ASXL1* versus control.

(B) Box plot of VANGL2 expression in CACO2 cells separated by treatment groups (control, mutant *ASXL1*, *ASXL1* siRNA and *VANGl2*)

(C) Box plot of VANGL2 expression in CACO2 cells separated by more specific treatment groups. The mutant *ASXL1* group is separated based on length of truncation, and control groups are separated by type of control.

From inspecting the volcano plot constructed from differential expression of the *ASXL1* siRNA treatment group, there is widespread significant downregulation of many genes. This is expected because ASXL1 regulates many genes involved in development, and is also involved in methylation at the promoters of many genes. There are not many genes that are highly upregulated in the *ASXL1* siRNA treatment group, which again corresponds to the expected function of ASXL1 and what would happen if that function was removed due to siRNA. VANGL2 is not highly upregulated or downregulated in this treatment group, which would seem to suggest that high expression levels of VANGL2 is not due to a loss of function of *ASXL1*, because the gene expression levels of *ASXL1* siRNA transfected cells are not similar to those in *VANGL2* transfected cells.

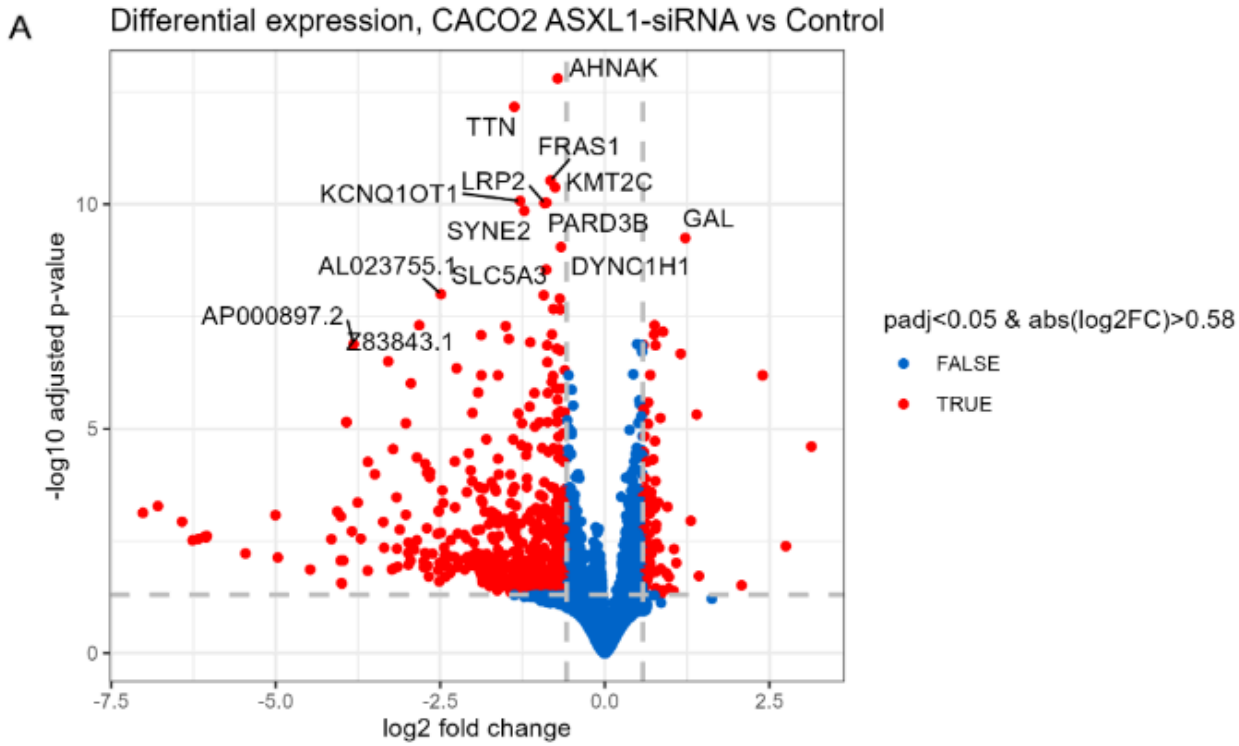


Figure 5

(A) Volcano plot of p-value plotted against log2 fold change for LFC shrunk *ASXL1* siRNA versus control.

One particular area of interest within the mutant *ASXL1* group is if there is a difference between the gene expression of "early" versus "late" truncations, as among the five samples that passed QC in this group, three are "early" truncations, from base pair one to base pair 479, while the two others are "late" truncations, from base pair one to base pair 1304. By separating these two groups in the barplot generation, we can see the difference in expression of specific genes between the early and late truncations. Of note is the ASXL1 gene. It is slightly higher in the late truncation, but not notably different. The median of the box plots are similar, while the range of the early truncating mutation is slightly lower than the range of the late truncating mutation. VANGL2 expression between the early and late truncations is also similar, with both of them being close to zero. Again, the very low level of VANGL2 expression even when these two lengths of truncations are analyzed separately shows that the upregulation of VANGL2 observed in the blood and fibroblast cells by Lin et. al. is not present in these transfected CACO2 cells. Very little can be concluded about the mechanisms behind VANGL2 overexpression in *ASXL1* truncated cells because of these results.

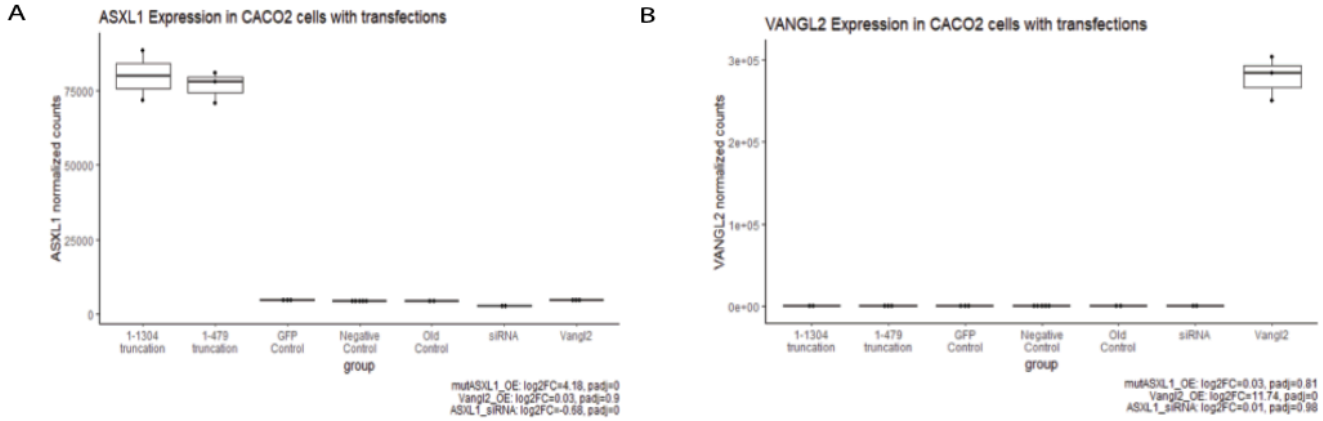


Figure 6

(A) Box plot of ASXL1 expression in CACO2 cells separated by more specific treatment groups. The mutant *ASXL1* group is separated based on length of truncation, and control groups are separated by type of control.

(B) Box plot of VANGl2 expression in CACO2 cells separated by more specific treatment groups. These groups are specified in (A).

Discussion

The increased expression of VANGL2 that was previously observed in blood and fibroblast cells from BOS patients with truncating *ASXL1* mutations, and of which this study sought to identify the mechanism behind, was not observed in these transfected CACO2 cells in neither the siRNA *ASXL1* transfection group nor in either of the early or late truncated *ASXL1* mutation transfection group. Due to this result and lack of similarities in the gene expression between either mutant *ASXL1* or siRNA *ASXL1* and *VANGL2* transfection groups, it is difficult to identify any potential mechanisms for action for ASXL1 specifically in relation to VANGL2. However, it is significant that the truncated *ASXL1* CACO2 cells had such a high LFC compared

to control cells because it allows, among the genes that were significantly dysregulated in the truncated *ASXL1* transfected cells, identification of other genes that are dysregulated and thus may possibly be involved in BOS phenotypes.

The high expression of ASXL1 in mutant *ASXL1* transfected cells may not be a direct result of the truncating mutation itself. CACO2 cells were chosen in this study because ASXL1 antibodies are not very sensitive, leading to poor experimental results when using other cells. CACO2 cells had natural high expression of ASXL1, and thus was better to measure the effect of the truncating mutations. Thus the high LFC of ASXL1 in the mutant *ASXL1* treatment group thus could have been a result of the change in cell type rather than a result of the truncating mutation itself. An ideal cell type for the experiment would have been blood or fibroblast cells in order to allow more direct comparison to the blood and fibroblast cell samples taken directly from BOS patients.

Another possible limitation in this experiment is the effect that transfection had on expression. From the PCA, it is clear that the GFP transfection cells separate from the other controls based on a number of genes. This may have affected the efficacy of the following DESeq2 analysis because the LFC was derived from comparing the expression of each treatment group against all the controls, and since the GFP control cells had different gene expression from the rest of the controls, this may skew the resulting LFC. On a functional level, the difference in gene expression between GFP transfection cells and the untreated controls indicates that the transfection process itself may affect gene expression in these cells, which would add a further variable that is not controlled for when comparing gene expression between blood and fibroblast cells taken directly from patients and transfected experimental cells grown *in vitro*.

The major limitation is in the difference that cell type causes. Follow up analysis could be done utilizing publicly available datasets on RNA-seq results of *VANGL2* transfections on different cell types to explore how the cell type affects VANGL2 expression. Furthermore, an ideal cell type to conduct transfections on would be blood or fibroblast cells. This would have allowed the resulting gene expression data to be more easily comparable to the gene expression data from patient sample blood and fibroblast cells. Another possible experimental design would be to isolate blood and fibroblast cells from the patient and introduce siRNA to a portion of the cells to observe how the added siRNA changes gene expression in those cells over time.

References

1. U.S. National Library of Medicine. (n.d.-b). *ASXL1 ASXL transcriptional regulator 1 [homo sapiens (human)] - gene - NCBI*. National Center for Biotechnology Information. https://www.ncbi.nlm.nih.gov/gene/171023
2. U.S. National Library of Medicine. (n.d.). *ASXL1 gene: Medlineplus genetics*. MedlinePlus. https://medlineplus.gov/genetics/gene/asxl1/
3. Sinclair, D. A., Campbell, R. B., Nicholls, F., Slade, E., & Brock, H. W. (1992). Genetic analysis of the additional sex combs locus of Drosophila melanogaster. Genetics, 130(4), 817–825. https://doi.org/10.1093/genetics/130.4.817
4. Lin, I., Wei, A., Awamleh, Z., Singh, M., Ning, A., Herrera, A., REACH Biobank and Registry, Russell, B. E., Weksberg, R., & Arboleda, V. A. (2023). Multiomics of Bohring-Opitz syndrome truncating ASXL1 mutations identify canonical and noncanonical Wnt signaling dysregulation. JCI insight, 8(10), e167744. https://doi.org/10.1172/jci.insight.167744
5. Aravind L, Iyer LM. The HARE-HTH and associated domains: novel modules in the coordination of epigenetic DNA and protein modifications. Cell Cycle. 2012;11(1):119–131.
6. Russell B, Tan WH, Graham JM Jr. Bohring-Opitz Syndrome. 2018 Feb 15. In: Adam MP, Feldman J, Mirzaa GM, et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2024. Available from: https://www.ncbi.nlm.nih.gov/books/NBK481833/
7. Russell B, et al. Bohring-Opitz Syndrome. In: Adam MP, et al eds. GeneReviews. University of Washington, Seattle; 2018:NBK481833.
8. Katoh M. Functional proteomics of the epigenetic regulators ASXL1, ASXL2 and ASXL3: a convergence of proteomics and epigenetics for translational medicine. Expert Rev Proteomics. 2015;12(3):317–328.

9. Balasubramani A, Larjo A, Bassein JA, Chang X, Hastie RB, Togher SM, Lähdesmäki H, Rao A. Cancer-associated ASXL1 mutations may act as gain-of-function mutations of the ASXL1-BAP1 complex. Nat Commun. 2015 Jun 22;6:7307. doi: 10.1038/ncomms8307. PMID: 26095772; PMCID: PMC4557297.
10. Pratcorona M, Abbas S, Sanders MA, Koenders JE, Kavelaars FG, Erpelinck-Verschueren CA, Zeilemakers A, Löwenberg B, Valk PJ. Acquired mutations in ASXL1 in acute myeloid leukemia: prevalence and prognostic value. Haematologica. 2012 Mar;97(3):388-92. doi: 10.3324/haematol.2011.051532. Epub 2011 Nov 4. PMID: 22058207; PMCID: PMC3291593.
11. Hagiwara, A., Yasumura, M., Hida, Y., Inoue, E., & Ohtsuka, T. (2014). The planar cell polarity protein Vangl2 bidirectionally regulates dendritic branching in cultured hippocampal neurons. Molecular brain, 7, 79. https://doi.org/10.1186/s13041-014-0079-5
12. Komiya Y, Habas R. Wnt signal transduction pathways. Organogenesis. 2008 Apr;4(2):68-75. doi: 10.4161/org.4.2.5851. PMID: 19279717; PMCID: PMC2634250.
13. U.S. National Library of Medicine. (n.d.-c). Vangl2 VANGL planar cell polarity protein 2 [homo sapiens (human)] - gene - NCBI. National Center for Biotechnology Information. https://www.ncbi.nlm.nih.gov/gene/57216
14. Brunt, L., Greicius, G., Rogers, S., Evans, B. D., Virshup, D. M., Wedgwood, K. C. A., & Scholpp, S. (2021). Vangl2 promotes the formation of long cytonemes to enable distant Wnt/β-catenin signaling. Nature communications, 12(1), 2058. https://doi.org/10.1038/s41467-021-22393-9
15. Love MI, Huber W, Anders S (2014). "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2." Genome Biology, 15, 550. doi:10.1186/s13059-014-0550-8.
16. Anqi Zhu, Joseph G Ibrahim, Michael I Love, Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences, Bioinformatics, Volume 35, Issue 12, June 2019, Pages 2084–2092, https://doi.org/10.1093/bioinformatics/bty895
17. H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

18. T Wu, E Hu, S Xu, M Chen, P Guo, Z Dai, T Feng, L Zhou, W Tang, L Zhan, X Fu, S Liu, X Bo, and G Yu. clusterProfiler 4.0: A
19. Universal enrichment tool for interpreting omics data. The Innovation. 2021, 2(3):100141