HHS Public Access

Author manuscript

Gastro Hep Adv. Author manuscript; available in PMC 2023 December 13.

Published in final edited form as:

Gastro Hep Adv. 2023 ; 2(8): 1077–1081. doi:10.1016/j.gastha.2023.08.008.

Evaluating Performance of IsoformSwitchAnalyzeR and mRNA Isoform Switching in Small Intestine Epithelial Differentiation

R. AITA[1], **L. CHEN**[1,2], **M. P. VERZI**[1]

[1]Department of Genetics, Human Genetics Institute of New Jersey, Rutgers Cancer Institute of New Jersey, Rutgers Center for Lipid Research, New Jersey Institute for Food, Nutrition & Health, Division of Environmental & Population Health Biosciences, EOHSI, New Brunswick, New Jersey

[2]School of Life Science and Technology, Key Laboratory of Developmental Genes and Human Disease, Southeast University, Nanjing, China

Adiverse set of molecular mechanisms contribute to the control of homeostatic differentiation in the intestinal epithelium, maintaining a balance of cells in proliferative crypts and on differentiated villi. However, genome-wide alternate messenger RNA isoform usage, also known as mRNA isoform switching, is largely understudied in the field of small intestinal epithelial homeostasis. Isoform switching can impact the stability, subcellular targeting, or functional domains encoded in an mRNA transcript. The study of genome-wide isoform switching may provide new insights into important intestinal health issues including inflammatory bowel disease, metabolic syndromes, malnutrition, and cancer. Here, we report that relatively small numbers of isoform switching events are apparent during differentiation of the intestinal epithelium and that there appear to be limitations in current isoform switching analysis packages in calling true positives in our system.

The Bone Morphogenetic Protein (BMP) signaling pathway is well-documented to promote cellular differentiation in the intestine.[1,2] Here, we catalog genome-wide isoform switching events in RNA-seq data from across the (1) duodenal crypt-villus axis in the epithelial fractions of mouse intestine (GSE133949)[3] (Figure 1A) and (2) in response to BMP treatment of duodenal organoids (GSE232589) (Figure 1A). These differentiation data sets were then validated using a third data set derived from flow-sorted intestinal stem

Correspondence: Address correspondence to: Michael P. Verzi, PhD, Rutgers University, Life Sciences Building, 145 Bevier Rd., Piscataway, New Jersey 08854. mv347@hginj.rutgers.edu.

Supplementary Materials

Material associated with this article can be found in the online version at https://doi.org/10.1016/j.gastha.2023.08.008.

Conflicts of Interest:

The authors disclose no conflicts.

Ethical Statement:

The corresponding author, on behalf of all authors, jointly and severally, certifies that their institution has approved the protocol for any investigation involving humans or animals and that all experimentation was conducted in conformity with ethical and humane principles of research.

Reporting Guidelines:

ARRIVE/Care and Use of Laboratory Animals.

cells (ISCs), transit amplifying (TA) cells, and adult enterocytes (AEs) from mouse small intestine (GSE127488)[4] (Figure 1A).

To identify mRNA isoform switching that occurs during epithelial differentiation in the intestine, we utilized genome-wide alternative isoform analysis (as outlined in the Supplementary Methods section). In brief, isoform switches were identified by using Kallisto-aligned RNA-seq data and IsoformSwitchAnalyzeR (v1.19.0) and DEXSeq (v1.43.0) R packages (Figure 1B). The underlying principle of this workflow is to statistically analyze transcript-based isoform switches and identify functional consequences of said switches.

A Venn diagram indicates the number of mRNA isoform switches called in each data set by IsoformSwitchAnalyzeR (Figure 1C). Following a robust set of filtering as outlined in Figure 1C, the comparison of crypts and villi yielded 139 isoform switches, BMP2 ligand-treatment of organoids yielded 302 isoform switches, the comparison of ISCs and AEs yielded 656 isoform switches, and the comparison of TA cells and AEs yielded 555 isoform switches. To better understand the isoform-switch calls made by IsoformSwitchAnalyzeR, we visually inspected the raw data traces of the top 25 calls in each data set (ranked by q value). Qualitatively, only a minority of the calls were easily identified as isoform switching events when inspecting the raw data, indicating either limitations of this analysis package or reflecting the relatively small number of isoform switching events that occur during epithelial differentiation in the intestine (Figure 1C). Nevertheless, clear isoform switching events were identified, with 4 genes exhibiting isoform switches across all data sets analyzed (*E2f3, Mxi1, Nfic, Gins2*).

With this in mind, we demonstrate *E2f3 as* an example of a significantly curated isoform switch (Figure 2A). Within *E2f3*, isoform ENSMUST00000102948.11 undergoes significant differential isoform usage in all 4 experimental pairwise comparisons and closely resembles E2f3a (Figure 2A), which acts as a transcriptional activator.[7] The promoter of the *E2f3a* isoform is more enriched in the less differentiated conditions in each data set comparison (Figure 2A) and is consistent with *E2f3a* promoting cellular proliferation in other systems and tissues.[7,8] Interestingly, the *E2f3b*-resembling transcript isoform was only called a significant switch in the crypts vs villi data as per IsoformSwitchAnalyzeR (v1.19.0) (Figure 2A). Integrated Genomics Viewer visualization of the data confirms that the promoter of *E2f3b* is enriched in the differentiated or villus-like conditions in each data set, consistent with a role in cellular differentiation and in mediating opposing effects of *E2f3a* (Figure 2A). This opposing effect is again supported by how *E2f3b* promotes differentiation in muscle cells.[9] Taken together, we hypothesize that *E2f3a* promotes proliferation in the crypts, while *E2f3b* antagonistically promotes a post-mitotic state in the villus.

Utrn is another example of a curated isoform switch event (Figure 2B). Previously, alternative promoter usage of *Utrn* has been shown to lead to altered protein isoform localization in a report investigating Duchenne muscular dystrophy.[10] ENSMUST00000076817.5 (*Utrn*-LF) and ENSMUST00000219003.2 (*Utrn*-SF) are 2 transcript isoforms undergoing a statistically significant isoform switch (q value < 0.001) in 3 conditions: Veh vs BMP2 treated organoids, TA cells vs AEs, and ISCs vs AEs.

We see increased expression of 3′ exons corresponding to the UTRN-SF isoform in the crypt-like or proliferative conditions (Figure 2B). It will be interesting to understand whether *Utrn* isoforms have distinct functions in intestinal differentiation and how the expression of these isoforms as potential therapeutics for Duchenne's muscular dystrophy[11] could impact intestinal functions.

The diverse number of RNA-seq analysis tools enables easier genome-wide differential expression analyses, but there has been a lag in research output for genome-wide alternative splicing studies. Isoform switch calling algorithms are in the early days of use, and different tools can yield rather distinct outcomes. For example, a study on garden pea roots in symbiosis with fungi yielded only 8 common genes among 3 different alternative splicing tools (DRIMSeq, SUPPA2, and IsoformSwitchAnalyzeR); the 8 common genes were about 19.0% of the total number of genes that showed differential splicing by DRIMSeq (42 total genes), 1.6% of the total number of genes that showed differential splicing by SUPPA2 (507 total genes), and 27.6% of the total number of genes that showed differential splicing by IsoformSwitchAnalyzeR (29 total genes).[12] More recently, in a multi-center cohort study on human patients with sepsis, the authors report that 1215 isoforms of 627 genes undergo at least one isoform switch in association with at least one sepsis type, but only 50 statistically significant and unique isoform switching genes are actually common between IsoformSwitchAnalyzeR and rMATS.[13] There is clearly limited literature on comparing genome-wide alternative splicing tools, especially in the field of intestinal biology and in mouse models, and this represents an untapped field of transcriptomics.

Ultimately, we find that IsoformSwitchAnalyzeR might give limited true isoform switching genes because (1) there are many false positives we found within the top 25 subsets of each pairwise comparison through qualitative visual validation and (2) there is no clear concordance among the tools available for identification of genome-wide isoform switching.[12,13] We hope our report both identified the existence of exciting isoform switching events during intestinal differentiation but will also serve as a call to attention on the lag that exists between the genome-wide differential expression analyses and genome-wide alternative splicing analyses that are currently available.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

The authors of this study would like to thank the Office of Advanced Research Computing at Rutgers University for providing invaluable resources such as the Amarel high performance computing cluster. The Verzi lab is supported as members of the NIEHS Center for Environmental Exposures and Disease (CEED), Rutgers EOHSI.

Funding:

This research was supported by the National Institutes of Health National Institute of Diabetes and Digestive and Kidney Diseases with the parent grant R01DK121915 funded to Dr Michael Verzi and the affiliated Diversity Supplement Program grant funded to Rohit Aita. The Verzi lab is also supported by R01DK126446 and from the Cancer Center Support Grant (CCSG, P30CA072720) from the National Cancer Institute.

Data Transparency Statement:

RNA-seq data exist already in the form of the National Center for Biotechnology Information Gene Expression Omnibus (NCBI GEO) for the crypts and villi data set (GSE133949) and the small intestinal validation data set (GSE127488). For the duodenal crypt organoids data set, RNA-seq data will be made available as a new submission to the NCBI GEO database (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE232589, token: uzcbusgkfjmptmf).

Abbreviations used in this paper:

AEs	adult enterocytes
BMP	BoneMorphogeneticProtein
ISCs	intestinal stem cells
TA	transit amplifying

References

1. Beumer J, et al. Cell Rep 2022;38: 110438. [PubMed: 35235783]
2. Chen L, et al. Nat Genet 2019; 51:777–785. [PubMed: 30988513]
3. Aita R, et al. J Biol Chem 2022;298: 102213. [PubMed: 35779631]
4. Kazakevych J, et al. Genome Biol 2020;21:64. [PubMed: 32160911]
5. Sprangers J, et al. Cell Death & Differentiation 2021;28:95–107. [PubMed: 33208888]
6. Oliveros JC. Venny. An interactive tool for comparing lists with Venn's diagrams; 2007–2015.
7. Chong J-L, et al. Mol Cell Biol 2009;29:414–424. [PubMed: 19015245]
8. Danielian PS, et al. Oncogene 2008;27:6561–6570. [PubMed: 18663357]
9. Asp P, et al. Genes Dev 2009; 23:37–53. [PubMed: 19136625]
10. Perkins KJ, et al. FEBS Lett 2018; 592:1856–1869. [PubMed: 29772070]
11. Baby SM, et al. Brain Pathol 2010; 20:323–342. [PubMed: 19486009]
12. Zorin EA, et al. Plants (Basel) 2020; 9:1700. [PubMed: 33287282]
13. Chen L, et al. Sci Rep 2022;12: 10276. [PubMed: 35715539]

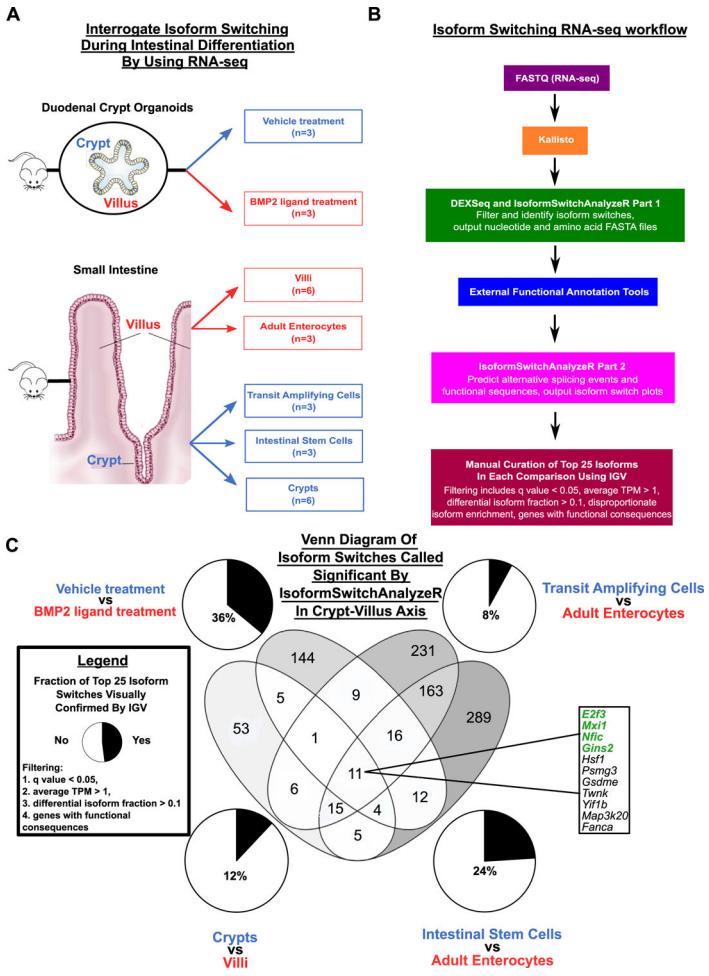

Figure 1.

(A) Three experimental designs are shown to interrogate isoform switching during intestinal epithelial differentiation: (1) duodenal crypt organoids (available on GSE232589) were cultured and treated for 24 hours on day 6 with either vehicle (n = 3) or 100ng/ml of BMP2 ligand (n = 3), (2) duodenal crypts (n = 6) or villi (n = 6) were isolated, or (3) adult enterocytes (n = 3) were sorted from murine small intestinal villi, while transit amplifying cells (n = 3) and intestinal stem cells (n = 3) were sorted from murine small intestinal crypts.[4] Organoid figure adapted from Sprangers et al 2021.[5] Crypt-villus figures are adapted from Wikimedia Commons [https://tinyurl.com/mrdjun7h]. The crypt-villus data are available on GSE133949.[3] The validation data set on the adult enterocytes, transit amplifying cells, and intestinal stem cells is available on GSE127488.[4] (B) Analysis workflow is depicted. Kallisto was used to align the paired-end RNA-seq generated FASTQ files using pseudoalignment to the GENCODE vM29 (mm39) annotation. IsoformSwitchAnalyzeR and DEXSeq were then used to identify isoform switches and output nucleotide and amino acid FASTA files. A set of external tools were then used to assess coding potential, protein domains, signal peptides, and intrinsically disordered regions. After annotation of the transcripts, IsoformSwitchAnalyzeR was utilized again to perform alternative splicing analysis and functional annotation, as well as generate isoform switch plots. Visualization of the merged bigwig output of the

Spliced Transcripts Alignment to a Reference RNA-seq workflow was done using Integrated Genomics Viewer (IGV). Additional analysis details can be found in Supplemental Methods. We curated the top 25 transcript isoform switching events (based on q value < 0.05, average transcript per million > 1, absolute differential isoform fraction > 0.1, disproportionate isoform enrichment, and genes with functional consequences) outputted from IsoformSwitchAnalyzeR part 2, using IGV output. (C) Pie charts indicate the fraction of called isoform switches that passed manual curation. A Venn (v2.1.0)[6] diagram was constructed using the ranked isoform switch lists filtered for absolute differential isoform fraction > 0.1, average transcript per million > 1, q value < 0.05, and genes with functional consequences (meaning isoforms with different coding or regulatory potential) for each of the 4 experimental comparisons mentioned to identify consensus isoform switching across all data sets. 4 out of the 11 common genes *(E2f3, Mxi1, Nfic, and Gins2)* across all 4 experimental group comparisons were validated using IGV in at least 1 comparison.

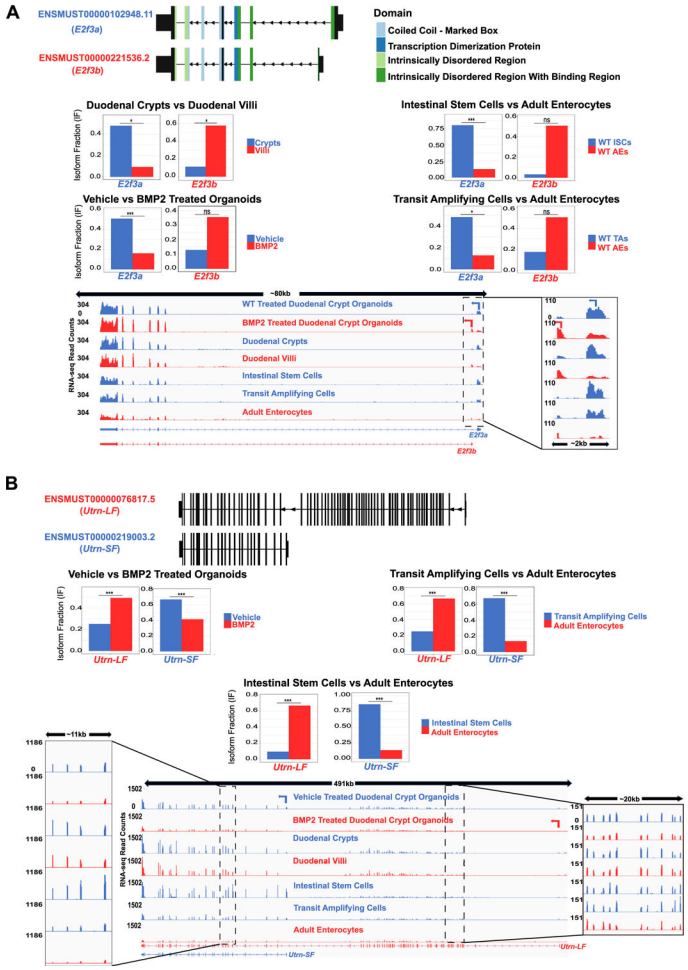

Figure 2.

(A) Two transcript isoforms (ENSMUST00000102948.11, ENSMUST00000221536.2) are shown from the corresponding isoform switch plot from IsoformSwitchAnalyzeR. Based on a previous report by Chong et al,[7] we think that ENSMUST00000102948.11 resembles *E2f3a*, whereas ENSMUST00000221356.2 resembles *E2f3b*. Both variants contain the coiled coil–marked box domain, transcription factor dimerization partner (TDP) domain, and 2 kinds of intrinsically disordered regions. Underneath the 2 isoform variants, there are isoform usage plots based on the IsoformSwitchAnalyzeR isoform fraction metric for the indicated comparisons. Bottom, 80 kilobases of the *E2f3* locus show Spliced Transcripts Alignment to a Reference-aligned RNA-seq bigwigs (see Supplemental Methods) on IGV from all sets of data in this report. Blue indicates the crypt-like or proliferative conditions, whereas the red indicates the villus-like or differentiated conditions. The inset of 2 kilobases near the promoter regions of both *E2f3a* and *E2f3b* highlights the alternative promoter usage. (B) Two transcript isoforms (ENSMUST00000076817.5, ENSMUST00000219003.2) from the *Utrn* locus are shown from the corresponding isoform switch plot from IsoformSwitchAnalyzeR. For the purpose of this report, ENSMUST00000076817.5 is called UTRN-LF (UTRN-long form) while ENSMUST00000219003.2 is called UTRN-SF (UTRN-short form). Underneath the isoform schematics, isoform usage plots from IsoformSwitchAnalyzeR indicate relative isoform fraction expression in each comparison.

491 kilobases of the *Utrn* locus are depicted via IGV. Blue indicates the crypt-like or proliferative role conditions, whereas the red indicates the villus-like conditions. The leftmost inset shows approximately 11 kilobases worth of disproportionately enriched exons toward the 3′ region of both isoforms. The rightmost inset shows approximately 20 kilobases worth of similarly enriched exons toward the 5′ end of both transcripts. *q value < 0.05, ***q value < 0.001.