

RESEARCH

Open Access



An alternative splicing signature in human Crohn's disease

Daowei Li¹, Yuanzi Liang¹, Jia Lu¹ and Yue Tan^{2*}

Abstract

Background: Although hundreds of risk loci for Crohn's disease (CD) have been identified, the underlying pathogenesis of CD remains unclear. Recently, evidence has shown that aberrant gene expression in colon tissues of CD patients is associated with the progression of CD. We reasoned that post-transcriptional regulation, especially alternative splicing (AS), may also play important roles in the pathogenesis of CD.

Methods: We re-analyzed public mRNA-seq data from the NCBI GEO dataset (GSE66207) and identified approximately 3000 unique AS events in CD patients compared to healthy controls.

Results: "Lysine degradation" and "Sphingolipid metabolism" were the two most enriched AS events in CD patients. In a validation study, we also sequenced eight subjects and demonstrated that key genes that were previously linked to CD, such as *IRF1* and *STAT3*, also had significant AS events in CD.

Conclusion: Our study provided a landscape of AS events in CD, especially as the first study focused on a Chinese cohort. Our data suggest that dysregulation of AS may be a new mechanism that contributes to the pathogenesis of CD.

Keywords: Alternative splicing, Crohn's disease, Post-transcriptional regulation

Background

Crohn's disease (CD), an inflammatory bowel disease (IBD), is thought to be caused by dysregulated mucosal immune responses to the gut flora in genetically susceptible patients [1]. To date, more than 200 risk loci for IBD have been identified, which are also enriched in multi-genic regulatory modules [2, 3]. Although genome-wide association studies and subsequent meta-analyses could explain the underlying pathophysiology of CD to some extent, the pathogenesis of CD remains unclear. Recent studies have shown that dysregulated gene expression in the gut mucosa of CD patients may determine the initiation and progression of CD [4]. A subsequent study has also showed that there is a large difference in the

gene expression of inflamed and non-inflamed intestinal mucosa from CD patients when compared the gene expression of healthy mucosa from controls [5]. Moreover, a study that focused on integrated analysis of micro-RNAs (miRNA) and mRNA expression have demonstrated that miRNAs also play critical roles in regulating gene expression in CD patients [6]. Some of these circulating miRNAs are also thought to be potential prognostic markers in CD patients [7]. These data suggest that transcriptional and post-transcriptional regulation may play important roles in determining the etiology of CD.

Protein diversity determines the complexity of the eukaryotic cellular processes. Alternative splicing (AS) of the precursor mRNA is one of the essential mechanisms for increasing protein diversity and regulating intricate protein-RNA networks [8, 9]. Nearly 95% of the total human multi-exon genes are involved in AS events [10]. Evidence indicate that AS plays crucial roles in not

*Correspondence: tanyue2006@163.com

² Department of Gastroenterology, Shengjing Hospital of China Medical University, 39 Huaxiang Road, Tiexi District, Shenyang 110022, China
Full list of author information is available at the end of the article



© The Author(s) 2021. **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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

only oncogenic processes, including proliferation, apoptosis, hypoxia, angiogenesis, immune escape, and metastasis [11, 12], but also in basic developmental processes and tissue identity [13]. A pioneering study, which have focused on the AS of pre-mRNA in IBDs, have identified 47 splicing factors and 33 intron retention events that are dysregulated in the mucosal tissue of patients [14]. However, due to the lack of next-generation sequencing technology, only 149 splicing factors and 145 intron retention events were screened in that pioneering study.

We reasoned that dysregulated AS could contribute to the pathogenesis of CD. Thus, we obtained public mRNA-seq data from the NCBI GEO dataset (GSE66207) [6], which included 33 colon tissue mRNA-seq data (20 CD and 13 healthy controls). Using Mixture-of-Isoforms (MISO) and related AS event analysis software [15], we identified approximately 3000 significant AS events in CD patients compared in controls. A total of 117 biological pathways were significantly enriched for these AS events, some of which were highly related to inflammation-related responses. Interferon regulatory factor 1 (*IRF1*), signal transducer and activator of transcription 3 (*STAT3*), and monoamine oxidase A (*MAOA*) were identified as key genes involved in these AS events. In a validation study, we also identified that 64% of genes of significant AS events that overlapped with genes present in results from the public dataset, indicating a comprehensive role of AS in CD. We believe our results may shed light on the mechanism of post-transcriptional regulation of CD progression.

Methods

Sample collection

A total of four patients with CD from the Shengjing Hospital of China Medical University, Shenyang, China, were

recruited between June 2020 and September 2020 and served as the experimental group. All eligible patients had an established diagnosis of CD based on endoscopic and histologic assessments. Colonic biopsy specimens were taken from the rectum, ulcer margin of the sigmoid colon, and inflamed portions. Four patients with normal distal colon confirmed by surgical pathology served as the control group. The study was approved by the institutional review board of the Shengjing Hospital of China Medical University, and informed consent was obtained from each patient.

Data preparation

RNA-seq data were obtained from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) dataset (GSE66207) [6]. Sequencing was performed via ILLUMINA (Illumina HiSeq 2500) mean run: 26.98 M spots. The mean size for the sequencing file includes 2.58G bases according to the SRX886282 project detail. This dataset included 20 Crohn's disease patients and 13 normal colon tissues' raw sequence data and mRNA expression data (patient information is included in Table 1). Fastq files were aligned on the human Hg19 genome using STAR-2.7.1a [16] (read length = 150, average read depth = 50). The indexed.bam files were generated by Samtools (1.10) [17] for the next step of the analysis. Raw files for validation of the experiment were prepared using the same method.

Identification of significant AS events

AS events were identified and quantified using MISO version 0.5.4 [15] for 33 samples. A read length of 48 was applied, and other parameters were kept as default in MISO. The level of AS events was defined as the percent spliced in (PSI). To obtain differentially expressed

Table 1 Clinical Characteristics of the Patients (A total of 4 patients with CD were included between June 2020 and September 2020)

	Patient1	Patient2	Patient3	Patient4
Age (years)	16	22	30	23
Gender	Male	Female	Female	Male
Course of disease (years)	1.5	0.5	4	3
Age at diagnosis	15	22	29	21
Disease phenotype	B3	B1	B1p	B2p
Disease location	L3	L3	L3	L2
CDAI	160	170	410	240
Medication	Anti-TNF-a therapy	5-ASA and Anti-TNF-a therapy	Anti-TNF-a therapy	Anti-TNF-a therapy and corticosteroids

The clinical characteristics of all patients are shown in Table 1. There were no significant differences with regard to age and sex between the experimental and control groups ($P > 0.05$).

Locations: L1 = terminal ileum; L2 = colon; L3 = ileocolon. Behaviors: B1 = nonstricturing; B2 = stricturing; B3 = penetrating; P = perianal disease

ASA, Aminosalicic acid

AS events between Crohn's disease patients and normal samples, we applied the Wilcoxon test on all AS events with at least three patients and one normal sample. Significant events were defined as those with P values < 0.05 . The same method was applied in the transferability study. The AS events functional impact analysis was performed as significant events imported into the ASpedia database (<http://combio.snu.ac.kr/aspedia/>). Inputs of ASpedia analysis were converted from the MISO results.

Pathway analysis

Pathway analysis of the splicing and data was performed using Enrichr [18, 19] with the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway library [20]. Only pathways with a P value < 0.05 were considered as related pathways. WebGestalt [21], SUMER [22] and Cytoscape (3.8.0) [23] were applied to the pathway analysis of the transferability study section. Overrepresentation analysis and gene ontology biological processes were applied as enrichment methods and functional databases, respectively and were conducted using WebGestalt. The enriched category with gene sizes less than 5-kb and a false discovery rate above 0.05 was removed from the WebGestalt results. Results of different AS types were input into SUMER to obtain the pathway network. Finally, Cytoscape (3.8.0) was applied to modify the color and text size of the network from SUMER.

Visualization

A Sankey diagram was generated using Plotly in Python. UpSet plots were generated by ComplexHeatmap [24] in R (4.1.0). Sashimi plots were generated using Integrative Genomics Viewer (version 2.8.3) [25]. To avoid random error, we randomly selected and merged RNA-seq alignment data from three patients with Crohn's disease and three normal samples. Other graphs are plotted using the ggplot in R.

RNA-seq expression analysis

Pair-ended RNA-seq libraries for the transferability study were prepared using the NEBNext Ultra RNA with a Poly-A selection kit and were sequenced on an Illumina Hi-Seq 4000 (Genergy, Shanghai, China). Kallisto software was subsequently used to quantify RNA-seq raw counts based on a pseudoalignment algorithm [26] (read length = 200, default; mean read depth = 50). Differential gene expression was determined with \log_2 fold-change > 1.5 and $P < 0.05$, genes with > 1 count per million. Any gene with a P value greater than the false discovery rate, after Benjamini–Hochberg correction for multiple testing, was deemed significantly differentially expressed under the test conditions as compared to the controls.

Quantitative real-time PCR (RT-qPCR) analysis

Colon tissues from four control and four CD patients were used for qPCR assays. Total RNA was extracted from cells using TRIzol reagent (BioTeke, Beijing, China) and then reverse-transcribed using the first-strand cDNA synthesis kit (TSK302S, RT6 cDNA Synthesis Kit Ver 2). Reverse-transcribed products were used as templates for qPCR using the $2 \times T5$ Fast qPCR Mix (SYBR Green I). The primers used for the different target genes were as follows: *STAT3* (exon1)-F: GGCGAGGATTGGCTGAAGGG; *STAT3* (exon1)-R: CAGGCCGAAGGGCCTCTC; *STAT3* (lengthened 3'UTR)-F: tttctggaagttaaagtagatagacga; *STAT3* (lengthened 3'UTR)-R: ggccactgcattcaattcc; *IRF1*(exon1)- F: GGCAGAGCTCGCCACTCCTTAGTC; *IRF1*(exon1)-R: AGGCAGAGGTTGCCGGGTT; *IRF1* (lengthened 3'UTR)-F: acatgtggctagtgcaccagt; *IRF1* (lengthened 3'UTR)-R: cacagatgctgtcctcaaaa.

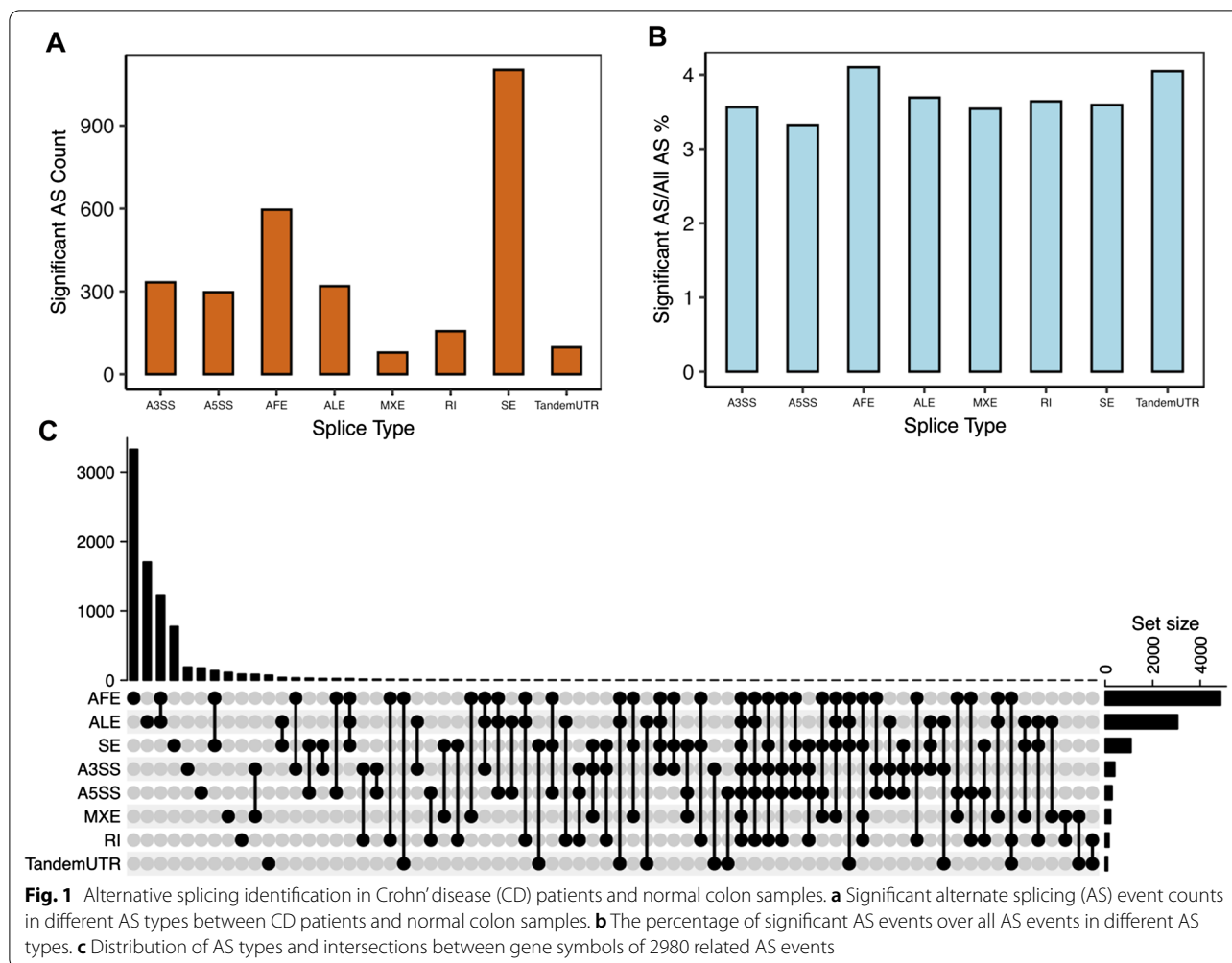
Results

Splicing types

In the study of identifying AS events in the data from GSE66207, we found eight types of AS events, including alternative 3' splice site (A3SS), alternative 5' splice site, alternative first exon (AFE), alternative last exon, mutually exclusive exons (MXE), retained introns (RI), skipped exon (SE), and tandem 3' UTR. A total of 2980 significant AS events were discovered in 20 patients with Crohn's disease (Fig. 1a) (Additional file 1: Table S1). Among these types, SE had the most significant AS events, but the percentage of the total number of SE events was low. A total of 4.1% of AFE events and 4% of tandem 3'UTR events were recognized as significant events, indicating a higher level compared to other events (Fig. 1b). The distribution and intersection of gene symbols among the eight event types for 2980 significantly different AS events are shown in Fig. 1c.

Pathway analysis

AS events related to Crohn's disease contributed to 117 different biological pathways (KEGG pathway) (Fig. 2). Among these significantly enriched pathways, "Lysine degradation," "Sphingolipid metabolism," and "Proteoglycans in cancer" were the most enriched terms, with P values of 3×10^{-5} , 3.6×10^{-4} , and 3.6×10^{-4} , respectively (Additional file 2: Table S2). Most of the pathways were associated with only one splicing type. However, "Human cytomegalovirus infection" and "Lysin degradation" pathways were affected by the four splicing types. We then combined related AS event pathway analysis results with RNA-seq expression results. We found that eight inflammation-related genes were regulated by CD-related AS events. Among these eight genes, *IRF1*, *STAT3*,



and *MAOA* were identified to have the differentially expressed AS events with the lowest *P* values of the Wilcoxon test and the most significant differential expression level between the Crohn's disease and normal group (Fig. 3a–c). Three genes were observed with SE splicing events (the first exon was skipped; red box in Fig. 3). *STAT3* has also been observed in AFE splicing events. According to the pathway analysis, *IRF1* and *STAT3* are involved in prolactin signaling pathways, which have a multitude of effects related to immunoregulation and protection [27]. Moreover, *MAOA* is related to tryptophan metabolism, which is also highly associated with inflammation, stress response, and microbiome homeostasis [28]. All three genes in the different sample groups showed significantly different RNA expression levels.

Principal component analysis (PCA) analysis

We summarized 2980 related AS events that occurred across all 33 samples with PSI values generated by MISO. To characterize the AS events between the disease and

normal samples, we performed a PCA on the MISO results of 2980 related AS events. PC1–PC3 accounted for 33% of the variance (Fig. 4). Biological differences between CD patients and control group were captured by the first, second, and third principal components.

Splicing types in the transferability study

To validate AS events that we identified in the GSE66207 datasets, we collected samples and performed RNA-seq on four colon tissues from patients with CD and four controls from the Shengjing Hospital of China Medical University. Between June 2020 and September 2020, a total of four patients diagnosed on the basis of the standard clinical, endoscopic, and histological criteria of CD at the Gastroenterology Department of the Shengjing Hospital of China Medical University were included and served as the validation group. Biopsies were performed during endoscopy. Patients with infectious colitis, indeterminate colitis, Behcet's disease, intestinal tuberculosis, and colorectal cancer were excluded. Colon biopsies

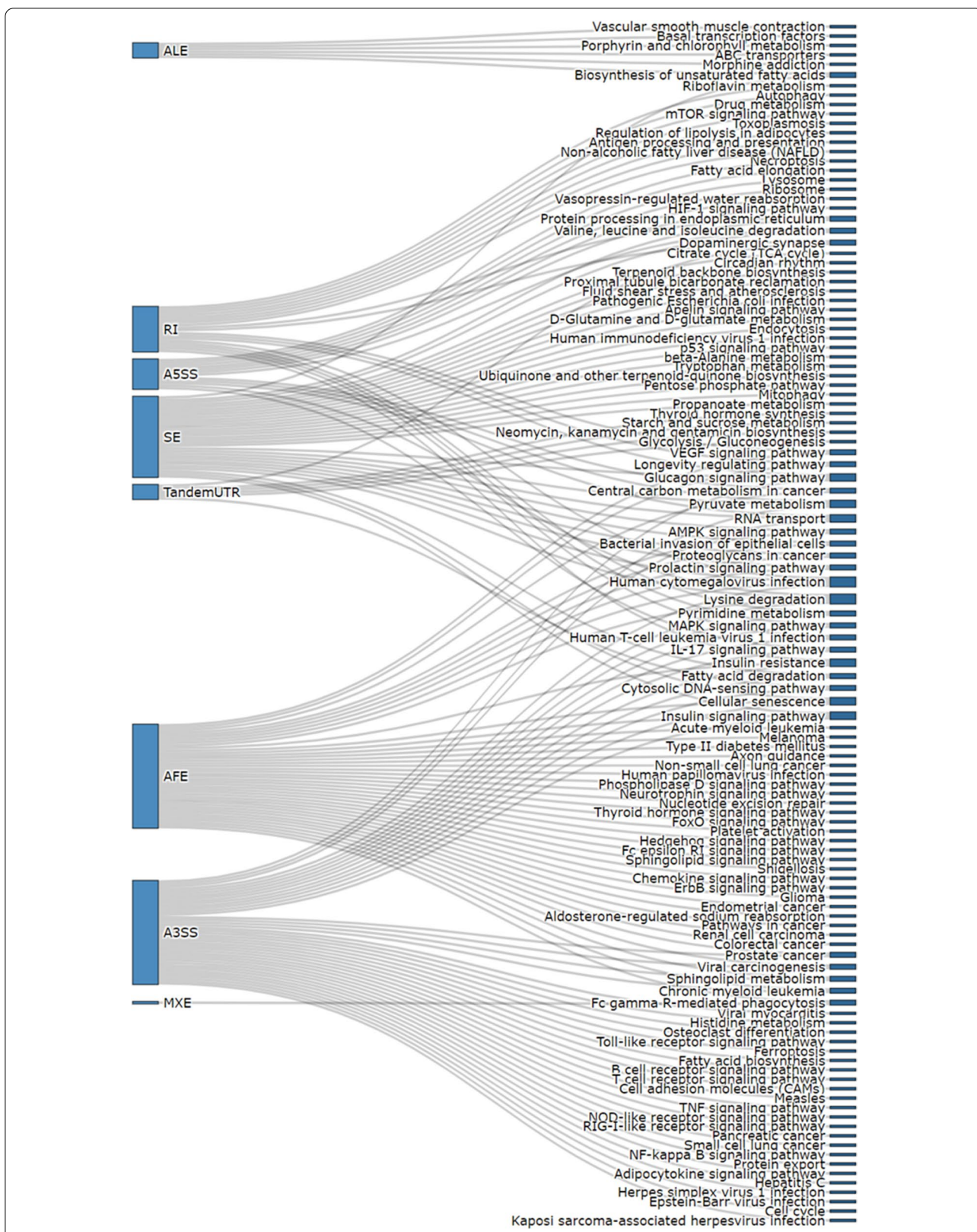
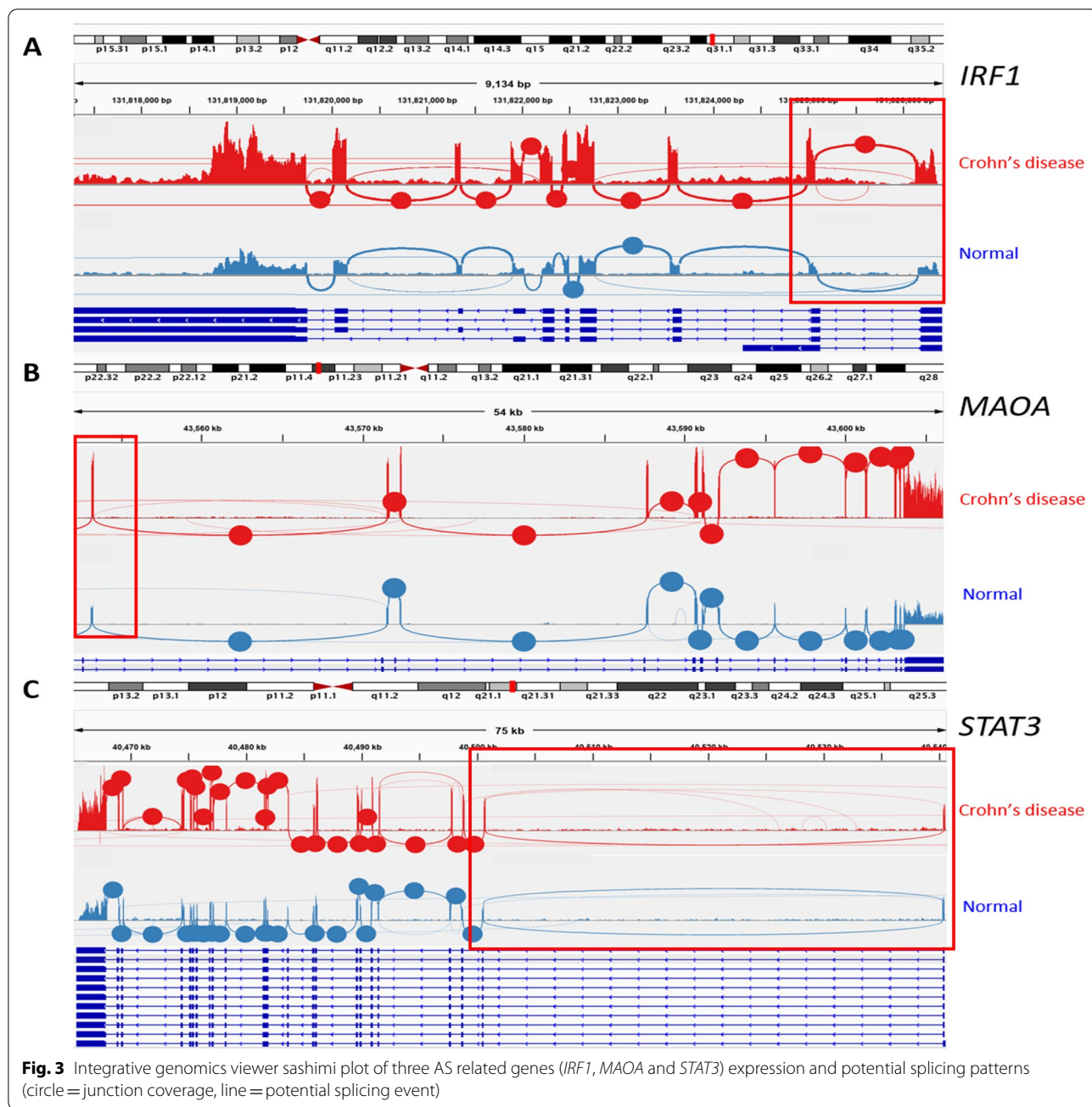


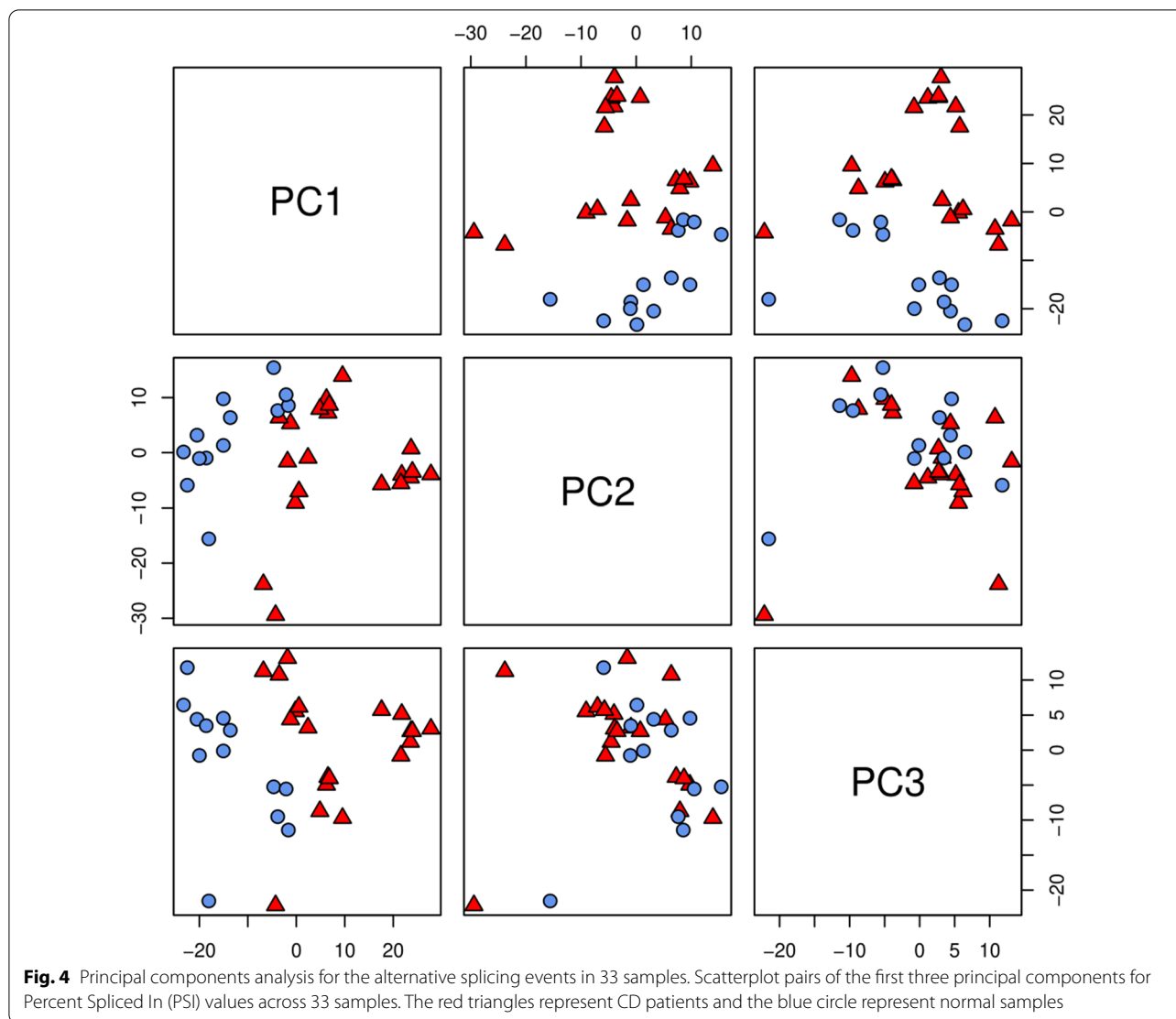
Fig. 2 Pathway enrichment analysis. Biological pathways associated with the related alternate splicing events in different splice types



obtained from healthy volunteers during endoscopy served as the control group. Clinical disease activity was assessed by measuring the Crohn's disease activity index for CD by a physician who was unaware of the patients' colonoscopy results. This study was approved by the institutional review board of the Shengjing Hospital of China Medical University, and informed consent was obtained from each patient (Table 1).

Using the same analysis pipeline, we identified 1715 significant AS events (Additional file 3: Table S3) in the

validation analysis. Interestingly, these results demonstrated that SE and AFE were still the two most common events, and MXE was the least common, which were similar to results from the public dataset (Fig. 5a). However, the percentage over the total number of all event types was relatively lower than that in the previous study (Fig. 5b). In 1715 significant AS events, AFE and alternative last exon were also identified across multiple genes, which had the most intersections (Fig. 5c).

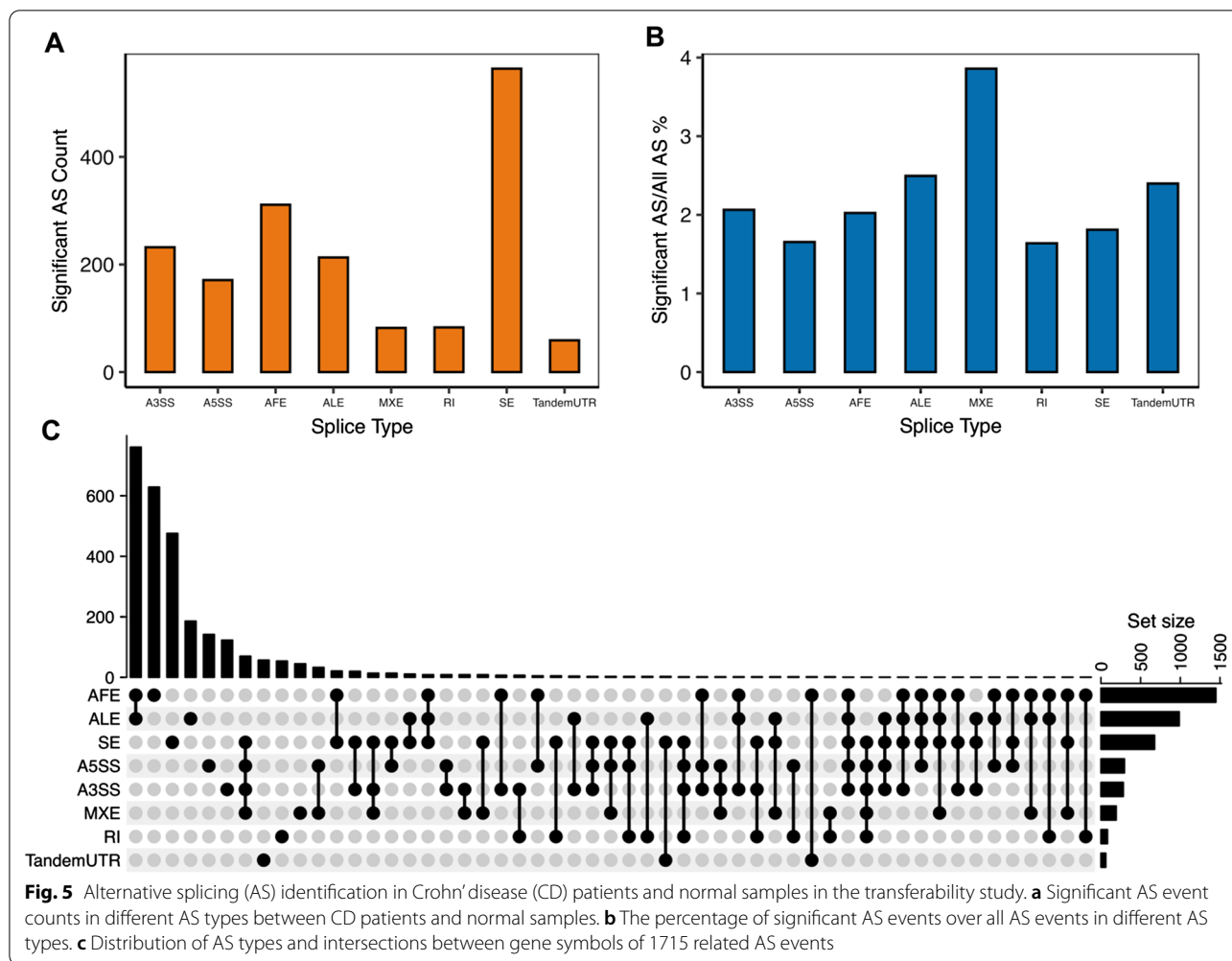


Biological process pathway analysis in the transferability study

We performed biological process pathway analysis in a transferability study based on the gene lists of 1731 significant AS events. Since there are less than 200 related genes in MXE, RI, and tandem 3' UTR event types, we did not identify significant biological processes of these three types of events. In the other five AS-type events, we found that several biological process networks were highly enriched using WebGestalt analysis. Among these significantly enriched networks, “defense response,” “cell killing,” “response to cytokine,” and their related modules were the most enriched clusters (Fig. 6). These results suggest that the potential dysregulation of AS may play important roles in host-microbiome interactions in patients with CD.

Comparison between the data from public datasets and our transferability study

Finally, we compared results from the analysis of the public dataset and results from the analysis of our own cohort. In terms of genes that had significant AS events, we identified 64% genes that were identified to include significant AS events in the transferability study and overlapped with genes from the results we found in the public dataset (Fig. 7a). Interestingly, *IRF1* and *STAT3* were still identified to have significant tandem 3'UTR events and showed significantly altered expression at the mRNA level in the patients (Fig. 7b). To validate these findings, we performed RT-qPCR assays on eight samples (4 CD vs. 4 control). We initially validated the upregulation of *STAT3* and *IRF1* expression in the CD samples (Fig. 7c), which is consistent with previous findings [29, 30]. We also performed RT-qPCR assays specifically targeting

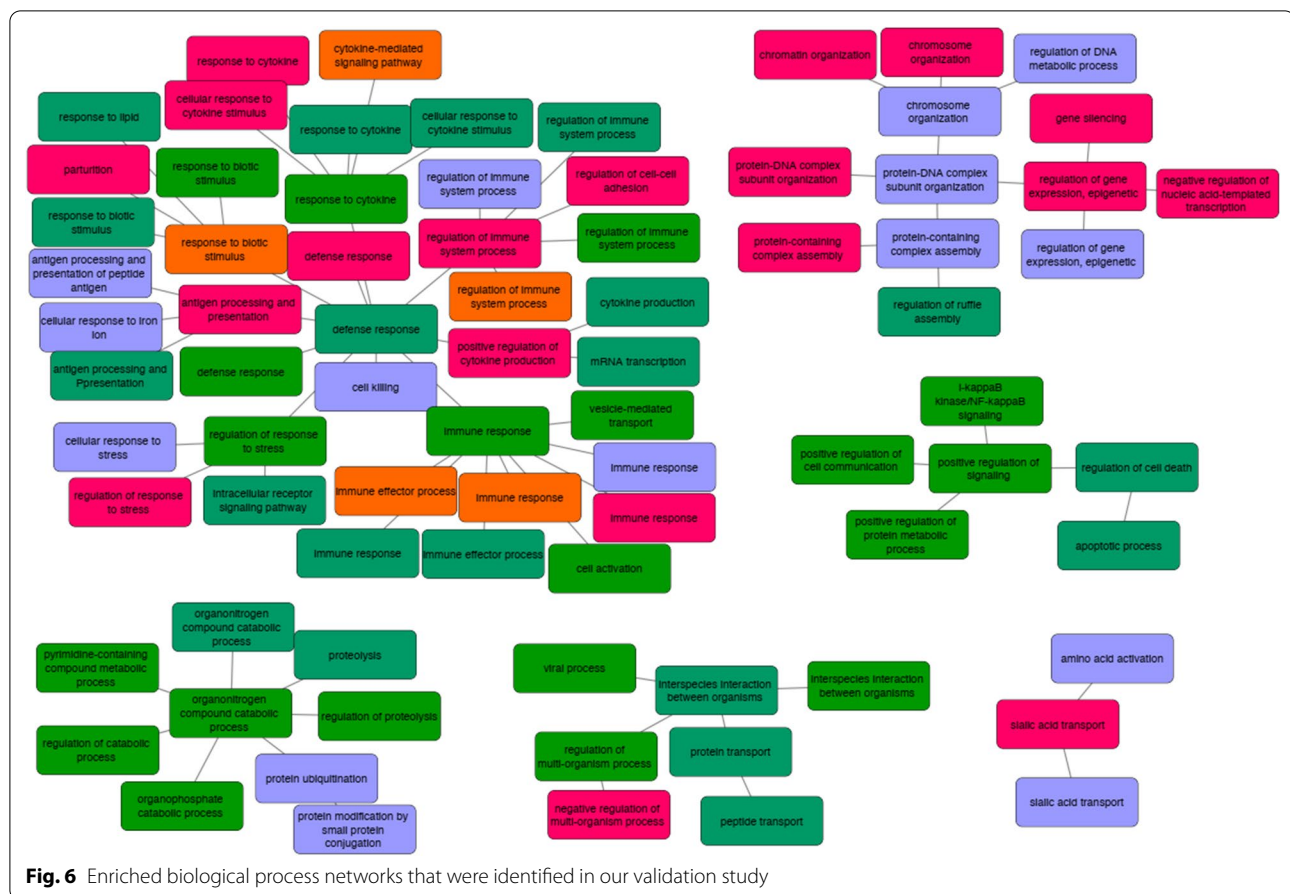


different uses of the 3'UTR regions of these two genes. Interestingly, the expression levels of the lengthened 3'UTR were significantly increased in the CD samples, compared to those in the control (Fig. 7d, e). Even after normalization to the overall mRNA expression (Exon 1), the lengthened 3'UTR expression was still profoundly upregulated in the CD samples (Fig. 7f). These results further validated our findings of significant AS in crucial genes related to CD, indicating the potential role of AS in regulating CD-related gene expression, thus contributing to the etiology of the disease. Finally, we collected 404 splicing factors from a large splicing factor database for expression analysis [31]. We excluded 34 factors that were barely expressed in the RNA-seq data (fragments per kilobase of transcript per million mapped reads (FPKM) < 1). As a result, we identified five splicing factor genes that were significantly differentially expressed (P value < 0.05) in both datasets (Table 2). We also noticed that these splicing factors had the same trend in terms of expression level alteration. We speculate that these

splicing factors may be important in regulating AS events in patients with CD. In conclusion, these results suggest a strong similarity between our transferability study and the public dataset. We also speculate the difference from these analyses may be due to different genetic backgrounds of these samples, as the first cohort was from the University of North Carolina (presumably a cohort of mixed ethnicity) and the validation cohort consisted of only Han Chinese. The average age of the first cohort was 45 years old, and the average age of our cohort was 21 years old, indicating a significant difference in age between the two populations. However, the sex ratio between the two cohorts was similar (1:1 vs. 0.7:1). These factors could also potentially affect AS results.

Discussion

Dysregulation of pre-mRNA splicing can cause human disease, and an increasing number of studies have shown that targeting AS could lead to the development of novel therapeutics [32]. To date, accumulated data have already



highlighted crucial roles of AS in many human diseases, including cancer, neurodegenerative disorders, and immune and infectious diseases [33–35]. Herein, we present a comprehensive analysis of AS events in CD using a well-established public dataset and a validation dataset from sequencing a Chinese cohort. In our study, we found that SE was the most significantly enriched AS event in patients with CD. This suggests that transcriptome of CD patients is much more complex and chaotic compared to that of healthy controls, due to a more sophisticated mRNA isoform system. The PCA plot also demonstrated that even if only different AS events were used as principal components, CD and control groups were still distinguishable, suggesting that the mRNA AS event itself, other than the expression level, could be a solid comprehensive signature for CD patients. All these results proved our hypothesis that dysregulated AS events may contribute to the etiology of CD.

Among AS events related to significantly enriched pathways, the lysine degradation pathway was the most enriched pathway. *EZH1* and *EZH2* are two genes involved in this pathway, which have been previously linked to the progression of IBD [36, 37]. “Sphingolipids

metabolism” is the most significantly enriched pathway in A3SS events. Interestingly, sphingolipids have recently been recognized as mediators of inflammation and are potential therapeutic targets for IBDs [38]. Our results suggest that the metabolism of sphingolipids in CD patients may differ from that in healthy individuals due to AS, resulting in a dysregulated sphingolipid pool. Moreover, the three most dysregulated genes in inflammatory response-related pathways were previously linked to IBD. *IRF1*, as an interferon regulatory factor, has been correlated with the CD activity index and CD endoscopic index of severity [39]. In a cohort study, *IRF1* showed a 72% increase in gene expression in patients with CD compared to that in controls [40]. If *STAT3*, as a central component in immune response signal transduction, has a gain-of-function mutation, it will cause multi-organ early auto-immune diseases [41]. Notably, *STAT3* has also been indicated as a crucial target for treating IBDs, although activation of *STAT3* is likely to occur in both innate and acquired cell types [29]. *MAOA*, as a monoamine oxidase, is not directly associated with IBDs and is highly related to the induction of oxidative stress in obese people suffering from chronic inflammation [42].

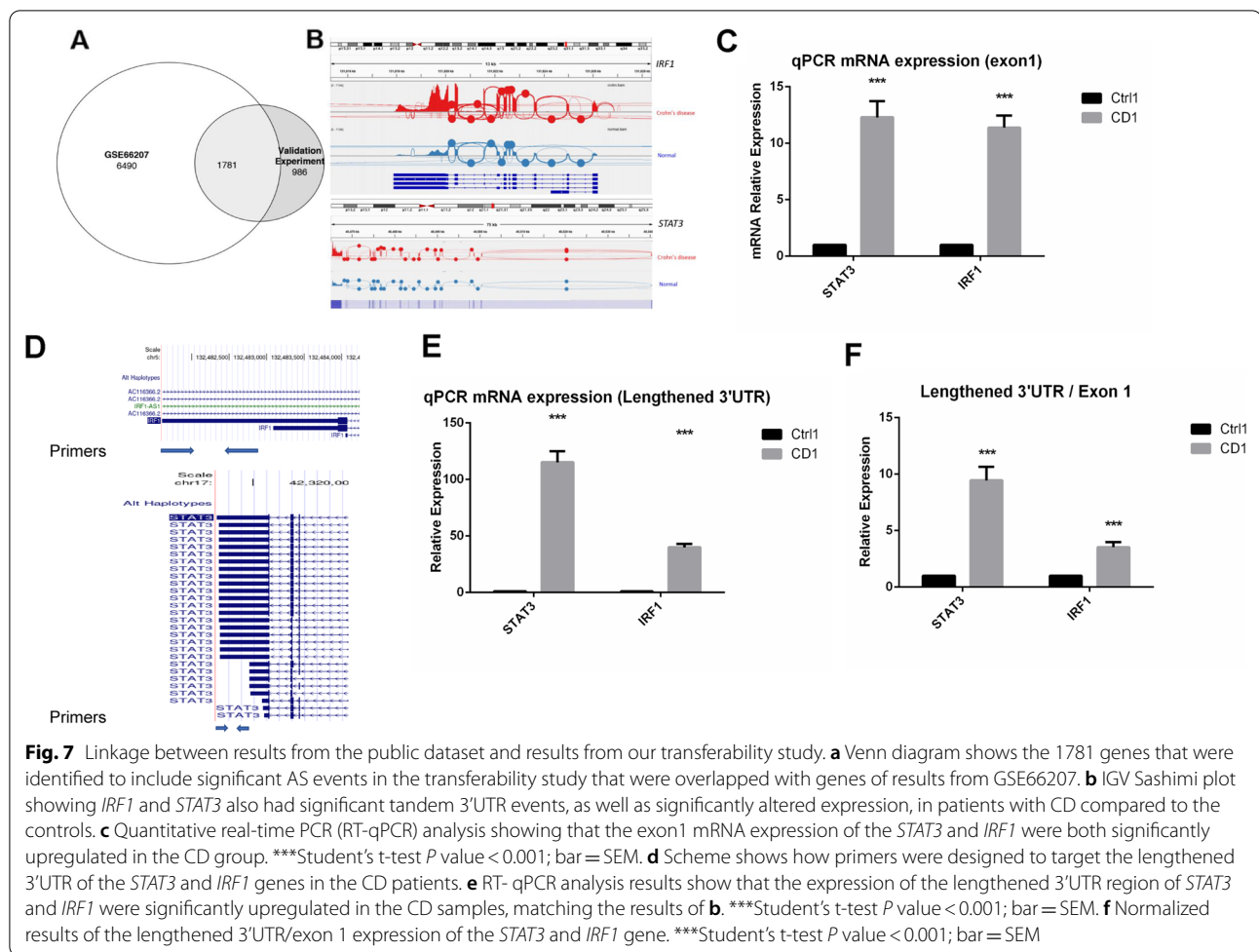


Table 2 Transferability study RNA-seq splicing factor raw count fold change of 8 samples

Gene symbol	Description	GSE66207-Log2Foldchange CD versus control	Validation experiment-Log2Foldchange CD versus control
HSPA1A	Heat shock protein family A (Hsp70) Member 1A	2.4950086	1.73285
RNF213	Ring finger protein 213	1.5825733	3.760895
MOV10	Mov10RISC complex RNA helicase	1.5125286	1.304929
HSPA5	Heat shock protein family A (Hsp70) Member 5	1.3313641	3.431851
LSM10	LSM10, U7 small nuclear RNA associated	1.2935081	1.631783

Recently, dysregulation of *MAOA* has also been linked to specific microbe alterations [43]. In conclusion, both top AS-related significantly enriched pathways and significantly dysregulated genes were highly related to IBD and CD. Our results provide new insights into the role of post-transcriptional regulation in IBD by regulating these key pathways and genes.

Although our study provided a comprehensive AS analysis of colon tissues from CD patients and also

identified nearly 3000 unique AS events in CD patients, our study still has certain limitations. First, although we validated the AS events of *STAT3* and *IRF1* genes, the experimental validation (protein level) of discovered AS events in many other genes is still lacking. In fact, recent updates on AS effect prediction algorithm have allowed us to predict the function alteration of the targeted proteins. 68 AS events were predicted to alter the downstream protein function or expression by using

ASpedia software. These affected AS events include 10 nonsense-mediated decay (NMD) related events, 22 protein domain related alterations, 57 posttranslational modification related events (Additional file 4: Table S4). However, whether these AS events affect protein diversity or not still requires experimental validation. Validation studies should examine protein levels of these AS-regulated genes and detect downstream effects on certain signal transduction pathways. Finally, even if we have detected several splicing factors that were dysregulated in CD patients, an RNA immunoprecipitation sequencing or cross-linking immunoprecipitation study should be performed to further detect responsible splicing factors and RNA-binding proteins for CD-related AS events. Nonetheless, our study demonstrated a new potential pathophysiology of CDs, which is regulated by AS. Our study also presented the first AS landscape in patients with CD and suggested that drugs that target AS-related genes and splicing factors should be considered for further screening of patients with CD.

Conclusion

Our study presented a landscape of AS in CD and provided the first AS-related transcriptional analysis in a Chinese cohort. Integrated analysis of the two datasets revealed that AS may play a crucial role in determining the pathogenesis of CD.

Abbreviations

CD: Crohn's disease; AS: Alternative splicing; IBD: Inflammatory bowel disease; miRNAs: Micro-RNA; KEGG: Kyoto encyclopedia of genes and genomes; FDR: False discovery rate; 3' UTR: Three prime untranslated region; *IRF1*: Inflammation-related responses. Interferon regulatory factor 1; *STAT3*: Signal transducer and activator of transcription 3; *MAOA*: Monoamine oxidase A.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12876-021-02001-2>.

Additional file 1: Table S1. Table for discovered significant AS events in the 20 CD patients from public dataset

Additional file 2: Table S2. Table for KEGG pathway enrichment of AS events in CD patients

Additional file 3: Table S3. Table for discovered significant AS events in our cohort

Additional file 4: Table S4. Table for AS related effects on protein function/diversity for the public dataset analysis

Acknowledgements

We appreciate the assistance from Intanx Life (Shanghai) Co. Ltd. in data processing and consulting.

Authors' contributions

Y. Tan designed research and collected the patients' samples; D. Li, Y. Liang and J. Lu analyzed data; D. Li and Y. Tan wrote the paper. All authors read and approved the final manuscript.

Funding

This work was supported partly by the Science and Technology Program of Liaoning Province (No. 2019-BS-140).

Availability of data and materials

RNA-seq data for 20 CD AS analysis was obtained from NCBI GEO dataset (GSE66207). "The datasets generated and/or analysed during the current study are available in the NCBI_GEO repository, [PERSISTENT WEB LINK OR ACCESSION NUMBER TO DATASETS]".

Declarations

Ethics approval and consent to participate

The study was approved by the institutional review board of Shengjing Hospital of China Medical University, and informed consent was obtained from each patient (2020Yc002). All the experiment protocol for involving human data was in accordance with the guidelines of national/international/institutional or Declaration of Helsinki in the manuscript.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Radiology, The People's Hospital of China Medical University and The People's Hospital of Liaoning Province, No. 33, Wenyi Road, Shenhe District, Shenyang 110016, China. ²Department of Gastroenterology, Shengjing Hospital of China Medical University, 39 Huaxiang Road, Tiexi District, Shenyang 110022, China.

Received: 5 May 2021 Accepted: 26 October 2021

Published online: 08 November 2021

References

1. Simmons A. Genes, viruses and microbes. *Nature*. 2010;466(7307):699–700.
2. Momozawa Y, Dmitrieva J, Théâtre E, Deffontaine V, Rahmouni S, Charlotiaux B, Crins F, Docampo E, Elansary M, Gori A-S, et al. IBD risk loci are enriched in multigenic regulatory modules encompassing putative causative genes. *Nat Commun*. 2018;9(1):2427.
3. Hong SN, Park C, Park SJ, Lee CK, Ye BD, Kim YS, Lee S, Chae J, Kim J-I, Kim Y-H. Deep resequencing of 131 Crohn's disease associated genes in pooled DNA confirmed three reported variants and identified eight novel variants. *Gut*. 2016;65(5):788–96.
4. van Beelen Granlund A, Flatberg A, Østvik AE, Drozdov I, Gustafsson BI, Kidd M, Beisvag V, Torp SH, Waldum HL, Martinsen TC. Whole genome gene expression meta-analysis of inflammatory bowel disease colon mucosa demonstrates lack of major differences between Crohn's disease and ulcerative colitis. *PLoS ONE*. 2013;8(2):1–13.
5. Hong SN, Joung J-G, Bae JS, Lee CS, Koo JS, Park SJ, Im JP, Kim YS, Kim JW, Park WY, et al. RNA-seq reveals transcriptomic differences in inflamed and noninflamed intestinal mucosa of Crohn's disease patients compared with normal mucosa of healthy controls. *Inflamm Bowel Dis*. 2017;23(7):1098–108.
6. Peck BCE, Weiser M, Lee SE, Gipson GR, Iyer VB, Sartor RB, Herfarth HH, Long MD, Hansen JJ, Isaacs KL, et al. MicroRNAs classify different disease behavior phenotypes of Crohn's disease and may have prognostic utility. *Inflamm Bowel Dis*. 2015;21(9):2178–87.
7. Jensen MD, Andersen RF, Christensen H, Nathan T, Kjeldsen J, Madsen JS. Circulating microRNAs as biomarkers of adult Crohn's disease. *Eur J Gastroenterol Hepatol*. 2015;27(9):1038–44.

8. Wang Y, Liu J, Huang B, Xu YM, Li J, Huang LF, Lin J, Zhang J, Min QH, Yang WM. Mechanism of alternative splicing and its regulation. *Biomed Rep*. 2015;3(2):152–8.
9. Chen M, Manley JL. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat Rev Mol Cell Biol*. 2009;10(11):741–54.
10. Black DL. Mechanisms of alternative pre-messenger RNA splicing. *Ann Rev Biochem*. 2003;72(1):291–336.
11. David CJ, Manley JL. Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged. *Genes Dev*. 2010;24(21):2343–64.
12. Oltean S, Bates DO. Hallmarks of alternative splicing in cancer. *Oncogene*. 2014;33(46):5311–8.
13. Baralle FE, Giudice J. Alternative splicing as a regulator of development and tissue identity. *Nat Rev Mol Cell Biol*. 2017;18(7):437.
14. Häslér R, Kerick M, Mah N, Hulstschig C, Richter G, Bretz F, Sina C, Lehrach H, Nietfeld W, Schreiber S, et al. Alterations of pre-mRNA splicing in human inflammatory bowel disease. *Eur J Cell Biol*. 2011;90(6–7):603–11.
15. Katz Y, Wang ET, Airoidi EM, Burge CB. Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat Methods*. 2010;7(12):1009–15.
16. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15–21.
17. Li H, Handsaker B, Wysoker S, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–9.
18. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR, Ma'ayan A. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinf*. 2013;14(1):128.
19. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucl Acids Res*. 2016;44(W1):W90–7.
20. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res*. 2016;44(D1):D457–62.
21. Liao Y, Wang J, Jaehng EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res*. 2019;47(W1):W199–205.
22. Savage SR, Shi Z, Liao Y, Zhang B. Graph algorithms for condensing and consolidating gene set analysis results. *Mol Cell Proteomics*. 2019;18(8 suppl 1):S141–52.
23. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498–504.
24. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*. 2016;32(18):2847–9.
25. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform*. 2013;14(2):178–92.
26. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol*. 2016;34(5):525.
27. Ben-Jonathan N, Hugo ER, Brandebourg TD, LaPensee CR. Focus on prolactin as a metabolic hormone. *Trends Endocrinol Metab*. 2006;17(3):110–6.
28. Gostner JM, Geisler S, Stonig M, Mair L, Sperner-Unterwieser B, Fuchs D. Tryptophan Metabolism and Related Pathways in Psychoneuroimmunology: The Impact of Nutrition and Lifestyle. *Neuropsychobiology*. 2020;79(1):89–99.
29. Sugimoto K. Role of STAT3 in inflammatory bowel disease. *World J Gastroenterol*. 2008;14(33):5110.
30. Clavell M, Correa-Gracian H, Liu Z, Craver R, Brown R, Schmidt-Sommerfeld E, Udall J Jr, Delgado A, Mannick E. Detection of interferon regulatory factor-1 in lamina propria mononuclear cells in Crohn's disease. *J Pediatr Gastroenterol Nutr*. 2000;30(1):43–7.
31. Seiler M, Peng S, Agrawal AA, Palacino J, Teng T, Zhu P, Smith PG, Buonamici S, Yu L. Somatic Mutational Landscape of Splicing Factor Genes and Their Functional Consequences across 33 Cancer Types. *Cell Rep*. 2018;23(1):282–296.e284.
32. Zhao S. Alternative splicing, RNA-seq and drug discovery. *Drug Discovery Today*. 2019;24(6):1258–67.
33. Wang G-S, Cooper TA. Splicing in disease: disruption of the splicing code and the decoding machinery. *Nat Rev Genet*. 2007;8(10):749–61.
34. Scotti MM, Swanson MS. RNA mis-splicing in disease. *Nat Rev Genet*. 2016;17(1):19.
35. Kim HK, Pham MHC, Ko KS, Rhee BD, Han J. Alternative splicing isoforms in health and disease. *Pflügers Archiv-Eur J Physiol*. 2018;470(7):995–1016.
36. Fernandez-Becker NQ, Moss AC. In silico analysis of T-bet activity in peripheral blood mononuclear cells in patients with inflammatory bowel disease (IBD). *In Silico Biol*. 2009;9(5–6):355–63.
37. Lou X, Zhu H, Ning L, Li C, Li S, Du H, Zhou X, Xu G. EZH2 regulates intestinal inflammation and necroptosis through the JNK signaling pathway in intestinal epithelial cells. *Dig Dis Sci*. 2019;64(12):3518–27.
38. Sukocheva OA, Lukina E, McGowan E, Bishayee A. Sphingolipids as mediators of inflammation and novel therapeutic target in inflammatory bowel disease. *Adv Protein Chem Struct Biol*. 2020;120:123–58.
39. Tang R, Yang G, Zhang S, Wu C, Chen M. Opposite effects of interferon regulatory factor 1 and osteopontin on the apoptosis of epithelial cells induced by TNF- α in inflammatory bowel disease. *Inflamm Bowel Dis*. 2014;20(11):1950–61.
40. Huff CD, Witherspoon DJ, Zhang Y, Gatenbee C, Denson LA, Kugathasan S, Hakonarson H, Whiting A, Davis CT, Wu W, et al. Crohn's disease and genetic hitchhiking at IBD5. *Mol Biol Evol*. 2012;29(1):101–11.
41. Milner JD, Vogel TP, Forbes L, Ma CA, Stray-Pedersen A, Niemela JE, Lyons JJ, Engelhardt KR, Zhang Y, Topcagic N. Early-onset lymphoproliferation and autoimmunity caused by germline STAT3 gain-of-function mutations. *Blood*. 2015;125(4):591–9.
42. Sturza A, Olariu S, Ionică M, Duicu OM, Văduva AO, Boia E, Muntean DM, Popoiu CM. Monoamine oxidase is a source of oxidative stress in obese patients with chronic inflammation (1). *Can J Physiol Pharmacol*. 2019;97(9):844–9.
43. Xie Y, Wang C, Zhao D, Wang C, Li C. Dietary proteins regulate serotonin biosynthesis and catabolism by specific gut microbes. *J Agric Food Chem*. 2020;68(21):5880–90.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

